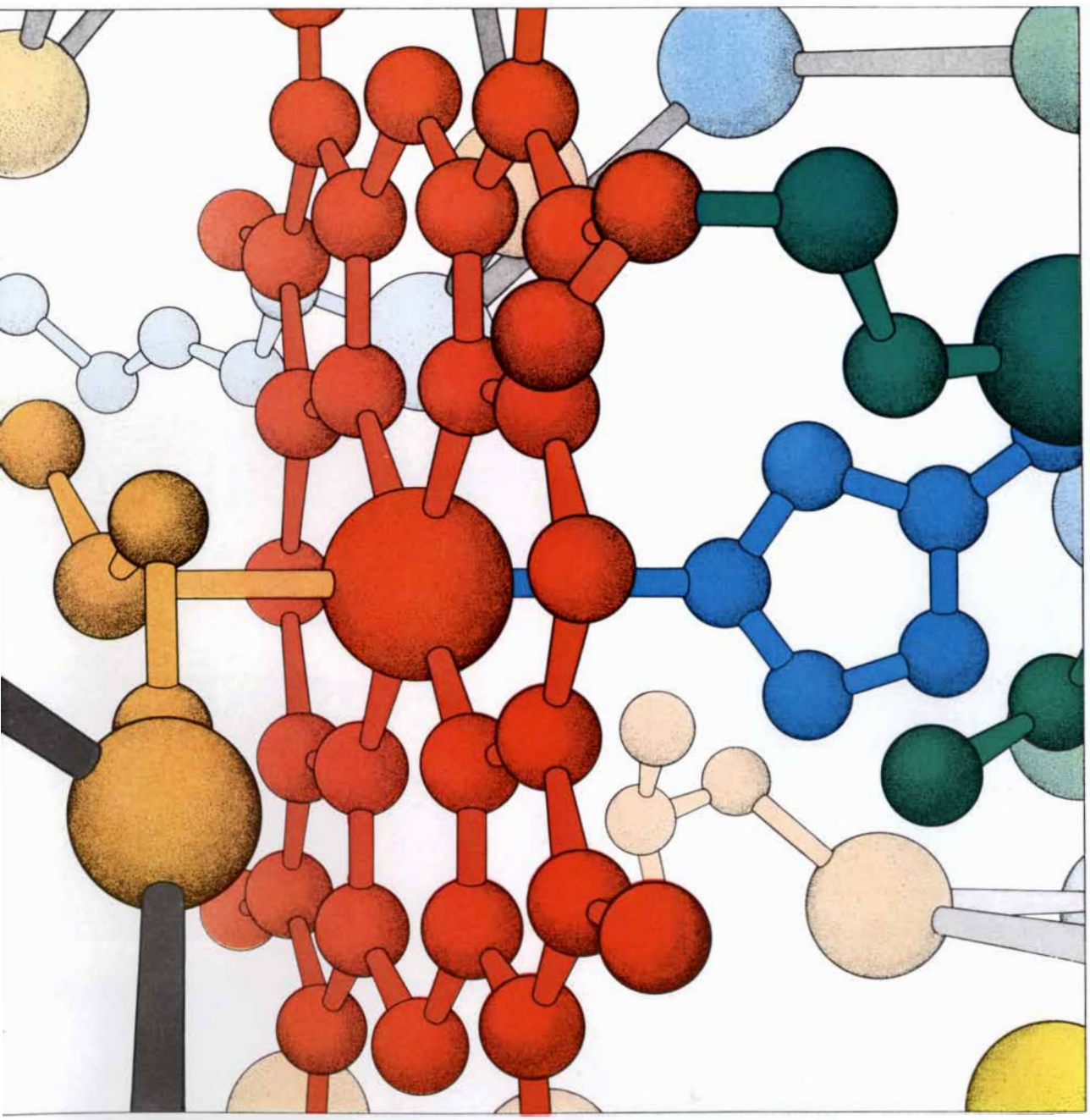
SCIENTIFIC AMERICAN



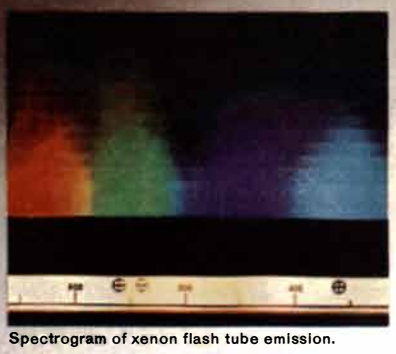
CYTOCHROME MOLECULE

ONE DOLLAR

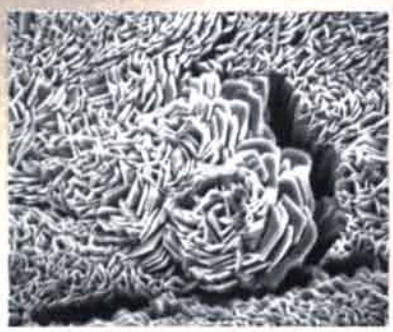
April 1972

Polaroid®

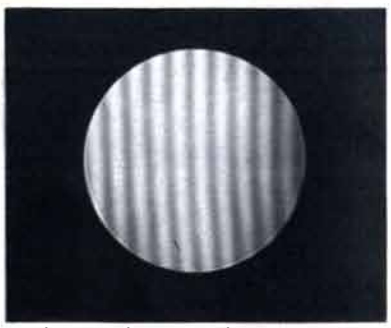
How one great idea leads to another.



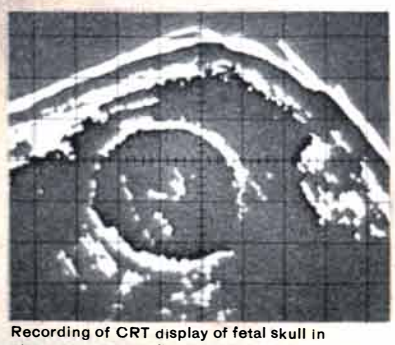
Spectrogram of xenon flash tube emission.



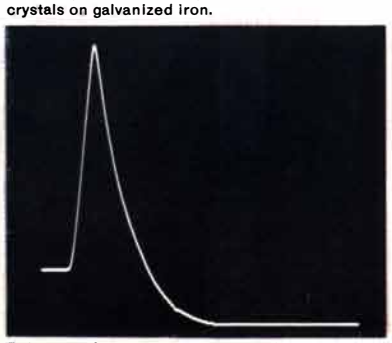
Scanning electron micrograph of zinc crystals on galvanized iron.



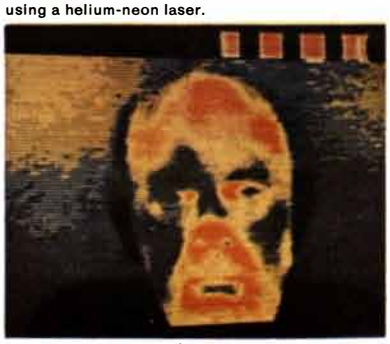
Interferogram of a mirror surface, using a helium-neon laser.



Recording of CRT display of fetal skull in utero made with reflected ultrasound waves.



Pulmonary function trace made with spirometer using fibre optics.



Thermograph showing blockage in carotid artery.

When you want a visual record of almost anything, Polaroid can help you get it in seconds.

When Polaroid succeeded in developing the process of instant photography, even we could not foresee its incredible variety of applications.

Today, equipment manufacturers have adapted Polaroid Land camera backs to more than 500 scientific, photographic and laboratory instruments for instant recording. And the number is growing.

As scientific advances proliferate, so does the need for a fast, convenient, accurate method of recording data. Even where original data are received via forms of energy other than visible light, more and more designers depend on Polaroid.

Our camera backs are available in two formats, 3¼ x 4¼ inch (either roll or pack) and 4 x 5 packets. And there are 13 different Polaroid self-developing films to choose from. Including high speed (10,000 ASA), high contrast, line and continuous tone transparency films, color, and one that produces a fully developed black-and-white negative with matching positive in seconds.

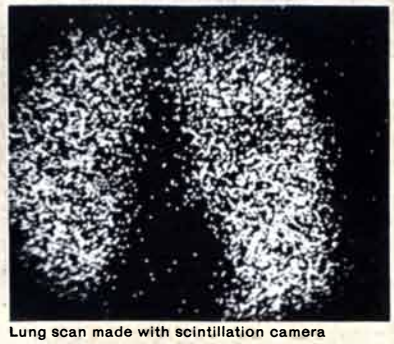
But Polaroid does more for original equipment manufacturers than sell camera backs and film. We have a wide range of experience, and are more than willing to share it with you. We can offer advice on making the right selections for your particular needs. And on how to adapt to your equipment in the most practicable way.

Perhaps you're working on a new development today. Or you might come up with one tomorrow. If making any kind of visual record is part of the plan, get in touch with us. Write to Polaroid Corporation, Dept. 57-235, 549 Technology Square, Cambridge, Massachusetts 02139.

Our ideas may give you an idea.

Polaroid Land Camera Backs

Radiograph of prisoner's shoe revealing narcotic injection kit.

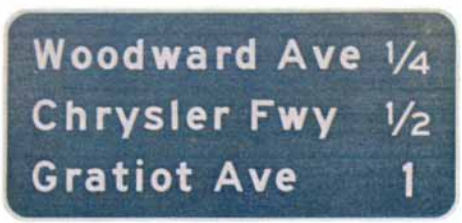


Lung scan made with scintillation camera using radioisotopes.

It doesn't matter who you are. You can't get a Chrysler Imperial that hasn't had a shakedown in Detroit traffic.

ARTHUR GODFREY

Nobody, but nobody gets an unused Imperial. Every Imperial that Chrysler builds gets somewhere between twelve and fifteen miles put on it before it ever gets delivered. If there are any problems, Chrysler wants to find them before you do.



Each car is checked on city streets and freeways, in all kinds of traffic conditions. The test drivers listen for wind noise and check for rattles. They even inspect *under* the car: Things like the fuel line and brake line. There's a check sheet with more than 200 items that gets filled out on each car. If it's not right, they don't want you to have it.

CHRYSLER
Plymouth

Coming through with the kind of car America wants.

That's their slogan this year, and I think they're living up to it.

I've seen the way they build the 1972 Chryslers. It's obvious, they want these cars to last longer and perform better than any car they've ever built before.

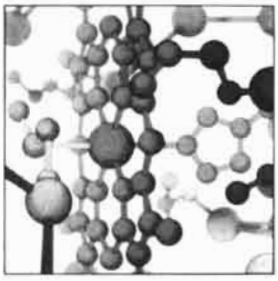
Your Chrysler-Plymouth dealer can let you take an Imperial out for a road test of your own. You'll see what I mean.

Established 1845 **SCIENTIFIC AMERICAN** April 1972 Volume 226 Number 4

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: $10 PER YEAR. EUROPEAN SUBSCRIPTION COPIES DISTRIBUTED C/O INTERNATIONAL DISTRIBUTORS, VLAGSTRAAT 128, BORGERHOUT 2200, BELGIUM.



THE COVER

The illustration on the cover depicts the central portion of the enzyme cytochrome *c*, a protein whose history can be traced back at least 1.2 billion years (see "The Structure and History of an Ancient Protein," page 58). Cytochrome *c* evolved as one of a group of proteins that act sequentially in the chemical process enabling cells to extract energy when food molecules are oxidized. The enzyme is present in all organisms whose cells contain a nucleus, from yeasts to man. The cover picture provides a close-up of the heme group (*red*), a rosette of carbon and nitrogen atoms surrounding a central atom of iron. The heme group is firmly held by the side chains of three of the 104 amino acids that are linked together to form the cytochrome *c* molecule of higher organisms. The side chains of methionine (*yellow*) and histidine (*blue*) form bonds with the central atom of iron; cysteine (*green*) is bonded to an atom of carbon at the edge of the heme group.

THE ILLUSTRATIONS

Cover illustration by Irving Geis

Page	Source
16–17	Jim Egleson
18–21	Graphic Presentation Services, Inc.
22	Jon Brenneis
24–26	Dan Todd
27–28	Physics International Company
29–31	Dan Todd
32	Denis Keefe, Lawrence Berkeley Laboratory, University of California
33	Dan Todd
35	Donn T. Bayard, University of Otago (*top*); Wilhelm G. Solheim II, University of Hawaii (*bottom*)
36	Alan D. Iselin
37	Chester F. Gorman (*top*); Lorelle M. Raboni (*bottom*)
38–41	Lorelle M. Raboni
42	Jon Brenneis
44–52	Jim Egleson
58–59	Irving Geis
60–61	Graphic Presentation Services, Inc. (*left*), Irving Geis (*right*)
62–64	Irving Geis
66–67	Richard E. Dickerson, California Institute of Technology
68	Graphic Presentation Services (*top*), Irving Geis (*bottom*)
77	Norman Geschwind, Harvard Medical School
78–83	Bunji Tagawa
85–91	Allen Beechel
92	Paul Weller
94–96	Tom Prentiss
97	Paul Weller
100–104	Enid Kotschnig
106–110	Roger Hayward

Working for National ought to be a snap.

After all, we're not one of the big boys in the business, right?

Wrong.

In fact, it's staggering to realize that we've actually got more new Chevys than anyone else in the world. Clean, shiny '72 Chevys. (Not to mention our other fine cars.)

That means that most every day upwards of fifty thousand businessmen and salesmen and vacationers and you name 'ems can count on us to put them on the road.

It's exhausting.

Because with National there's no fooling around.

Doesn't matter if we're renting you a car in Orinda, California or Boston, Massachusetts, there's only one way to do it. The right way.

Keeping customers waiting, that's a no-no. So those rental forms have to be filled out fast, fast, fast.

And all those Green Stamps we have to count out. It's not enough to be like everyone else and just smile and say thanks.

No, we have to smile and say thanks with S&H Green Stamps.*

And then there are all those reservations we have to keep track of. With business so good, you wouldn't think that National would get so uptight about, well, goofing a reservation now and then.

With us, though, reservations are sacred.

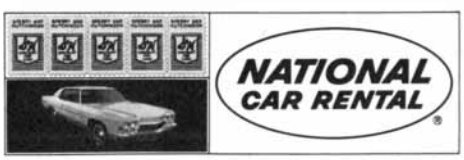
As you can see, it's no cinch working for a company that does things the way National does.

But they figure that the more you have to offer the more people will take advantage of what you offer.

So far, it looks like they're right.

To reserve a car in the U.S. and elsewhere in the world call 800-328-4567 toll free.†

Take advantage of us.



National Car Rental System (in Canada it's Tilden Rent-A-Car). *Valid on U.S. rentals only except where prohibited by law. †In Minnesota call 612-944-1234 collect. ©National Car Rental System, Inc. 1972

LETTERS

Sirs:

In the article "Checks on Population Growth: 1750–1850" [SCIENTIFIC AMERICAN, February] William L. Langer performs a most valuable service in documenting the extent of infanticide during those years.

However, his casual comment that "strong maternal feelings evidently are aroused only on the second or third day after giving birth, when nursing begins" is profoundly ill-based. A baby not separated from its mother will start nursing within moments of birth. Prolactin (mothering hormone) is present even during the latter half of pregnancy and in large amounts at birth. The pleasure-sensation-causing hormone oxytocin is also present at birth. Thus the biological factors for the mother-infant bond are present immediately at birth.

There is one other, more probable explanation. Prolactin actually causes a restless, anxious mood at first, and only after a mother has nursed and cared for her baby does the pleasure of mothering behavior itself replace the anxiety. In a society of highly negative feelings toward infants this anxiety can be directed against the newborn. Since Professor Langer reports that most infants were abandoned as newborns, the mothering behavior never had a chance to become established. If at the same time nursing was delayed two to three days by social custom, the difficulty of establishing maternal attachment was doubly compounded. Nevertheless, even in societies that separate mother and child as part of institutionalized births, most mothers overcome this anxiety, at least in part and with difficulty, and would not allow harm to come to their newborns.

On the other hand, mothers who gestated in a society hostile toward infants (such hostility might be caused by the stresses of overpopulation) would find the infant an easy target for the postnatal anxiety.

The willingness to abandon infants must be attributed not to any biological delay in nursing but rather to an entire social environment.

It is interesting to apply this to current practice in the U.S., where most hospitals do not allow a mother to have or nurse her baby for 12 hours and as much as 48 hours, nor to assume responsibility for its care for four to seven days, thus effectively obstructing the initial imprinting period. Medical practice has extensively worked to interfere also with breast feeding, which prevents the continuous flow of prolactin and oxytocin. Juxtaposing the historical record with the fact that the U.S. ranks 13th in infant survival among industrialized societies [see "Government Investment in Health Care," by Irving J. Lewis; SCIENTIFIC AMERICAN, April, 1971], one must uneasily suspect that, quite predictably, man shares with many other species some innate responses to overpopulation, and that this infanticidal attitude is one of them.

LUCY E. CUTLER

St. Paul, Minn.

Sirs:

The point raised by Miss Cutler is of some importance and, so far as I know, has been but little explored. Lacking medical knowledge, I meant merely to throw out as a possibility the idea that attitudes toward babies and infanticide may have changed with time, much as we see attitudes in matters of sex changing before our eyes. My attention was first aroused by comments along this line in a book by Joseph Kay: *The Social Condition and Education of the People of England and Europe* (London, 1850), pages 445–447, and were then reinforced by the discussion in Philippe Ariès' noteworthy book *Centuries of Childhood* (New York, 1962), pages 38 ff.

I am grateful for the biological data provided by Miss Cutler, but note that in the end we are not very far apart. She speaks of the restless, anxious mood induced by prolactin. I am sure she would agree that the vast majority of women finding themselves with an unwanted child were not only anxious but also desperate, so that their one initial thought may well have been to get rid of the object of their trouble. I would say that for a time at least their desperation would serve as a block to all normal maternal feelings. I hope that my essay in *Scientific American* will evoke further comment from those who have the clinical experience that I lack.

WILLIAM L. LANGER

Harvard University
Cambridge, Mass.

Sirs:

The spectacular pictures illustrating the article on the scanning electron microscope by Thomas E. Everhart and Thomas L. Hayes [SCIENTIFIC AMERICAN, January] brought immediately to mind a 1944 offering of e. e. cummings. One stanza of "Pity This Busy Monster, Manunkind" reads:

pity this busy monster,manunkind,

not. Progress is a comfortable disease:
your victim (death and life safely
beyond)

plays with the bigness of his littleness
—electrons deify one razorblade
into a mountainrange;lenses extend

unwish through curving wherewhen till
unwish
returns on its unself.

(*Poems: 1923–1954,* Harcourt, Brace & World, Inc.)

BRUCE A. HOPKINS

Hamilton, Ohio

Scientific American, April, 1972; Vol. 226, No. 4. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer.

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 Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 25 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 3578.

Please enter change of address ☐
Please enter subscription ☐
☐ 1 year, $10 ☐ 2 years, $18 ☐ 3 years, $25

NAME

NEW ADDRESS

OLD ADDRESS

SCIENCE/SCOPE

The mobile earth station that linked Peking with the world during President Nixon's historic mission was built and operated by Hughes, under contract to Western Union International. It provided capacity for one color TV channel, nine voice commentaries, and 60 two-way telephone channels for use by the Presidential party and by the press to transmit teletype, telephotos, and radio reports. The air-portable terminal was similar to those which Hughes operated in Bogota, Colombia in 1968 for Pope Paul's visit and in Iran last year for that country's 2500th anniversary.

Communications from China were received by Intelsat IV satellites built by Hughes for Comsat, manager for the 83-nation International Telecommunications Satellite Consortium. Stationed over the Pacific and the Atlantic, these satellites carried TV and all press communications from Peking and relayed them to Intelsat's worldwide satellite communications network. Each satellite can carry 5,000 phone conversations, or 12 television programs, or tens of thousands of teletype circuits.

The first 27 Maverick missiles tested by the U.S. Air Force Systems Command surpassed all contract requirements and scored a better than 90 percent success rate in flight tests, enabling USAF to reduce the missile firings by one-third and complete the tests two months ahead of schedule. Maverick also demonstrated lower maintenance requirements and faster aircraft loading time than required.

The air-to-ground missile, built by Hughes, is guided by a miniature TV camera in its nose which the pilot locks onto the target. After launch, Maverick is independently guided to the target and the pilot is free to leave the area safely.

Fingerprint facsimiles and criminal records were sent via satellite between California and Florida recently in a two-week test directed by Project Search (System for Electronic Analysis and Retrieval of Criminal Histories), a consortium of 20 states that account for 80 percent of U.S. crime. Project Search was formed to develop a prototype communications system for law enforcement agencies which would make it possible to exchange police records more rapidly than the 10 days to two weeks it now often takes. Hughes-built equipment used in the test included NASA's 5-year-old ATS-1 satellite and three portable ground stations.

Opportunities for graduate engineers: The continued growth of Hughes' solid-state microwave product line has created an immediate need for engineers with experience in the design and development of either 1) microwave transistor amplifiers, phase lock sources and multipliers, and Gunn and IMPATT diode sources and amplifiers; or 2) millimeter-wave mixers and detectors. U.S. citizenship is required. Please write: Mr. R. F. Wolfe, Hughes Electron Dynamics Division, 3100 W. Lomita Blvd., Torrance, CA 90509. Hughes is an equal opportunity M/F employer.

A target acquisition system, now in engineering development at Hughes, will give individual ships of the U.S. Navy a new means of defending themselves against surprise attack. Integrated with the NATO Sea Sparrow missile, it will enable each ship to defend itself against threats that penetrate the umbrella-like fleet area defenses. The Hughes system integrates a short-range radar and infrared sensors with an IFF system that separates enemy targets from friendly aircraft. It will employ the latest techniques for cancelling out natural and man-made "clutter".

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

APRIL, 1922: "The United States has never until now possessed a truly national road policy. The passage in November of the Federal Highway Act is the first formulation of such a road policy. This bill opens a new epoch in our road construction because, first, it outlines the beginnings of a national highways system and, second, it establishes new principles in the legal aspect of governmental road construction. Hitherto Federal aid has been distributed among the states according to a formula taking into consideration existing road mileage, population and the number of rural free delivery routes. States were allowed to spend their Federal aid upon any roads in the state they desired that might meet with the approval of the Secretary of Agriculture. Now all Federal aid must be spent on no more than 7 percent of such state road mileage, and three-sevenths of this mileage must be 'primary' or 'interstate' highways. The road plan of the State Road Commission is subject to review and approval by the Secretary of Agriculture so that the 'interstate' highways of adjoining states may be made to join. The remaining four-sevenths of the state road mileage that may receive Federal aid must 'connect or correlate' with the three-sevenths."

"It is with mingled emotions that we contemplate a recent announcement detailing the 'improvements' in a certain car for 1922. Without being too specific we can say that these involve, among other things, radical changes in the lubrication, the carburetor, the intake manifold, the radiator, the fan belt, the clutch, the torque member, the braking, the gear ratio, the springs, the rear axle, the front spindles, the steering gear, the frame, the body and the fenders. We wonder just what the owner of last year's model will think when he learns the extent to which his car has been reduced overnight to a pile of junk. The ability to get a car right in the first place and the balance to keep it right thereafter—are not these the things that give us confidence in the designer and builder? The practice of other makers of admitting more or less candidly in their advertising matter that previous models have had serious weaknesses or have been of poor caliber throughout, but that *this* year's car is better, leads to a strong presumption that a saving in dollars and cents will be absent when the time comes to dispose of a car of this kind."

"According to Mr. G. H. Knibbs, the present population of the world, and its recent rate of increase, is such that it cannot be maintained even for a few centuries, since it can easily be shown that no possible accession of food supplies could meet the requirement. Even a very moderate rate of growth cannot indefinitely postpone trouble, and a people can save itself in this respect only in so far as it can increase its efficiency to the highest possible limit, and also restrict itself in regard to all unnecessary luxury. Whatever advances are made in science and its applications in industry, and in food supply, they cannot of themselves postpone the ultimate resistance to further development."

SCIENTIFIC AMERICAN

APRIL, 1872: "After a few days' illness that, with his great age, led to anticipation of a fatal result, Professor Samuel F. B. Morse died at 10 minutes before eight o'clock in the evening of Tuesday, the 2d of April. His long and varied life, and his universal renown, will give interest to the following particulars. He was born at Charlestown, Mass., on April 27, 1791. His early education was acquired at Yale College, and his career would have been through life that of a painter had not circumstances directed his attention to scientific pursuits. In the year 1826 or 1827 he became interested in electro-magnetism and cognate subjects, of which his education at Yale had given him sound practical knowledge. Morse went to Europe in 1829 and returned in 1832, and in conversation with his fellow passengers on the ship concerning the recent obtaining of an electric spark from a magnet he mentioned the idea of an electro-magnetic and chemical recording telegraph. Other inventors in England and France, in the latter part of the 18th century, had been working toward a similar object, and it is impossible and unnecessary to decide to whom the thought first came. Certain it is that it had been largely canvassed long before Professor Morse's labors had begun. But it is to Morse that the credit of long and enduring perseverance in introducing a practical and efficient telegraph is due."

"There is at last a prospect that China, the most impenetrably conservative nation in the world, may yield somewhat to external influences and allow the introduction into her interior of the productions of America and Europe. This hope is held out to us by the conclusion of a treaty between China and Japan, arranging the terms of commercial intercourse between the two countries, stipulating the conditions on which certain ports, to be afterward selected, in each country shall be open to the commerce of the other and appointing a system of arbitration for the settlement of disputes. In spite of the rigorous conditions under which the English have traded with the Chinese in the five treaty ports, Amoy, Ningpo, Shanghai, Foochow and Hong Kong, much good effect on the prejudices of the Orientals has been made by the traders there, and the opening of some more localities to similar influences, particularly to a nation of the same branch of the human family, is likely to widen the beneficial result of that interchanged commerce which has done so much, and which will in the future do more, for the cause of civilization."

"It is the business of science to discover, record and classify facts. Whether these facts conflict with or confirm the religious faith of any does not concern in the least the scientific investigator. If he discovers that the ancestors of mankind were apes, it is his duty to announce his real or supposed discovery; it is not his province to turn upon those who have held a different view and hold them up to scorn or ridicule because they believe they sprang from a higher source and repudiate their anthropoid ape ancestors. Science and religion should not be directly antagonistic; neither one nor the other is benefited by such controversies. All of this we can say, while we own to a decided leaning toward the theory of evolution. It seems more consistent with the way in which an all-wise being would work, that through eternal and immutable laws He should evolve the varied complex structures that people the universe, than that each should be the result of a special act of creation. In this we see nothing that conflicts with such an interpretation of the Mosaic account as would harmonize with the now very generally conceded allegorical and poetical character of that portion of the Bible."

There are two cars built in Sweden. This is the one with front wheel drive for a firmer grip on rain-slick roads.

When we built the first Saab, we built it with front wheel drive for roads that are one big puddle.

From that car to today's Saab 99E, we've seen no reason to change. Because, with the engine weight directly over the drive wheels, you get a better grip on the road.

And with the engine pulling you around curves instead of pushing you, you're less likely to skid. (In a rear wheel drive car, the rear wheels have a tendency to keep going straight even as the front wheels start to turn).

But front wheel drive isn't the only thing that makes us different from the other car built in Sweden.

We have rack and pinion steering for a better feel of the road and quick response.

Impact-absorbing bumpers that reduce your collision insurance rates 15% at Allstate.*

Headrests with a hole in them so they don't block your view.

The world's only electrically-heated driver's seat to keep you warm until the rest of the car warms up.

And "roll cage" construction. The kind that soon, by law, may be required on all cars.

But the best way to tell the difference between the two cars built in Sweden is to find a road that is really "slippery when wet" and drive both cars over it.

We think you'll want to keep driving the Saab.

SAAB 99E

Before you buy theirs, drive ours.

*Allstate discount available in most states. For the name and address of the dealer nearest you, call 800-243-6000 toll free. In Connecticut, call 1-800-882-6500.

THE AUTHORS

GUNNAR MYRDAL ("Political Factors in Economic Assistance") is professor emeritus of international economy at the University of Stockholm, founder and former director of the Institute for International Economic Studies at the university and chairman of the Stockholm International Peace Research Institute. Receiving a degree in law from the University of Stockholm in 1923, he practiced law in Sweden for a time but in 1927 began his long association with the economics faculty at the university. He has also served the Swedish government in such capacities as adviser, member of the Senate and minister of commerce. Myrdal's most recent book is *The Challenge of World Poverty: A World Anti-Poverty Program in Outline,* published in 1970.

DENIS KEEFE ("Collective-Effect Accelerators") is at the Lawrence Berkeley Laboratory of the University of California, where he is in charge of a group investigating collective-effect acceleration. Born in Dublin, he began his university education in engineering, turned later to mathematics and then to experimental physics. Receiving his Ph.D. from the University of Bristol, he taught for some years at University College Dublin, working on cosmic ray air showers, atmospheric electricity and elementary-particle physics. He joined the Lawrence Laboratory in 1960, intending to stay two years, and has been there since.

WILHELM G. SOLHEIM II ("An Earlier Agricultural Revolution") is professor of anthropology at the University of Hawaii and a research associate in the department of anthropology of the Smithsonian Institution. He was graduated from the University of Wyoming in 1947 and obtained his Ph.D. from the University of Arizona in 1959. He joined the faculty of the University of Hawaii in 1961.

PETER GOLDREICH ("Tides and the Earth-Moon System") is professor of planetary science and astronomy at the California Institute of Technology. He was graduated from Cornell University in 1960 and obtained his Ph.D. in physics there three years later. After a year as a postdoctoral fellow at the University of Cambridge he was a member of the faculty of the University of California at Los Angeles until he went to Cal Tech in 1966. Goldreich writes: "My scientific interests are the understanding of natural phenomena in terms of physical laws. Most of my work has been in astrophysics, but I have also done a little geophysics and hope to do more. I have published papers on solar-system dynamics, galactic dynamics, pulsars and, most recently, astrophysical masers. I spend much of my time and energy in teaching, and I derive a great deal of pleasure and frustration therefrom. My main interests outside my work are my family, vigorous physical exercise and reading."

RICHARD E. DICKERSON ("The Structure and History of an Ancient Protein") is professor of chemistry at the California Institute of Technology. He received his bachelor's degree at the Carnegie Institute of Technology in 1953 and his doctorate in physical chemistry from the University of Minnesota in 1957. After a stay in England, where he worked with J. C. Kendrew on the first high-resolution X-ray analysis of the structure of a crystalline protein (of sperm whale myoglobin) he spent four years on the faculty of the University of Illinois. He moved to Cal Tech in 1963. His main interests are in analyzing the crystal structure of proteins, investigating the relation between protein structure and function and studying molecular evolution. In addition to the work he describes in his article, he and his colleagues have solved the structure of the enzyme trypsin and are now looking at several bacterial cytochromes that are believed to be evolutionary near-relatives of eukaryotic cytochrome *c*. Dickerson is the author or coauthor of four textbooks. He describes his spare-time activities as "hiking, harpsichord and German numismatics."

NORMAN GESCHWIND ("Language and the Brain") is James Jackson Putnam Professor of Neurology at the Harvard Medical School and director of the neurological unit of the Boston City Hospital. He received his bachelor's degree from Harvard College in 1947 and his M.D. from the Harvard Medical School in 1951. After interning at Beth Israel Hospital in Boston he spent three years in London at the National Hospital in Queen Square, one of the major centers for the study of diseases of the nervous system. Returning to Boston in 1955, he spent a year at the City Hospital and two years working in the biology department of the Massachusetts Institute of Technology. In 1958 he joined the neurology department of the Veterans Administration Hospital in Boston, which he says "presented a magnificent opportunity for the study of aphasia" because the hospital had a special aphasia ward. Geschwind eventually became chairman of that department and later of the department of neurology at the Boston University School of Medicine. He returned to Harvard in 1969. He has published extensively on disorders of the nervous system and is preparing a paper on Dostoevski's epilepsy and the descriptions of epilepsy in Dostoevski's novels.

DONALD P. SNOWDEN ("Superconductors for Power Transmission") is in the special nuclear systems division of Gulf General Atomic, where he headed while it was active a group on applied superconductivity. His degrees are in physics: a bachelor's degree from the California Institute of Technology in 1953 and master's and doctor's degrees from the University of California at Berkeley. Snowden writes that apart from science he spends a "considerable amount of time on church-related activities," designs and builds furniture, "although perhaps that should be in the past tense because I seem to have run out of space to put anything else," and is a member of the San Diego Theatre Organ Group, which restored the four-manual, 32-rank organ in the Fox Theatre in downtown San Diego. The organ, which had not been played for 30 years, is now used regularly for concerts. Snowden has also built an electric organ; he says, however, that he seldom plays organs and does so "only when no one is around to hear."

ROGER A. MORSE ("Environmental Control in the Beehive") is professor of apiculture at Cornell University. He has been involved with bees since he was 10 years old, when he was given a hive by his father. During World War II he and his father maintained several hundred bee colonies in the Hudson valley and the Catskill Mountains, which they operated for honey production. Morse took his bachelor's, master's and doctor's degrees at Cornell; his work for the master's involved making mead, or honey wine, and last year he and an associate at Cornell received a patent for a mead-making process they developed. After receiving his Ph.D. Morse worked for the Florida State Plant Board as a bee specialist and was assistant professor of horticulture at the University of Massachusetts. He returned to Cornell as a member of the faculty in 1957.

Conversation Pieces

Technically intriguing items from TRW, guaranteed to add luster to your conversation and amaze your friends.

It Takes A Very Big Thing To Make A Very Small Thing

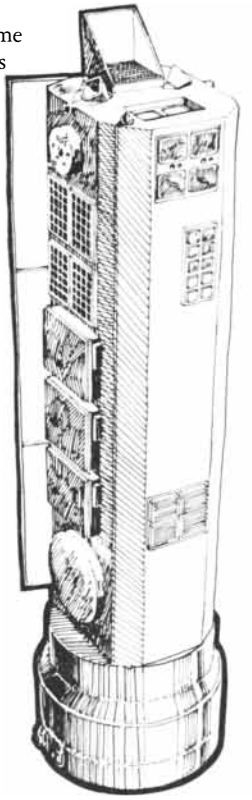
When they're not actually living together in tight quarters, those positively charged particles called protons can't stand one another's presence. If you try to bring two of them together, they fly apart at a great rate. Yet the carbon in your body has six protons sharing the small apartment of its nucleus, and the oxygen you are breathing has eight. Eighty protons are crammed together in the nucleus of a heavy element like Mercury, while the nuclei of the heaviest elements entertain lavish parties of more than 100. How were these mutually repelling protons assembled in the small compass of a nucleus? What force in nature is strong enough to have squeezed such inimical particles into such a tiny space?

The answer seems to lie in the stars. Stars are huge chemical factories within which matter is made. The great temperatures and pressures found inside stars drive together the nuclei of light elements like hydrogen (which contains a single proton) to form the heavier elements. The process is called *nucleosynthesis.* Our own star, the Sun, for example, fuses four hydrogen protons together and makes helium. At a later stage in its life, it will begin building other elements from the helium like carbon, then oxygen and neon. Eventually, it will make magnesium, silicon, and, finally, iron.

Many stars stop there and end their lives as "white dwarfs." In some, however, the process continues to even heavier elements. A few make all of the elements. These erupt in gigantic explosions called supernovae and hurl out the matter they have made in great clouds or nebulae. Some scientists believe that this matter eventually gathers together to form planets. All of the matter that makes up the Earth and its inhabitants, for example, may well be stardust. (Incidentally, Sir Arthur Eddington, not Hoagy Carmichael, is responsible for this insight.)

As the stars drive the nuclei of the elements together in thermonuclear reactions, a small fraction of their mass is converted to energy. Part of this energy comes forth as visible light, making the stars shine. Other parts of it stream forth as invisible radiation—X-rays, gamma rays, radio waves, heat, and so on. These energy by-products are valuable information, for they indicate what the stars are making, how old they are, how far away they are, and a number of other things.

Beginning in 1975, NASA's High Energy Astronomical Observatory (HEAO) will collect data on the high energy electromagnetic output of stars and the nuclei and electrons they sometimes eject (known to us as cosmic rays). It will give us fundamental information on some very important questions such as how matter is made, how old and how big the universe is, and where our planet may have come from.



To intercept these relatively rare high energy particles and short wave lengths, you need large, heavy detectors. Thus HEAO is big. In fact, it's three stories high and weighs around 10 tons. TRW is building it for the Marshall Space Flight Center.

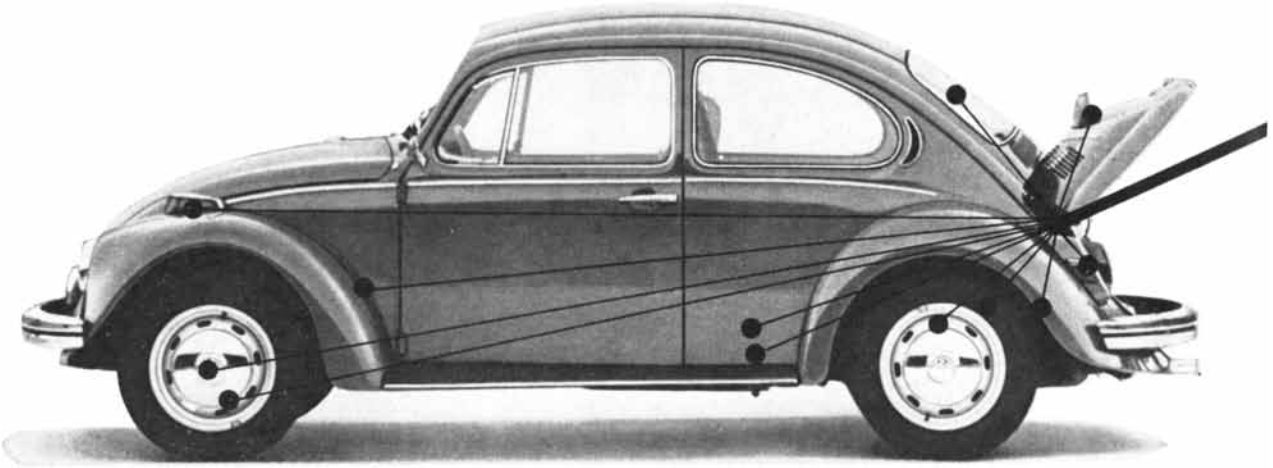
We're very happy to be associated with this program. Because stars are a primal source of matter and energy (and hence you and me), we think HEAO is one of the most scientifically fascinating programs to come along in a long time. After HEAO goes up, we'll keep you posted on what the stars tell us.

For further information, write on your company letterhead to:

TRW®
SYSTEMS GROUP

Attention: Marketing Communications, E2/9043E
One Space Park Redondo Beach, California 90278

A car so advanced, it can tell you just about everything that's right or wrong with it.



Remarkable new invention in every 1972 Volkswagen.

We've all been through it.

The agony of an automotive check-up.

You sit nervously thumbing through old magazines, praying that somehow you can get out of there for less than $50.

Then, the moment of truth:

A Service Manager telling you that one of his best mechanics thinks you need a new generator.

Those days will be over soon if you own a 1972 Volkswagen.

For instead of a mechanic telling you what he thinks is wrong with your car, now your car can tell you for sure.

Sound amazing?

It is.

A car wired like a space capsule.

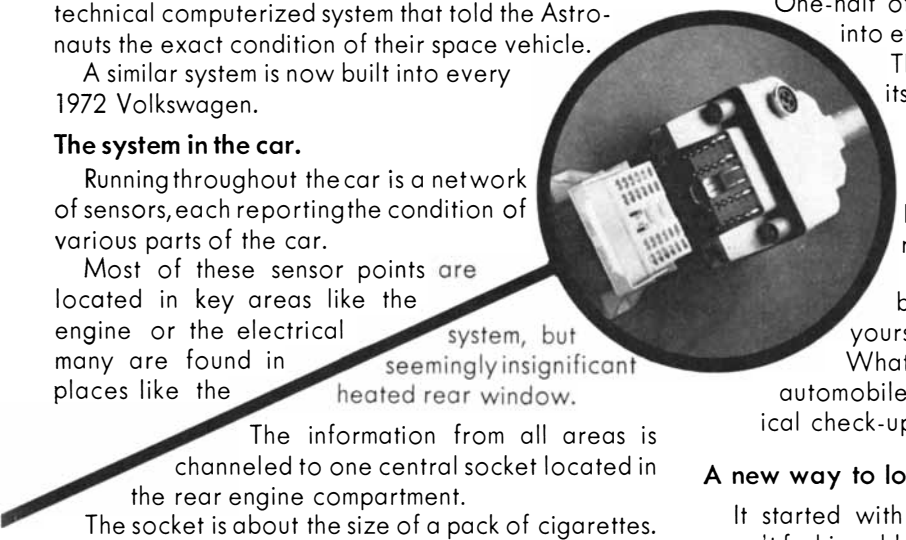
When Man went to the Moon, the success of each mission depended a great deal on a highly technical computerized system that told the Astronauts the exact condition of their space vehicle.

A similar system is now built into every 1972 Volkswagen.

The system in the car.

Running throughout the car is a network of sensors, each reporting the condition of various parts of the car.

Most of these sensor points are located in key areas like the engine or the electrical system, but many are found in seemingly insignificant places like the heated rear window.



The information from all areas is channeled to one central socket located in the rear engine compartment.

The socket is about the size of a pack of cigarettes.

We mention the size only because of what happens next. And that's what this amazing socket can do.

60 vital service checks.

Soon you'll be able to take any 1972 Volkswagen into an authorized VW dealership for the most advanced automotive check-up in the world today.

At that time, your car will actually be plugged into a computer.

And in half the time it takes to perform a conventional check-up, 60 vital service checks will be made and recorded.

Checks wheel alignment in 10 seconds.

In 10 seconds, you'll know if your front wheels are properly aligned.

In a minute, you'll know the condition of the compression of all engine cylinders.

Without a mechanic so much as taking a peek, you'll know whether or not your battery needs water.

Ignition, cylinder compression, dwell angle, generator, electrical system—

All checked out without human error.

In effect, your car will be telling you how it feels directly.

And once again, this information is emanating from that one tiny socket built into the back of every 1972 Volkswagen.

Results printed out in plain English.

One-half of the system is already here. Built into every new Volkswagen.

The other half, the computer, is on its way.

Imagine.

A computer five feet away from your car is printing out in plain English just about everything that's right or wrong with that car.

When all 60 service checks have been made, the print-out sheet is yours to keep.

What better proof to show that your automobile has finally had a thorough physical check-up?

A new way to look at a VW.

It started with economy, back in 1949, when it wasn't fashionable for an automobile to be economical.

But since when has a VW been fashionable?

Since never.

Obviously, the Volkswagen Beetle hasn't made it on looks alone.

But then, that's always been the plan.

While everyone else has been worrying about how their cars looked, we've been worrying about how ours acted.

And now, after all that time, we've even advanced it to a stage where it can speak.

Burn garbage, and what have you got?

Airborne garbage.

Dump garbage, and what have you got? Less usable land. Or polluted oceans.

In all of history, these have been the only ways to dispose of solid waste.

Now, not a minute too soon, there is another way; Landgard™ system.

Developed by Monsanto's subsidiary Enviro-Chem, Landgard is a system that bakes garbage rather than burns it. And then, to avoid air pollution, it inhales its own gases and consumes them.

All types of municipal solid waste go into our process.

And what comes out is 94% smaller than what went in.

In the course of doing this, we can recover the metal for re-use. And we can use the energy produced by the system to make steam and electricity.

We've never seen a cleaner way to build our business

at **Monsanto: the science company.**

Established 1845 SCIENTIFIC AMERICAN April 1972 Volume 226 Number 4

Political Factors in Economic Assistance

It is argued that the global downward trend in aid from the rich countries to the poor countries can be reversed only by stressing the moral argument

by Gunnar Myrdal

In October of last year the United States Senate voted overwhelmingly to kill the country's annual appropriation for economic assistance to the underdeveloped countries. Although funds were later restored at a lower level, the Senate vote stood as a declaration of disillusionment and frustration in which the most ardent supporters of foreign aid had joined with its most consistent opponents. The outlay for foreign aid—totaling $70 billion over the years since the end of World War II according to some reckonings—was held by all parties to have failed every purpose asserted by the confusion of motives for which the money had been appropriated. The "foreign handout program," in the view of its opposition, had won the U.S. "no gratitude" and "no friends"; in the United Nations Assembly a few days before the Senate vote 46 beneficiary nations had voted against the U.S. resolution on the admission of the People's Republic of China (that would have kept Taiwan in). Speaking for the friends of foreign aid, Senator Frank Church deplored the confounding of humanitarian motives with cold-war objectives and of economic assistance with military assistance in the country's transactions with the nations of the poor; he echoed Captain Cook's remorseful statement about the natives of the Pacific Islands: "It would have been better for these people never to have known us."

The question of foreign aid has come up again this year on the agenda of the second session of the Ninety-second Congress, but the compromise foreign-aid bill that has emerged has certainly not ended debate on this contentious issue. I propose to consider here the motivations and objectives advanced for the economic-assistance program in the hope that such reflection may help to place debate about the issue on a sounder and hence securer ground in the future. In particular I shall examine the rationalization most often embraced by the warmest proponents of the program. This is a variant of the classical liberal argument that trade promotes world peace. Aid is advanced as a kind of extension of trade that according to the classical doctrine works to bring about international harmony. Thus in the resolution of the UN General Assembly proclaiming the Second Development Decade development and international cooperation to promote it are characterized as "the essential path to peace."

To come straight to the point, I do not believe economic policies can make much of a contribution to peace-keeping. During the entire century before World War I and right up to the present there has been a tendency to give too much stress to economic factors in international relations. Liberal economic theory, from the classical writers onward, has not differed on this issue from the Marxist position as expressed in the official doctrine of the Communist countries and expounded in general terms by their statesmen and scholars. It is assumed in both camps that trade leads to peaceful political relations. Although there may be an element of reason in this consensus, I hold that that element is a small and heavily conditioned one. The much more important causal relations go in the opposite direction: a low level of trade and a distortion of trade patterns are the result of political conditions and are indications that the political climate is not favorable to peace.

Politics is sovereign. No political division line becomes established as the boundary between states without affecting the economic relations between those states. Even a boundary that puts the most "unnatural" economic barriers in the way of trade and economic relations tends in time to become "natural" in the sense that the economies on both sides become adjusted to the political conditions so created. Thus the partition lines that divided India and Pakistan terminated trade and financial links established over centuries. Such a division not only may set tremendous inhibitions and obstacles against the economic development of both countries but also may imply, as recent history has shown, grave dangers for peace. Yet this adverse development is caused not by economic determinants but by political ones. Even cities, such as Berlin and Jerusalem, can be brutally divided by a political line, physically expressed by a wall or a no-man's-land. The adjustments called forth by political fiat thereupon create condi-

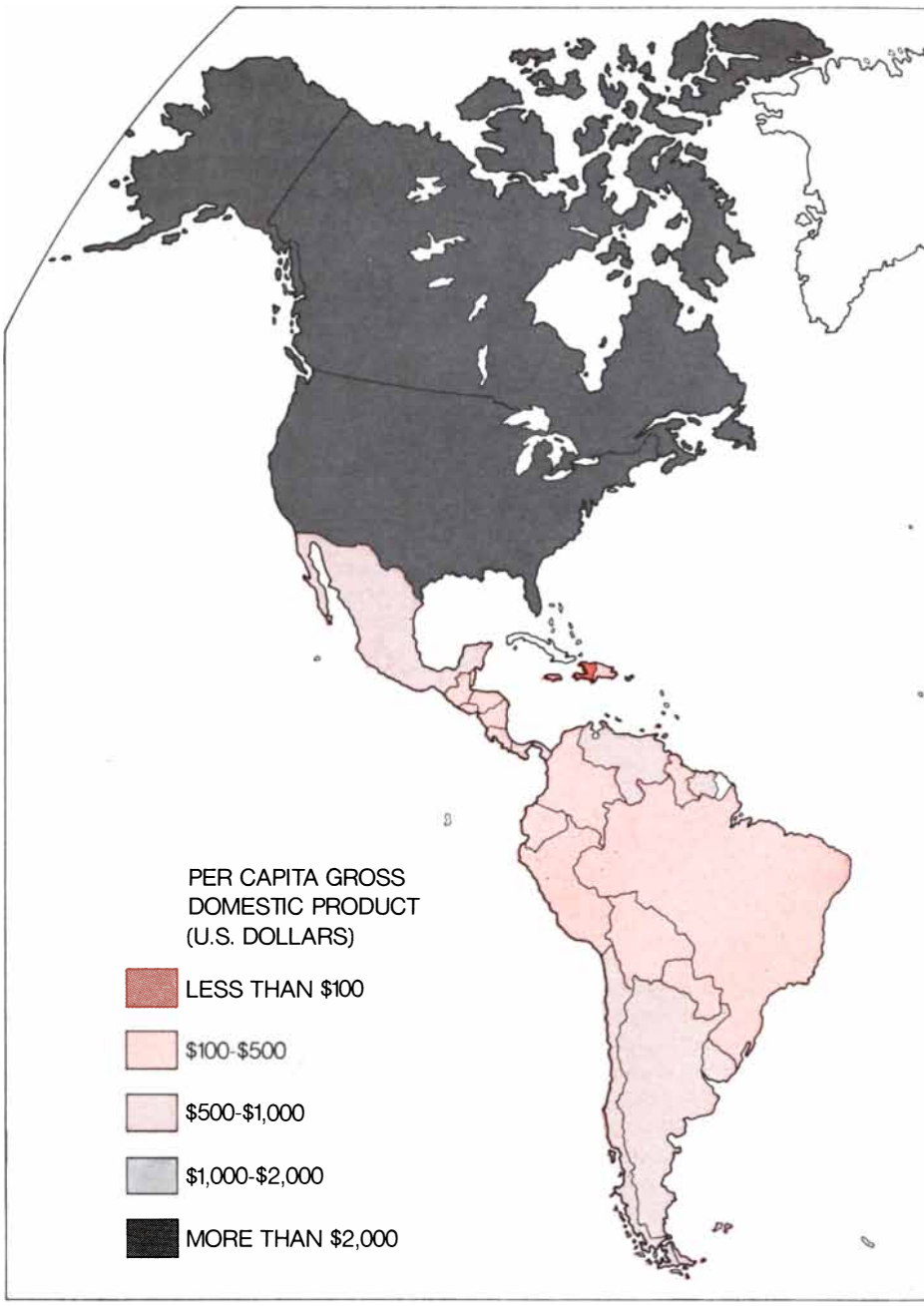
tions of permanency for the broken trade and financial relations such that these conditions make the "unnatural" "natural," or at least "more natural."

Political forces, moreover, can strangle economic relations almost instantaneously. To reverse the process and reestablish normal relations usually proves to be a difficult and time-consuming process. Even though a new political situation might allow it, the preexisting economic relationship may never be fully restored. The adjustment of the economies to the lower level of trade is bound to have changed resource allocations in the direction of self-sufficiency. As the cumulative outcome of vicious-circle processes set in motion by political events, the return to *status quo ante* may be largely foreclosed.

Recognition of the overriding importance of political conditions and political changes is essential to an understanding of the course of economic relations between western Europe and eastern Europe in this century. Throughout the years between the two world wars trade and financial relations between the U.S.S.R. and the rest of Europe (including, incidentally, the countries in central Europe and eastern Europe now counted in the Soviet bloc) remained at a very low level. The explanation lies largely, of course, in the strained political relations that prevailed from the October Revolution on between the insecure revolutionary country and its hostile, worried and even frightened neighbors. Whatever international trade and economic policy the U.S.S.R. had during this period was mainly directed to accommodation with political reality. The wartime alliance against Germany and Italy brought economic and military assistance from the U.S. to the U.S.S.R. At the end of the war East-West economic relations returned promptly to their low prewar level and stayed there for nearly a decade. The reason was again political, this time the steadily intensified "cold war." The Soviet trading bloc, however, included this time the socialist countries of eastern Europe bound to the U.S.S.R. by political ties.

A major political instrument in enforcing the inhibition of trade between the East and the West was the strategic export licensing enforced by the U.S. These controls were accepted, albeit reluctantly, by the western European countries as well as by the client countries of the U.S. in other parts of the world. Hindsight shows that this policy of the U.S. was not unwelcome to Stalin. During those years, when the U.S.S.R. was seeking to solidify its Eastern bloc, Stalin undoubtedly found the U.S. policy useful as evidence that the Soviet satellite countries had to rely on themselves, the U.S.S.R. and one another not only for their defense but also for their economic development. The economic side of the cold war presented during those years a textbook demonstration of the process of circular causation with cumulative results. In this case the vicious circle was tightened by a kind of "cooperation" between the two superpowers, each overreacting in anticipation of the other's reaction to its own initiatives in the conflict.

The death of Stalin in 1953 started significant political changes on the Soviet side. At the same time political changes were under way in western Europe. Marshall Plan aid from the U.S. was petering out, the flow of American military aid was decreasing and the economies of western Europe were getting back on their feet. The western Eu-

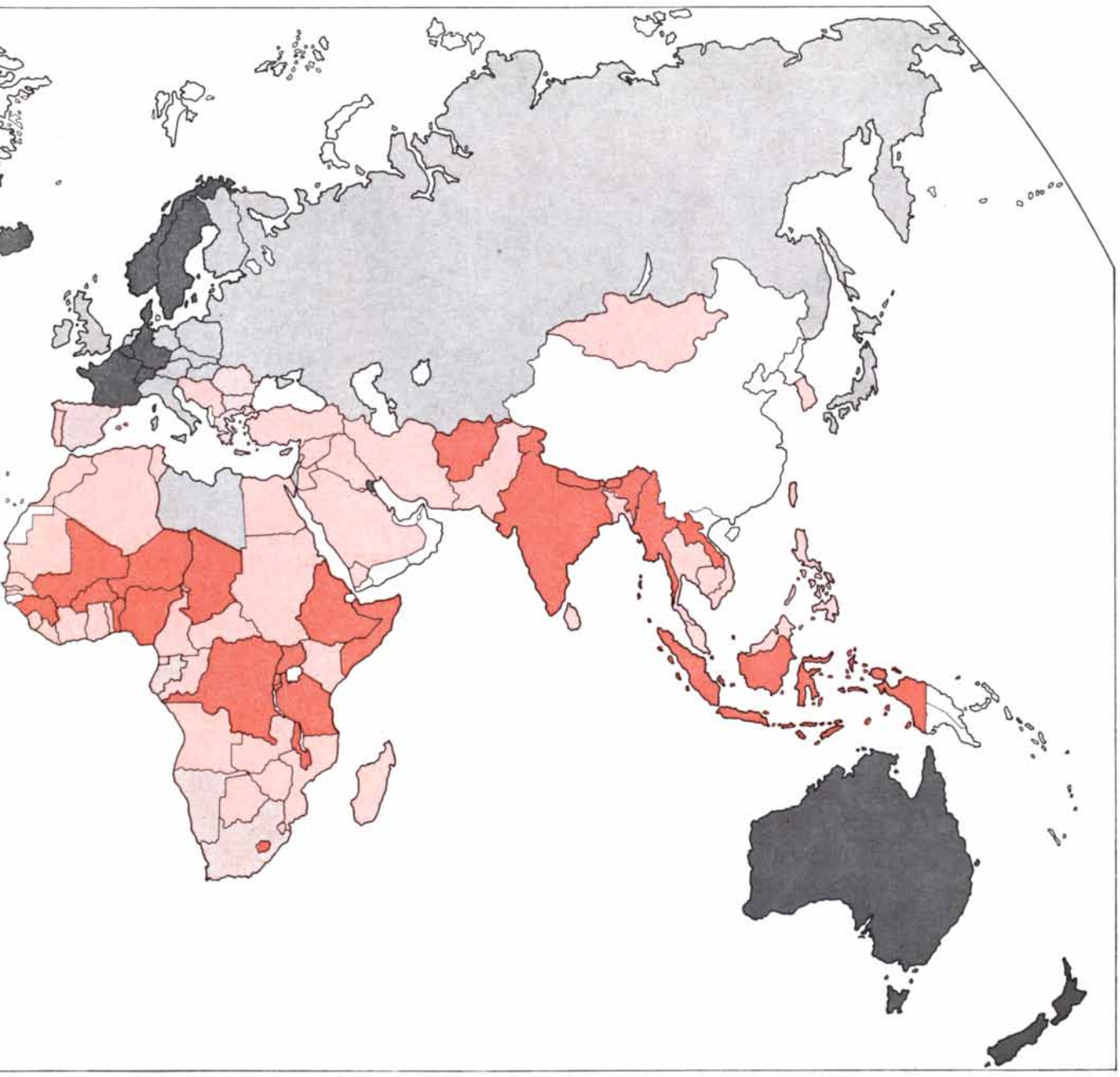


RICH COUNTRIES AND POOR COUNTRIES of the world are roughly categorized on this map according to the level of their estimated annual per capita gross domestic product, expressed in U.S. dollars (*see key at lower left*). The statistics on which the categories are based are for 1967, the latest year for which such figures are available for all the countries shown. The data were gathered from various official and unofficial sources by the Centre for Development Planning, Projections and Policies of the United Nations Secretariat. The

ropean countries used their greater independence to cut back the very inclusive U.S. trade-embargo list. Today, as far as Europe is concerned, the remnant list of forbidden exports is more of a nuisance than a serious trade barrier.

As early as 1954, in my capacity as executive secretary of the Economic Commission for Europe, I was able to call a meeting of trade consultants from countries on both sides of the cold war. In the guise of consultation with the secretariat, government representatives from East and West met for the first time, bilaterally and multilaterally, to discuss increased trade. Between several of the countries on opposite sides economic relations were at that time completely or almost completely extinguished. Following these consultations and the ensuing official negotiations between capitals East-West trade increased rapidly to substantial, although still "unnaturally" low, levels.

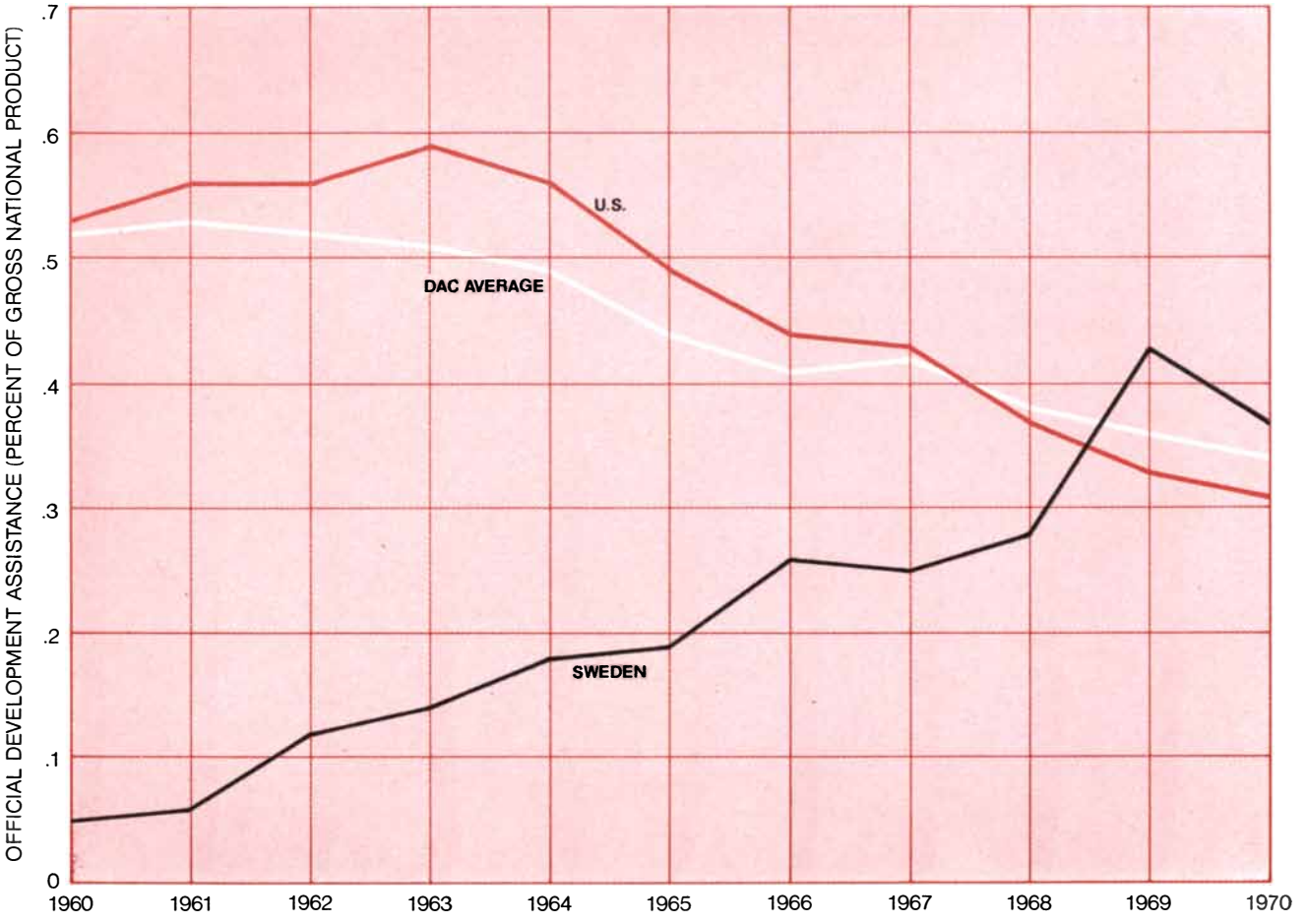
I cite this history in order to stress the decisive importance of political factors to the flow of trade. The record controverts the widely held, diffuse idea that a low level of economic relations presents a major cause of political tension and the corollary idea that political tension may be cured or substantially relieved simply by raising the level of trade. It was improvement in political relations between the Eastern and Western blocs that allowed the rise in trade between countries



figures for most of the developed and underdeveloped countries were estimated at current market prices. The figures for some of the centrally planned economies were estimated by another method and thus are not fully comparable with those of the other countries. Economic data for 1967 were unavailable for the countries and territories that are shown in white. In the author's view the available estimates of per capita gross domestic product and similar economic-development statistics are "highly unreliable and often totally useless." The illustrations accompanying this article are provided by the editors of SCIENTIFIC AMERICAN in order to make some general points, in this instance to show the rough geographic distribution of the developed and the underdeveloped regions.

on the two sides. Without doubt the increase in trade has in turn lessened to some extent the political tension between them. This round of circular causation with cumulative effects has led, in fact, to some constructive political negotiation between West and East. That I should deny the ameliorative effect of trade on politics is out of the question; for periods of my working life it has been my main occupation to strive for this desirable end. Nevertheless, economic influences must continue to be minor in the politically divided world of today. Resistance to such strivings comes mainly from the political sphere.

The prospects for the economic development of the poor countries of the world and for the delivery of any significant assistance to this process from the rich countries present the same issues in a theater of wider dimensions and with probably more decisive consequences for the destiny of the human species. In the early stages of the development of welfare-state policies in the rich nations of the West, it was argued that assistance to the poor constituted a kind of insurance against violent revolt from below. Now it is argued that aid to the poor countries is in the "self-interest" of the rich countries. Almost every statement on the subject, by statesmen, by journalists and even, I am sorry to have to add, by economists, contains the glib assertion that the poor countries must be aided in order to preserve the peace of the world. The inexplicit premise is that reasonable progress in economic development will make the underdeveloped countries and their governments more peaceably disposed toward one another and toward the rich countries, and that otherwise they will be pugnacious and rebellious. This is an entirely unsupported rationalization of how people who are well off think they would feel and act if they lived in misery and saw hope, or alternatively no hope, of



RECENT DECLINE in the commitment of most of the developed countries, including the U.S., to foreign aid is reflected in this graph, which charts the flow of "official development assistance" as a percentage of the gross national product of the donor nation (or nations). "Official development assistance" is defined by the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD), the compiler of the data used to make this graph, as "all flows to less-developed countries and multilateral institutions provided by official agencies, including state and local governments, or by their executive agencies, which meet the following tests: (*a*) they are administered with, as the main objective, the promotion of the economic development and welfare of developing countries, and (*b*) they are intended to be concessional in character." The figures for official development assistance are presented on a net basis, that is, after first deducting repayments on earlier lending. Measured by this standard, the average aid-giving performance of the 16 DAC countries (Australia, Austria, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Sweden, Switzerland, the United Kingdom and the U.S.) can be seen to closely parallel the faltering performance of the U.S. in recent years. In sharp contrast, the commitment of Sweden to foreign aid has been growing rapidly as a percentage of G.N.P. in the same period. The "International Development Strategy of the Second United Nations Development Decade," adopted by the UN General Assembly in October, 1970, states that "each economically advanced country will progressively increase its official development assistance to the developing countries and will exert its best efforts to reach a minimum net amount of .7 percent of its gross national product at market prices by the middle of the Decade." The actual performance of the advanced countries with respect to this UN goal is even poorer than such a graph suggests, since, as the author remarks, "figures for aid and assistance are opportunistically falsified upward, particularly those for the U.S."

improvement. Known facts provide no proof for such assertions.

If any generalization can be made, it would be rather that people get restless and rebellious when they find their material circumstances improving but not improving fast enough. In fact, some scholarly individuals have constructed models, incorporating selected evidence, to put forward this contrary theory. They have presented it as valid for racial tension in Detroit or Newark, for the Sukarno "confrontation" with Malaya and for wars and rebellions everywhere. Like other single-factor theories of human behavior, it is characterized by insufficiently critical study of empirical facts.

Another theory, equally unsupported by evidence but invoked to impress on politicians and the public the desirability of foreign aid, holds that an underdeveloped country will become more "democratic" if it makes a little progress. It is, of course, an open question what democracy can mean in largely illiterate countries, with masses of people apathetic, ignorant of their interests and still less organized to seek or to protect their interests. Nor is there any empirical substantiation for the proposition that "democracy" depends on or results from economic progress. The further idea, so often pressed on the heels of this one, to the effect that democracy is favorable for peace between countries is quite obviously a false proposition. Among developed countries, where democracy can be invoked as a more meaningful concept, history shows that the people of democratic nations are often more belligerent than their government.

Behind the whole farrago of theories about peoples' political reactions to the economic facts of poverty stands the popular notion of the "revolution of rising expectations." This sometimes appears in the apparently more learned guise of what is called the demonstration effect. At bottom, even though the idea is propounded by the most irreproachable personages of the West, it turns out to be nothing more than an elaboration of Marx's theory of the political consequences of the impoverishment of the workers. As a prerequisite of rebellion, according to this modern theory, actual impoverishment is not necessary; all that is needed is the failure to realize expected improvement. Marx himself was careful to exclude from the working of his theory the *Lumpenproletariat:* the ignorant, apolitical masses. Marx aside, no honest empirical research has ever supported this proposition. It belongs in the category of statements by people who are well off, whether they are citizens of rich countries or poor, of how they think they would feel if they were as poor as the poor of the poor countries.

The notion of the revolution of rising expectations has some relevance, however, to the state of mind and the political behavior of the tiny upper class (including the so-called middle class of the educated) in the poor countries. When they voice such sentiments in international conferences, they may be speaking with real sincerity, since they may be expressing their own feelings and convictions about themselves and their nation. The poor countries, rather independently of whether their constitutions specify universal suffrage or more authoritarian bases for governmental power, are ruled by shifting alliances of people in the upper strata. What these people have to say, therefore, has something to do with how their countries behave politically and militarily in their international relations. But then two questions remain unanswered: To what extent are these oligarchies moved by the economic progress they themselves may be experiencing compared with the degree to which such progress may be shared by the rural and urban masses of their country? What particular degree of economic progress is required to satisfy their expectations and keep them peaceful?

In sum, the ideas now so generally in circulation about the connection between economic progress and peacefulness are empty of objective knowledge and give no reliable guide to the questions of peace and war between countries or of tranquillity and rebellion within countries. This does not imply that I do not deem it urgent to preserve the peace. Nor do I mean to suggest that the rich countries do not have responsibility for promoting the economic development of the poor ones.

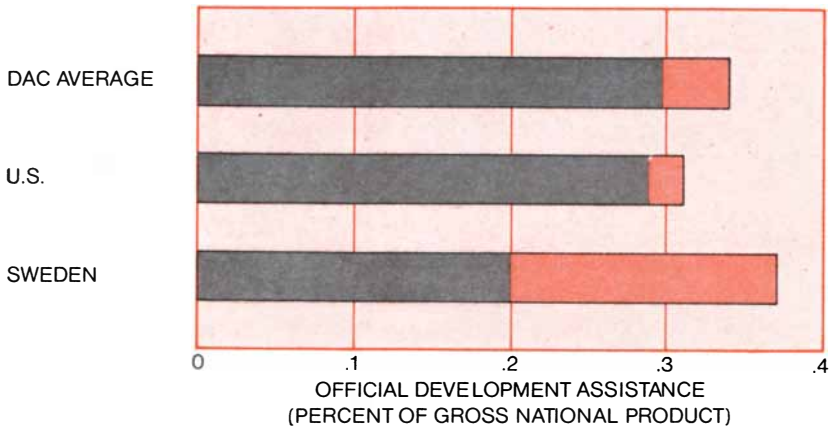
The present world situation is characterized, on the one hand, by a nervous peace among the developed countries. This order of things is upheld by the balance of terror between the two superpowers. A full-scale war cannot any longer be won, in any meaningful sense, by either of them. The minor powers, for their part, dare not risk war with one another and dare even less to risk it against either superpower. Most of the smaller and weaker developed countries have acknowledged the situation by committing themselves to political and military alliances with the U.S. or the U.S.S.R. There are, of course, elements of horrifying uncertainty in the balance of terror. The continued rapid rise in arms expenditure is rending the world economy; it has destroyed, for example, the monetary and trade arrangements of the West established at the end of World War II. For the time being, however, peace prevails among the developed countries.

The underdeveloped world presents, on the other hand, a contrasting panorama of domestic unrest and rebellion and of war or the threat of war between nations. This aspect of the contrast between rich and poor is most ominous. It cries out for measures to bring peace to the great majority of mankind that inhabit the poor countries. Such measures have not been taken. On the contrary, the rich countries—more specifically, some of the rich countries—have promoted unrest and disorder in the underdeveloped world.

Back in colonial times the metropolitan countries certainly maintained peace in and among their colonies. They left legacies everywhere, however, of potential conflict created by their exploitation of ethnic and religious divisions among their subject peoples. Economic relations among the countries that formerly were colonies, therefore, often reflect the isolation of the colonies from one another by policies that lined them up in economic dependence on cultural and general identity with their so-called mother countries. Strong elements of neocolonialism persist in the economic relations of the rich and poor countries. Former colonies are now client states of their former and, for some, new masters; their governments maintain unbroken the old economic dependencies.

Some of the mother countries, notably France and the Netherlands, tried to preserve their dominance in the present postwar period by conducting bloody colonial wars. The U.S. took over from France the war in Vietnam and has now spread its military intervention over the whole of Indochina. In this war the U.S. has used methods of warfare that the great majority in the UN consider forbidden by international law and that, significantly, it would not use in a war against another developed country.

The underdeveloped world is everywhere divided by the inconsiderate pursuit of political and strategic interests conducted by the superpowers and their allies. The two networks of alliances stand in the way of economic and other collaboration among the underdeveloped countries. Most of the aid from the rich to the poor nations has gone to build systems of satellite countries in the underdeveloped world. The poor have been encouraged to use their resources to buy armaments and have had weap-



COMPOSITION of official development-assistance programs also varies significantly. In particular, as this bar chart shows, the ratio of multilateral contributions (*color*) to bilateral grants and loans (*gray*) is substantially higher for Sweden than it is either for the U.S. or for most of the DAC countries. The statistics were compiled by DAC for the year 1970.

ons bestowed on them as military aid. Even the underdeveloped countries that seek neutrality in the great terror conflict have tended to become dependent on the economic and military aid that is pressed on them from both sides. As a survey of the Stockholm International Peace Research Institute (SIPRI) has just shown, the flow of armaments from the rich to the poor countries has mounted to the stupendous total of $17.4 billion over the years since the end of World War II, two-thirds of it from the U.S. ($7.7 billion) and the U.S.S.R. ($4 billion). To this one-way commerce, I am distressed to say, even Sweden has made a contribution ($80 million), although that country does not give away any weapons and has much more restrictive legislation governing the conditions and politics of the importing country.

The underdeveloped countries have in fact been the pawns of the cold war. The principal burden of that conflict is now visited on them. The reason is that the superpowers and their allies among the developed countries find it safer to pursue their conflict in the "third world." They have shown themselves willing to allow open warfare there; they have even promoted it. Thus the conflict in the Middle East has taken on the character of a war fought by proxies of the two superpowers, with the other veto-holders in the UN Security Council playing ambiguous roles. The principal common interest in the cushioning of this conflict is to prevent it from leading to a head-on collision between the U.S. and the U.S.S.R. So also it is known that the tragic civil war in Nigeria was fought with weapons given or sold by Great Britain and the U.S.S.R. to one side and by France to the other. The large-scale butchery in East Pakistan before the recent war was conducted with weapons furnished largely by the U.S. to the Pakistan government; the Indian intervention was conducted with weapons supplied by the U.S.S.R. and under the shelter of the new treaty between India and the U.S.S.R.

It was this kind of war that the UN Security Council was created to prevent and stop. The dependence of this machinery on the agreement of the two superpowers has now been demonstrated both positively and negatively in the stubborn conflict between India and Pakistan—by success in the war between the two countries in 1965, when the two superpowers could agree, and by failure in the conflict of last year, when the superpowers could not agree.

If this brief review and analysis of how wars get started in the underdeveloped world is at all correct, it shows that the causes are political. No economic measure, including assistance to spur development, can reach the roots of those causes, because they lie outside the relations among the underdeveloped nations. If the developed nations could agree that they would not use aid to foment and supply civil wars and other wars in the underdeveloped world and would instead use such influence to discourage war among the poor countries, then it could be argued that economic measures can safeguard the peace in the world. Such a change in policy, however, would have to be based on an antecedent change in the way the superpowers and their allies make use of the underdeveloped countries in their own political conflict. If such a change were possible, then many more political measures of even more direct importance for peace-keeping would also be possible. These are envisioned in the chapters of the charter of the UN that deal with the functioning of the Security Council.

The entire literature on the development of the underdeveloped countries is heavily biased. So much of the literature is semiofficial that it is scarcely to be dignified by close examination of the theoretical apparatus through which its biases and inexplicit premises operate. Suffice it to say that it is opportunistic, as all knowledge (and ignorance as well) turns out to be when it is critically examined. If the overoptimistic views that pervade the literature were correct, effective aid would be a bargain for the developed countries, and the ruling oligarchies of the underdeveloped countries could hope to see their countries developing without themselves having to put up with the inconvenience and hazards of otherwise inescapable domestic reforms.

Rectification of this situation in both theoretical and empirical research is needed for two reasons. First, economic literature ought to prize the truth as the common standard of scientific endeavor. Second, the truth must be known if the peoples of the rich countries are to be impressed with the urgent need for sacrifices on their part in providing help for the poor countries and equally to give the support of truth to the liberal forces in the poor countries that are struggling, against heavy odds, for domestic reforms. The statistics on aid have been juggled and falsified in a most scandalous way to give the appearance that the flow of aid is much bigger than it really is. As a result the people of the so-called donor countries have never been imbued with the motivation necessary to generate aid in really substantial and effective volume. Today the flow of aid in real terms is diminishing in most countries. The nature of the aid available has also been deteriorating, as loans take the place of gifts and as the flow is bound increasingly to exports from the "donor" country.

The wrong motives are commonly put forward for the giving of aid. This applies not only to the peace-keeping fallacy I have dealt with here at length but to the "down-to-earth practical" arguments that find favor in some quarters. When aid is justified as being "in the best interests of the U.S.," those interests when specified turn out to be political, strategic and military advantages of that country. Americans are also urged to appreciate the prospective commercial advantages, whereas the French are exhorted to find motivation in the prospect of

the continued cultural domination by France of her former colonies.

My first explanation for the faltering and decline of aid from the U.S. and elsewhere is that the standard arguments advanced for it do not appeal to and are not believed by the people who make up the majority of the electorate. The disastrous course of the diplomatic and military policies pursued by the U.S. in the underdeveloped countries of Asia and Latin America has confirmed the worst misgivings of the American people. Concurrently, as is to be expected, the country has been steeply reducing its commitments to foreign aid.

The trend in aid is moving in exactly the opposite direction in Sweden for reasons that sustain this explanation. In that small country, which has had no colonies for more than a century, it took a long time for people to recognize the need for aid to the underdeveloped countries. Now, for several years, the country's outlays for aid have been increasing by 25 percent annually. Something similar is happening or can be foreseen in a few other small countries on the sidelines of power politics.

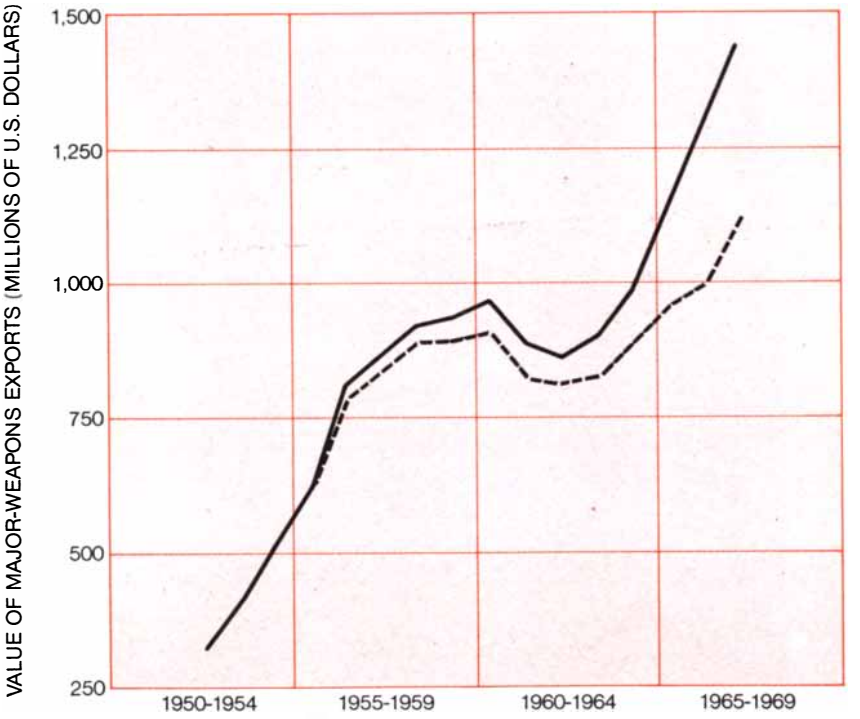
None of the familiar motivations so often advanced in other countries can be discerned in the Swedish aid program. If only for geographical reasons, it is difficult, perhaps impossible, to persuade the Swedes to fear an onslaught from the poor nations if they are not aided. Sweden is not a participant in the cold war; Swedes therefore have no stake in getting underdeveloped countries on their side against the Communist powers. Still less are Swedes moved by ambition to spread their language or culture. Old-fashioned free-trade traditions keep the Swedes, who are conservatives in this respect, from seeking commercial advantage from aid. As a result Sweden has not related aid to deliveries from its own country, as almost every other country has done.

In Sweden the only motivation that could effectively be presented to the people has been human solidarity and compassion. To say this, I realize, implies an invidious comparison between the U.S. and Sweden. I am not arguing, however, that Swedes are different in this respect from Americans or people in other developed nations. The statement illustrates my central thesis. It is my firm conviction, not only as a moralist but also as a social scientist who has studied these problems, that the moral argument is the only one that holds. The moral argument must be stressed if the global downward trend in aid to the poor nations is to be reversed. That argument, honestly advanced and honestly observed in the actuality of the aid program, will carry as much weight with Americans as it does with Swedes.

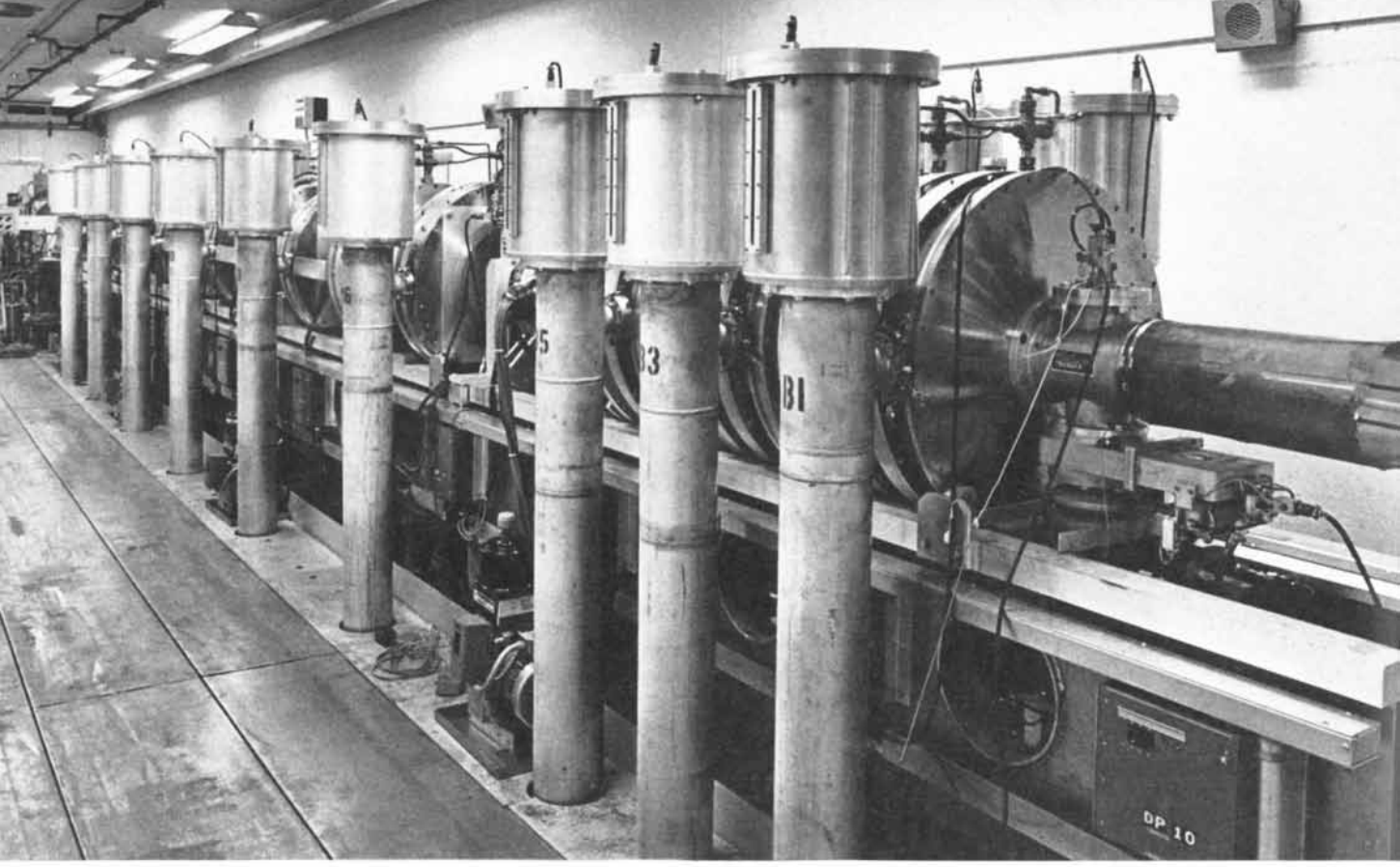
For the Swedes, it is true, the uninterrupted and successful progress of the welfare state at home has provided favorable conditions for this development in their foreign policy. No one for a long time has suggested in Sweden that aid to poor people constitutes a form of "insurance" for the better-off. Furthermore, with the really needy in Sweden reduced to a very small and steadily shrinking minority, it is less possible than it is in the U.S., for instance, to argue that charity begins at home and that aid should be limited by national boundaries.

The moral justification for aid has a reciprocal meaning for the underdeveloped countries receiving help, a meaning that is crucial to the development of those countries. As long as they are ruled by rich oligarchies, as long as they show great and increasing inequalities, as long as land reform remains a sham and the opportunity for education remains undemocratic and therefore antidevelopmental in its effects, it will be difficult to argue for aid on the moral ground of human solidarity and compassion. The ordinary man will ask: "Why don't they tax their own rich and reform their own countries before they come to us with their begging bowl?"

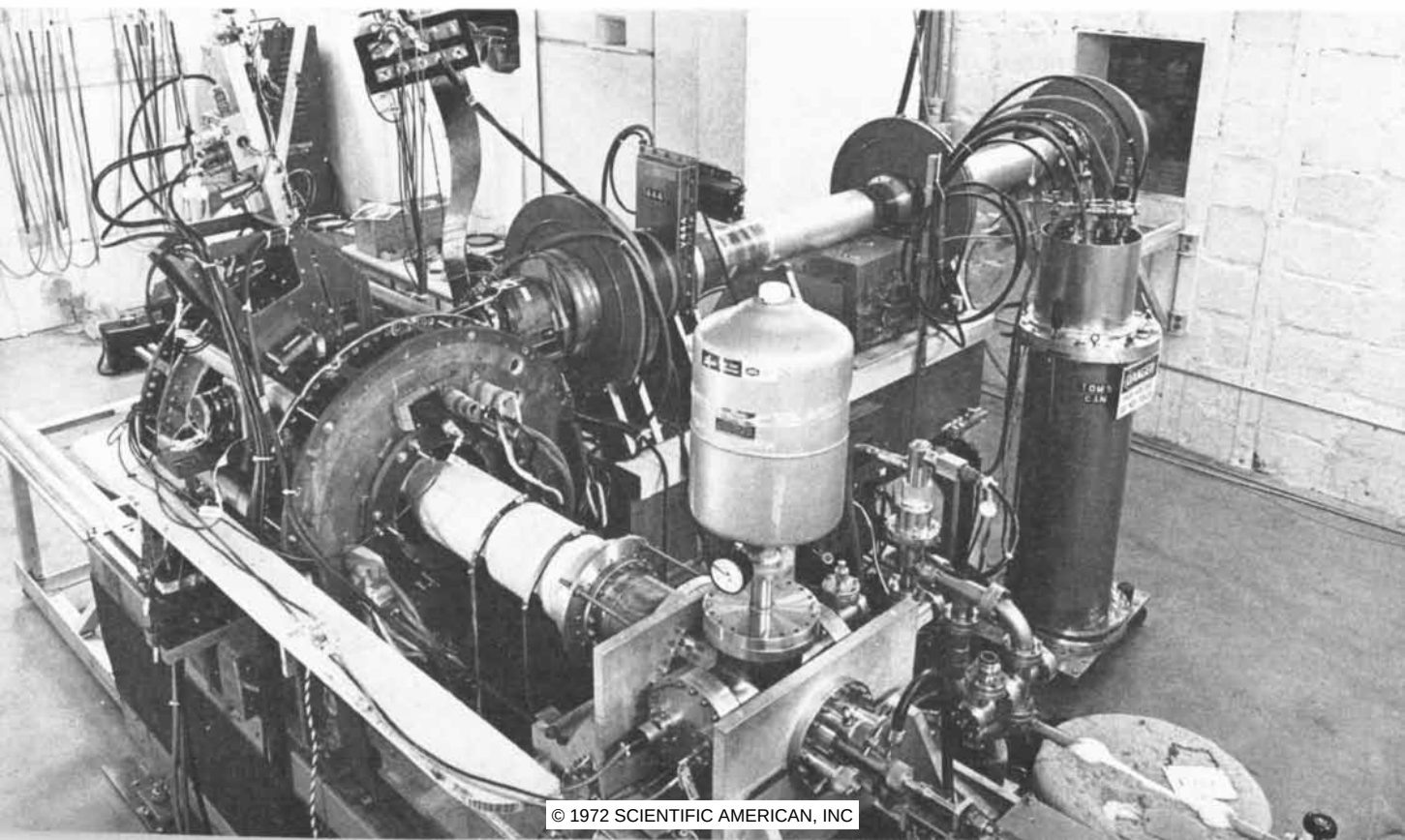
No aid can be morally neutral. That much has been taught by the history of economic assistance as administered by the superpowers and their allies. If the flow of aid is now to be increased substantially, it will be necessary to set moral and political conditions on its use in the recipient country. Preference must be given to underdeveloped countries committed to internal reforms. If aid is given to other countries, it must be on terms that call for increased equality in the sharing of the sacrifices and the benefits of development. Such rules have been explicitly laid down—although they have not yet been radically enforced—in the aid policies of Sweden. They are stated here as being not only essential to the revival of aid from the rich countries but also equally in the interest of rapid and steady economic development in the poor countries.



MILITARY EXPENDITURES by the poor countries consume an increasing proportion of the resources needed for their economic development. This graph charts one aspect of that diversion of resources: the flow of major weapons (naval vessels, aircraft, missiles and armored vehicles) from the rich countries to the poor countries from 1950 to 1969. (Dollar amounts indicated are five-year moving averages at constant 1968 prices.) During this period the gross national product of all the poor countries grew at an average rate of about 5 percent per year, while their major-weapons imports grew at an overall rate estimated at either 9 percent per year, excluding Vietnam (*broken curve*), or 11 percent per year, including Vietnam (*solid curve*). The graph is from a recent study, *The Arms Trade with the Third World*, prepared by the Stockholm International Peace Research Institute.



ELECTRON-RING ACCELERATOR under study at the Lawrence Berkeley Laboratory of the University of California is shown in these two views. The photograph at top shows the electron injector: the part of the machine that gives the electrons their initial high energy. The electrons are produced in intense pulses by a field-emission source at the near end of the injector. They are then accelerated down an evacuated tube by a series of pulsed induction cavities, which are located inside the large rings connected to the chimney-like vertical structures arrayed along both sides of the tube. After passing through a hole in the far wall the high-energy electron beam enters a device called a compressor, which is shown in the bottom photograph. The function of the compressor is to shape the electrons into rings, which are then further accelerated at right angles to the original beam. It is during this second stage of acceleration that the collective "self-fields" set up by the electron rings will be made to carry along with them protons or other positive ions. Schematic diagrams of the entire layout and of the details of the compressor appear on pages 30 and 31 respectively.



COLLECTIVE-EFFECT ACCELERATORS

The collective "self-fields" generated by intense beams of electrons can be harnessed to accelerate positive particles to high energies. In one case the electrons are first curled into ring-shaped clusters

by Denis Keefe

The electrically charged subatomic particles of a particle accelerator are accelerated and otherwise manipulated by electric and magnetic fields generated by the accelerator. Such particles of course have fields of their own, and the fields of many particles add up. Can the collective fields of many particles, appropriately marshaled in a different kind of accelerator, be used to accelerate other particles? Such accelerators, which open up many attractive possibilities, are being vigorously investigated in numerous laboratories around the world.

The first attempt to accelerate one species of particle by means of the collective effects of another species was made some 20 years ago in Stockholm by Hannes Alfvén and Olle Wernholm, who tried to accelerate positive ions by carrying them along inside focused clusters of electrons. The experiment was unsuccessful because it was not possible at that time to obtain an intense enough electron beam. After a long hiatus interest in collective effects was rekindled, largely because of two developments. First, the technology of producing intense electron beams has advanced considerably in the past few years. Second, a few of the many suggested methods of applying this mode of particle acceleration have proved to be susceptible to theoretical treatment. Although this article will be concerned primarily with the exploitation of collective effects for the purpose of accelerating protons and other positive ions, several quite different applications of the collective-effect principle will be mentioned in order to give some idea of the wide range of potential uses that have been suggested for this powerful new experimental method.

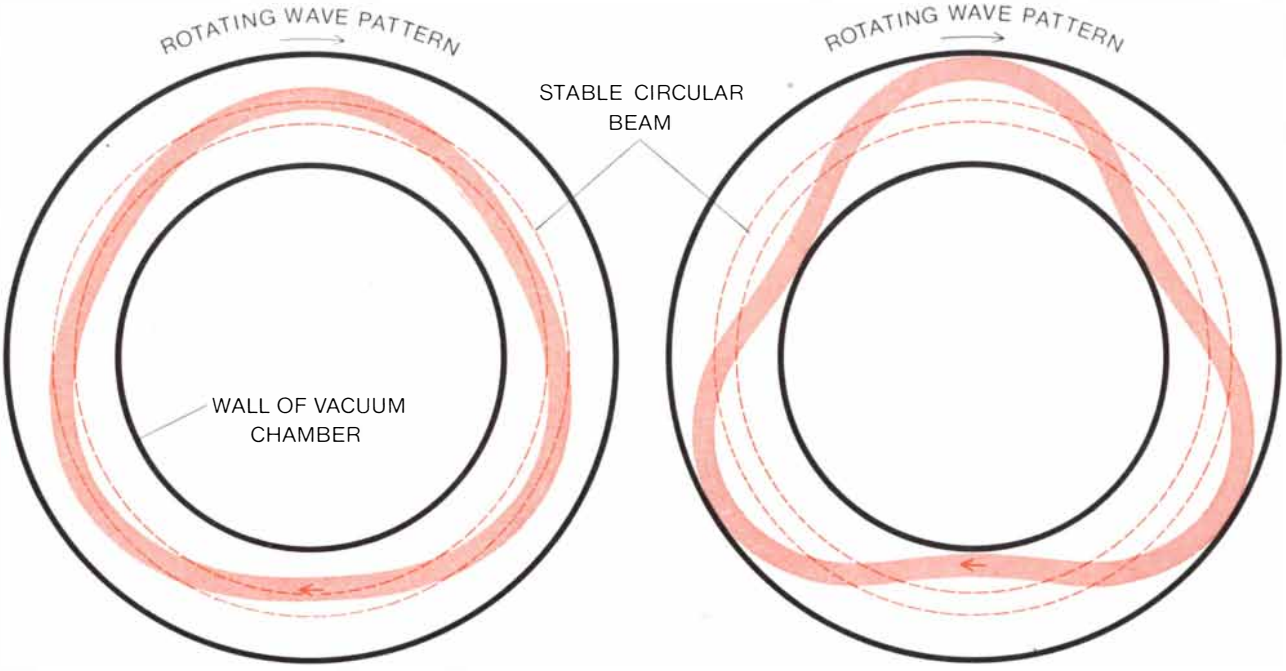
Let us begin by considering some of the salient features—and limitations—of conventional particle accelerators, in order to clarify just what it is the collective-effect approach has to offer. All modern high-energy accelerators have two basic features in common: an electric field to boost the energy of the particles and a magnetic field to guide them in the desired direction. In a linear accelerator the particles are sent once through a long succession of radio-frequency cavities; the magnetic field guiding them in a straight line is rather small and is set up by means of a few magnetic "lenses" situated between the cavities. Typical rates of energy gain in a linear accelerator are seven million electron volts (MeV) per meter for electrons and 1.5 MeV per meter for protons.

For accelerating particles to very high energies linear accelerators become inconveniently long. This disadvantage was bypassed by Ernest O. Lawrence's invention of the cyclotron in 1931. In a cyclotron particles are recirculated through a single small radio-frequency device by a magnetic guide field that curves them into circular orbits. The invention of the synchrotron and later of the alternating-gradient principle in turn led to dramatic simplifications in the circular guide-field system; these innovations meant that the bulky cyclotron magnet could be replaced by a slender ring of magnets, thus making it possible to achieve much higher particle energies at reasonable cost. For each meter of the accelerator structure a typical rate of energy gain in a proton synchrotron is 40 MeV per meter. (For example, the large alternating-gradient synchrotron at the Brookhaven National Laboratory has an energy of 30,000 MeV and is housed in a tunnel 800 meters long.) The rate of energy gain, measured in MeV per meter, is an important figure of merit to bear in mind, since the total cost of a high-energy accelerator scales roughly as the total length of its structure (magnet, tunnel, shielding, utilities and so on). The larger this figure of merit, the cheaper the accelerator, or alternatively the higher the energy that can be achieved for a given cost.

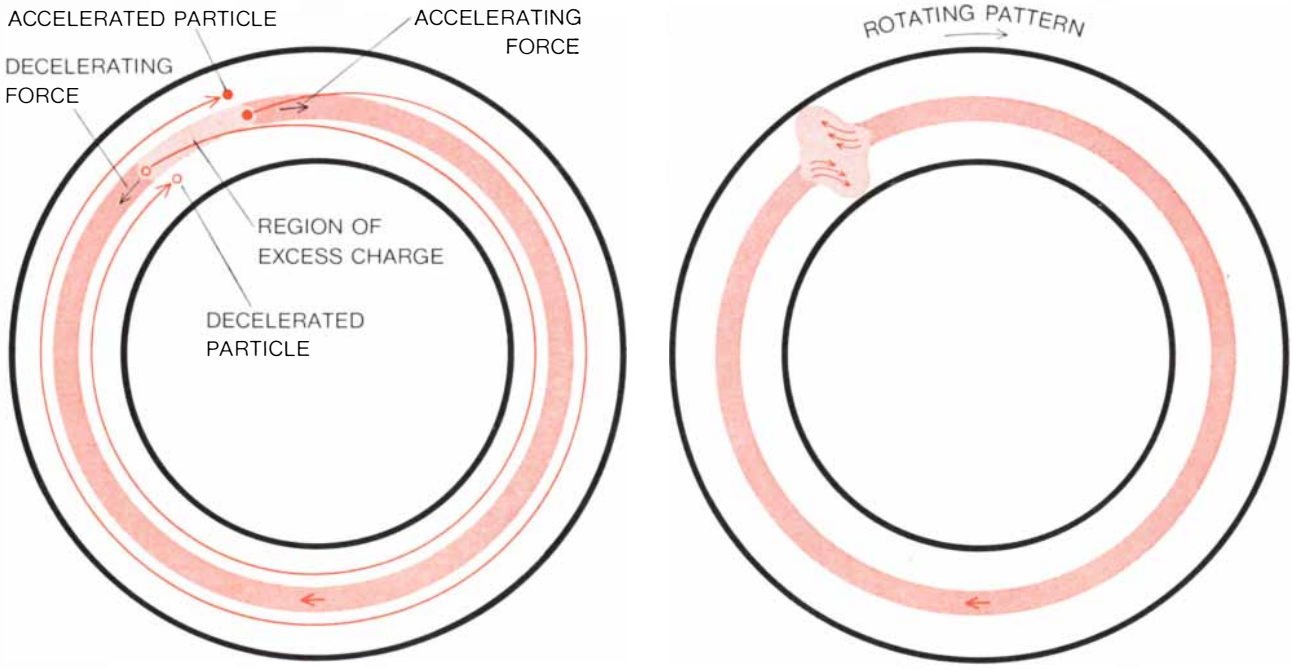
Intimately related to the economics of the accelerators are the technological factors that limit conventional machines. For brevity one can crudely define "conventional field" limits of about 10 million volts per meter for the electric accelerating field (the sparking limit) and about 20,000 gauss for the magnetic guide field. The present active development of superconducting magnets and superconducting radio-frequency cavities shows promise of leading in time to an extension of these limits by another factor of two or so. For conventional accelerators such technological limits in essence determine both the size and the cost.

Technology apart, what are the limits on beam performance imposed by the laws of physics? At low currents, that is, with a small number of particles present, the "single particle" stability problems are clearly understood in both linear and circular accelerators. For example, it is now known how to choose the guide field and the accelerating field for a linear or a circular machine in such a way that the particles are contained in well-defined bunches of stable size riding synchronously with the radio-frequency fields. There can still be forces that could cause the particle bunches to diffuse and eventually be lost by striking the walls of the vacuum chamber, but such forces can be eliminated by careful design.

In the past decade another class of stability problems has become increasingly important. This development coincides with the push to higher beam-intensities in accelerators, and with the



COHERENT TRANSVERSE INSTABILITY is an example of an undesirable collective-effect phenomenon that can limit the beam intensity of conventional high-energy particle accelerators. In this situation the circular beam as a whole behaves like a snake or a rope, developing an integral number of waves in stable positions around its circumference (*left*). The instability grows in amplitude until the beam is "wiped off" against the walls of the vacuum chamber (*right*). Such instabilities can become troublesome when the collective self-fields generated by the beam reach a level that is still very small compared with the externally applied fields.



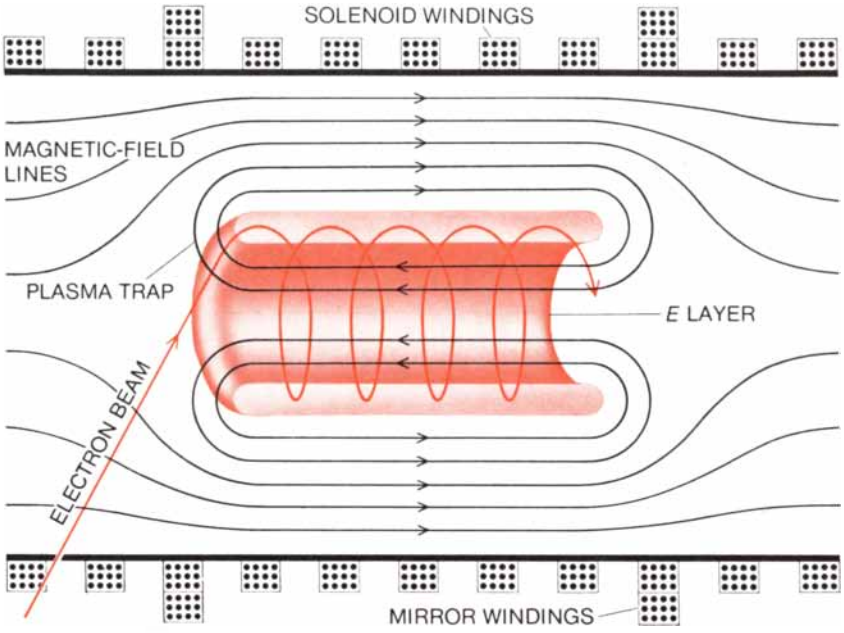
NEGATIVE-MASS INSTABILITY is another disadvantageous result of collective behavior that can interfere with the proper functioning of a conventional particle accelerator. By a statistical fluctuation an excess of charge can appear in a circular particle beam (*left*). Electrostatic forces would accelerate particles near the leading edge of such a zone and decelerate particles near the trailing edge. The accelerated particles would travel in a larger orbit with a longer revolution time and hence after another turn would tend to recede somewhat in the bunch. Conversely the decelerated particles would travel in a smaller orbit with a shorter revolution time and therefore would advance in the bunch. As the particles pushed forward recede and the particles pushed backward advance, this so-called negative-mass effect can lead to large radial spreading and the destruction of the beam as a continuous circular thread (*right*).

advent of particle-storage rings, which need intense beams. At a certain level the electric field generated by the beam (because it is a line of charges) and the magnetic field generated by the beam (because it is an electric current) cease to be negligible and can act back on the beam in a variety of disturbing ways [*see illustrations on opposite page*]. These collective "self-fields," which are generated by the beam as a whole, that is, the beam acting as a collection of charged particles, can become troublesome when they reach a level that is still very small (1 percent or less) compared with the externally applied laboratory fields. Many categories of collective instabilities are known but not all are understood; it is worth noting that essentially all accelerators and storage rings are today limited in performance by unfriendly collective phenomena.

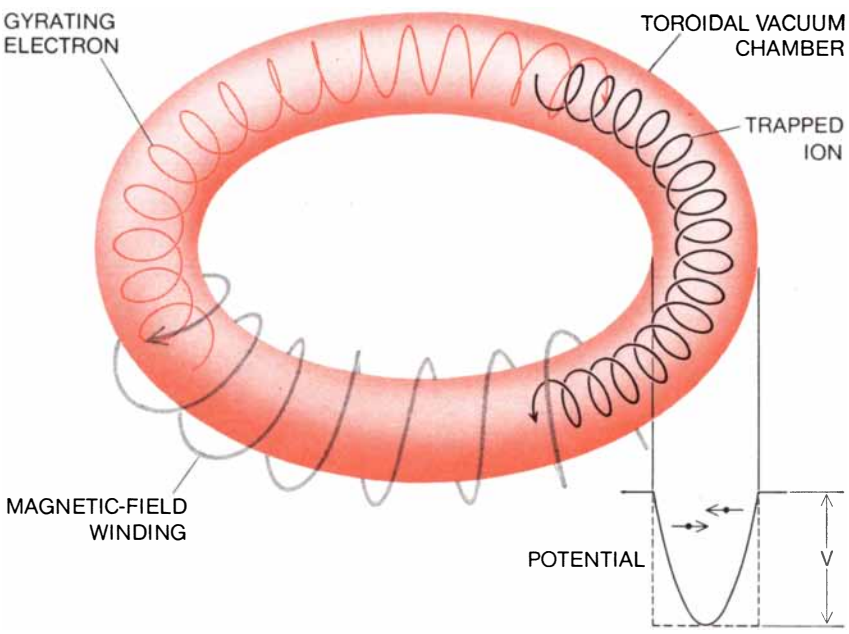
So much for the disadvantages of collective behavior. It may seem like a giant mental leap to consider next beams of charged particles capable of producing electric and magnetic fields that are not just tiny fractions of conventional laboratory fields but of comparable strength or even much stronger! A number of investigators speculated seriously on the benefits that might be gained from such effects more than two decades ago. Two main advantages could be foreseen from producing strong fields with intense particle beams. First, one would be able to create specially shaped magnetic or electric fields in regions of space where electrostatic electrodes or magnetic-field coils would not be feasible. Second, if collective-effect fields much stronger than conventional laboratory fields could be achieved (and controlled), one could break through the technological barriers limiting conventional accelerator guide fields and accelerating systems. This could lead either to more compact accelerators or to superhigh-energy devices.

These two rather generally stated advantages can be best illustrated by looking at a few examples. (We can return to the question of the feasibility of creating and controlling the intense particle concentrations that are needed.) A good example of the first class of applications is found in the controlled-fusion research program and not the accelerator-research field; it is the plasma-containment device called Astron, invented by Nicholas C. Christofilos of the Lawrence Livermore Laboratory of the University of California.

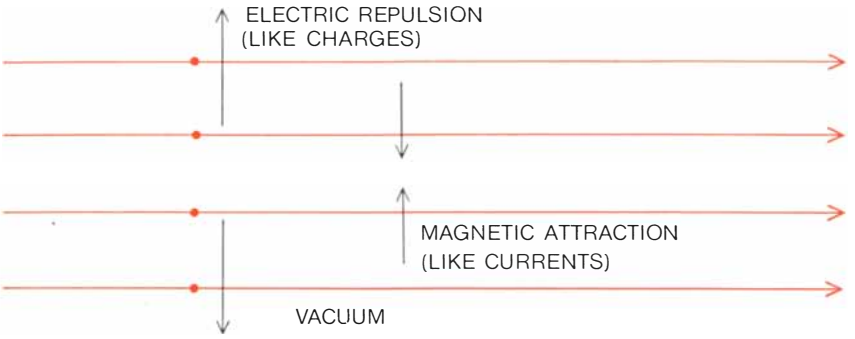
In principle an isolated short solenoid magnet could be made to generate a



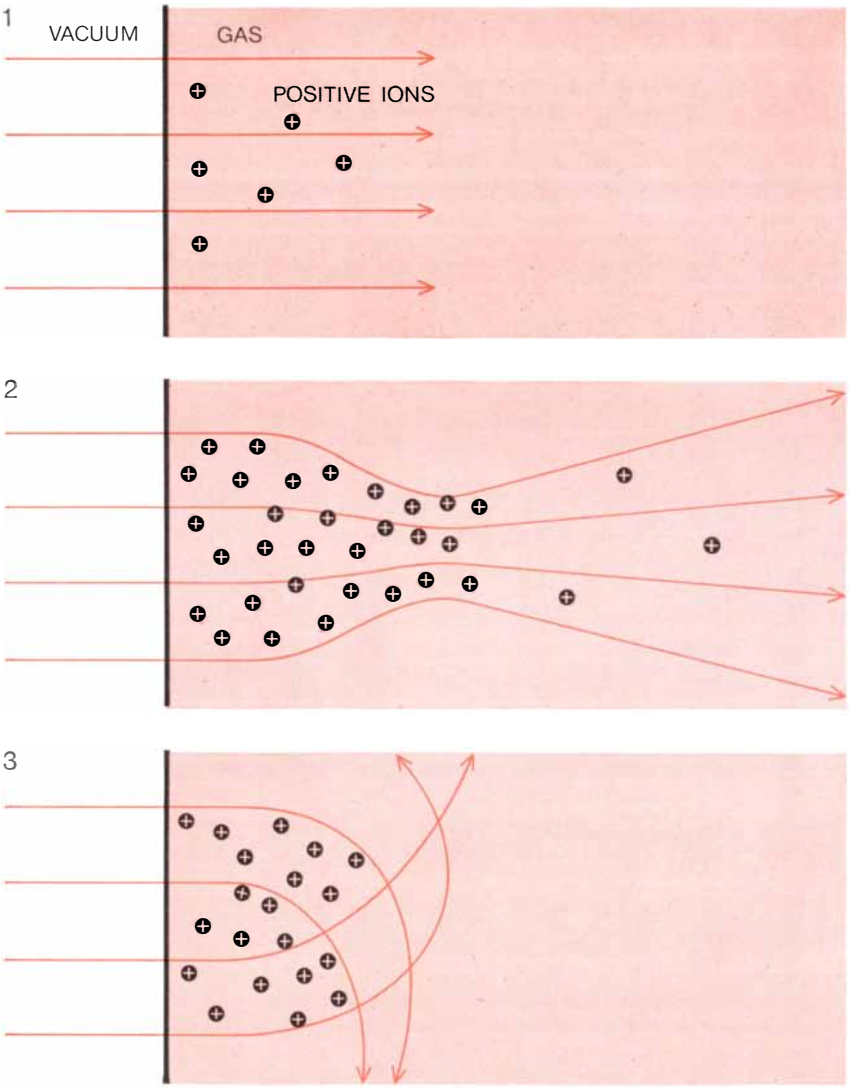
ASTRON CONCEPT, invented by Nicholas C. Christofilos of the Lawrence Livermore Laboratory, is an application of the collective-effect principle to the problem of magnetically confining a hot plasma for research on controlled fusion reactions. In essence the device consists of a suspended, nonmaterial solenoid made from a sheet of circulating electrons. The electron sheet, called the *E* layer, is made by injecting an intense beam of relativistic (very-high-energy) electrons into a large conventional solenoid. If enough electrons can be stacked in the *E* layer, the magnetic field created by the electrons can overcome the field created by the external solenoid, whereupon the field along the axis reverses, closed field lines are set up around the electron sheet and the "magnetic bottle" is complete.



HIPAC (an acronym for heavy-ion plasma accelerator) is an example of a collective-effect application that exploits the electric field produced by a beam of electrons rather than by the magnetic field. The device, built by Richard H. Levy and his co-workers at the Avco Everett Research Laboratory, consists of a toroidal, or doughnut-shaped, vacuum chamber around which current-carrying wires are wound to generate a magnetic field of the solenoid type. The chamber is flooded with low-energy electrons, which are forced by the field to gyrate in orbits that keep them from striking the walls. A small amount of gas is introduced and becomes ionized. The positive ions find themselves trapped and oscillating in a deep electric-potential well (*lower right*) created by the surrounding sea of electrons. If the potential well could be made deep enough, the resulting ion-ion collisions would be energetic enough for nuclear transformations to occur and transuranic elements to be formed.



IN A VACUUM a relativistic electron beam is acted on by two kinds of forces. The electrostatic forces (the Coulomb attraction between like charges) tend to make the beam grow laterally, whereas the magnetic forces (the attraction between the parallel currents formed by the moving charges) tend to make the beam shrink laterally. The repulsive forces slightly exceed the attractive ones, so that the cross-sectional area of the beam slowly expands.

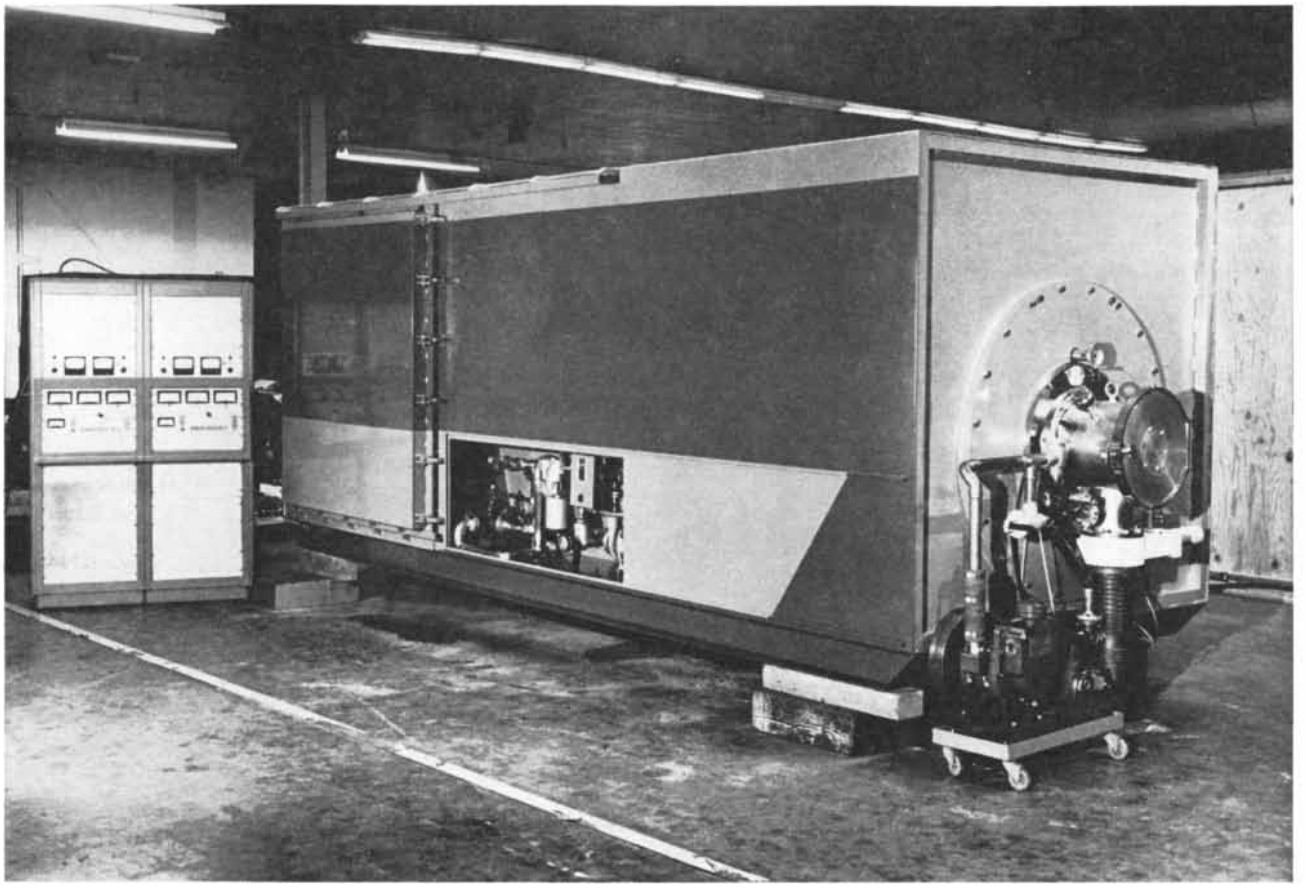


IN A GAS the relativistic electrons collide with the neutral gas atoms and in the process ionize them (*1*). After a certain length of the beam has passed a particular spot, enough positive ions are formed for them to begin to neutralize the average density of charge of the beam's electrons. As a result the beam "necks down" (*2*). At a certain limiting current the positive ions come close to totally neutralizing the negative charges of the electron beam, and the magnetic field at the edge of the beam bends the electrons sharply around (*3*).

suitable "magnetic bottle" for trapping a plasma of charged particles, with the ions and electrons spiraling along the closed lines of force that thread the coil. In practice it will not work; the plasma will quickly be lost when it strikes the windings of the coil, the mechanical supports and the leads carrying current in and out of the magnet. Moreover, in an ultimate thermonuclear reactor the coils could not survive the high temperature.

Christofilos' proposed solution to all these problems was to create a suspended "nonmaterial" solenoid made from a sheet of circulating electrons, which he calls an *E* layer [*see top illustration on preceding page*]. To make this layer he starts with a large conventional solenoid wound around a cylindrical vacuum vessel, into which he injects, off axis, an intense beam of relativistic electrons (electrons with velocities close to the speed of light). The electrons are curved by the field into orbits about a meter in diameter, and they migrate axially to form the cylindrical sheet. (A pair of extra "mirror" coils keeps the electrons from spreading too far.) If enough electrons can be stacked in the *E* layer, the magnetic field created by the electrons can overcome the field created by the external solenoid, whereupon the field along the axis reverses, closed field lines are set up around the electron sheet, and the magnetic bottle is complete. In the present large-scale experiments with Astron another factor of 10 in electron current is needed to achieve the desired result. In a small-scale version of the experiment at Cornell University field reversal was recently achieved.

Another machine in this general category is called HIPAC (heavy-ion plasma accelerator). It exploits the electric field of the electrons rather than the magnetic field. This device, built by Richard H. Levy and his co-workers at the Avco Everett Research Laboratory, has been described as a "tabletop heavy-ion cooker" [*see bottom illustration on preceding page*]. The inside of a toroidal, or doughnut-shaped, vacuum chamber is flooded with electrons of modest energy. Current-carrying wires wound around the chamber to generate a magnetic field of the solenoid type cause the electrons to gyrate in orbits that keep them from striking the walls and being lost. A small amount of gas is introduced and becomes ionized. Once an atom has lost an electron it becomes a positive ion that finds itself in a deep electrostatic potential well created by the surrounding sea of negative electrons. The positive ions held trapped and oscillating in



LINEAR ELECTRON ACCELERATOR is typical of the kind of machine currently being used to study the collective effects of very intense linear electron beams passing through gases. The accelerator was designed and built by the Physics International Company.

this potential well occasionally collide with one another. The original aim was to create a well so deep (from 10 to 20 million volts deep) that the ion-ion collisions could be energetic enough for nuclear transmutations to occur and transuranic elements to be formed. Potential wells some 10 times shallower than the one desired were achieved before the experimental work ended.

The second major category of collective-effect application includes a variety of schemes for jumping beyond the present technological limits of conventional accelerators. As early as 1956 the Russian physicist Gersh I. Budker described how an intense beam of electrons might be used to provide the guiding magnetic field for protons in a circular accelerator. One method would involve combining a low-power ring magnet with a small amount of electric acceleration; into such an arrangement one would then inject an enormously intense beam of electrons (amounting to a current of several thousand amperes) with a moderate energy (only a few MeV). In moving through a magnetic field electrons radiate electromagnetic energy by what is known as the synchrotron process, and it can be shown that this process would lead to a shrinking of the cross section of the beam. One might start, say, with an electron beam a few centimeters across and in time it would shrink to a narrow thread much smaller than a millimeter across. A quick calculation shows that the magnetic field close to the surface of such an intense filamentary beam can approach a million gauss! Even very energetic protons could be guided by such a field and circulate in the accelerator, oscillating around the thread of electrons. Thus one could achieve a very-high-energy proton accelerator in compact form. Although Budker's suggestions are now believed to be impractical in the form he proposed them, they nevertheless stimulated much new theoretical and experimental work on harnessing the collective effects of one particle beam to guide another.

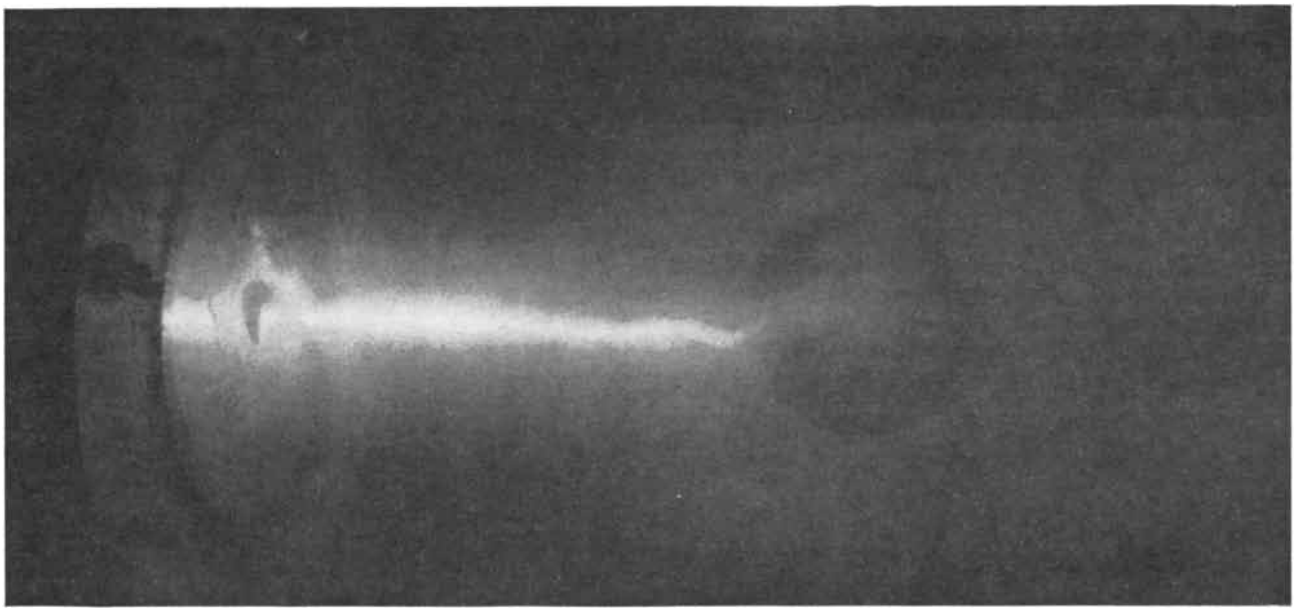
Recent research on collective fields shows much more promise of making improvements in the electric-field acceleration system than it does in the magnetic-field guide system. Accordingly the remainder of this article will be devoted to schemes designed to exploit the large electric self-fields of particle beams. These approaches divide into two main classes: those that rely on the complicated effects of a continuous intense stream of electrons and those that are based on very dense clusters of electrons. A prominent member of this second class of accelerators generates a cluster of electrons in the shape of a ring.

Many years before intense, very-high-energy electron beams became available in the laboratory there was considerable speculation about how such beams would behave in passing through a gas. It was believed that a beam of electrons in the million-electron-volt energy range could not "propagate" (that is, survive with the characteristics of a forward-moving beam) if the current exceeded a few tens of thousands of amperes. Alfvén and the British physicist John D. Lawson suggested that under certain conditions a limiting number of amperes about 34,000 times the energy of the electrons in MeV might exist; at or at about this value the collective field of the beam could literally tear it apart.

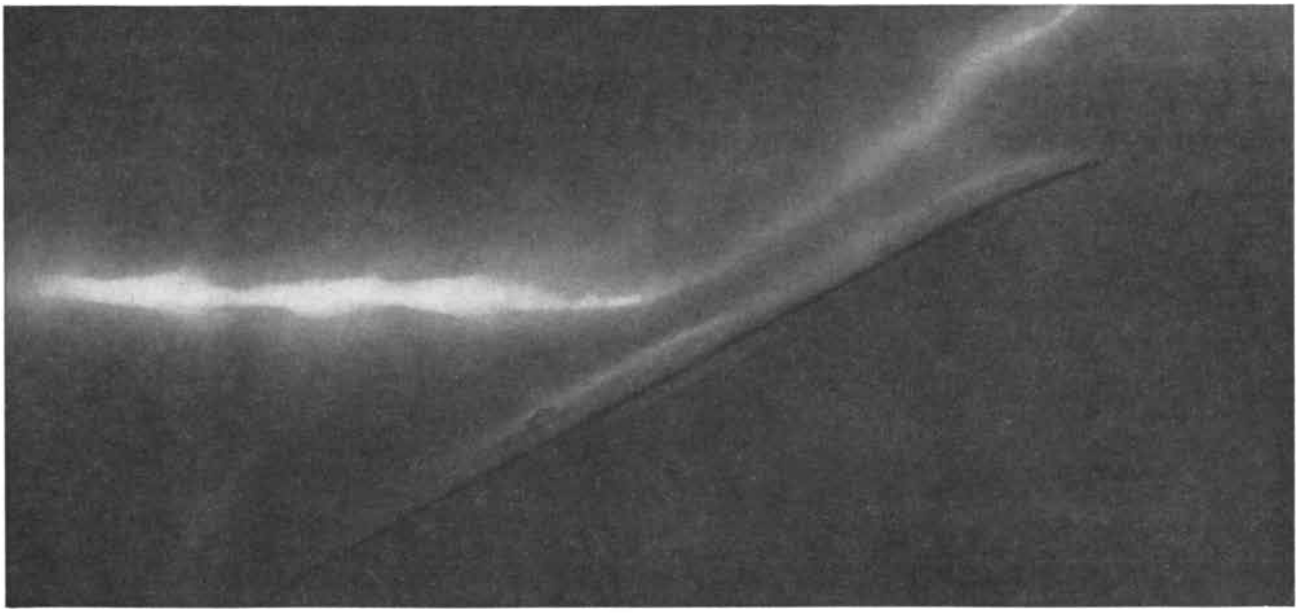
To understand the reasoning behind

their suggestion, let us examine what happens when a relativistic-electron beam leaves a vacuum and enters a region occupied by a gas. In a vacuum the electrostatic forces (the Coulomb repulsion between like charges) tend to make the beam grow laterally, whereas the magnetic forces (the attraction between the parallel currents formed by the moving charges) tend to make the beam shrink laterally. The repulsive forces always exceed the attractive ones, but only slightly in the case of a relativistic beam (at 5 MeV the magnetic attraction cancels 99 percent of the electric repulsion), so that the cross-sectional area of the beam would slowly expand.

When the electron beam passes through a gas, however, the electrons collide with the electrically neutral gas atoms and dislodge atomic electrons. The atomic electrons are promptly kicked out of the beam by the strong Coulomb repulsion, leaving comparatively immobile positive ions behind in the beam. After a certain length of the beam has passed a particular spot, enough positive ions are formed for them to begin to neutralize the average density of charge of the beam's electrons. Thus the repulsive beam-dilating forces are reduced, the attractive magnetic forces take over and the beam "necks down" to a smaller size. This "pinch effect" can be a dramatic one when the positive ions come close to



DISRUPTION of a very intense electron beam shortly after leaving a vacuum and entering a gas is clearly visible in this photograph. The only illumination used was that produced by the recombination of ionized gas molecules in the path of the beam. The photograph, made with a Physics International linear electron accelerator, corresponds to step *3* in the bottom illustration on page 26.



CURIOUS BEHAVIOR of a very intense, high-energy electron beam in a low-pressure gas is captured in this photograph, also from Physics International. When a tilted metal plate is placed in the path of such a beam, intense electric and magnetic "images" in the metal lead to strong deflecting forces, which make it appear that the electron beam bounces off the plate without actually striking it.

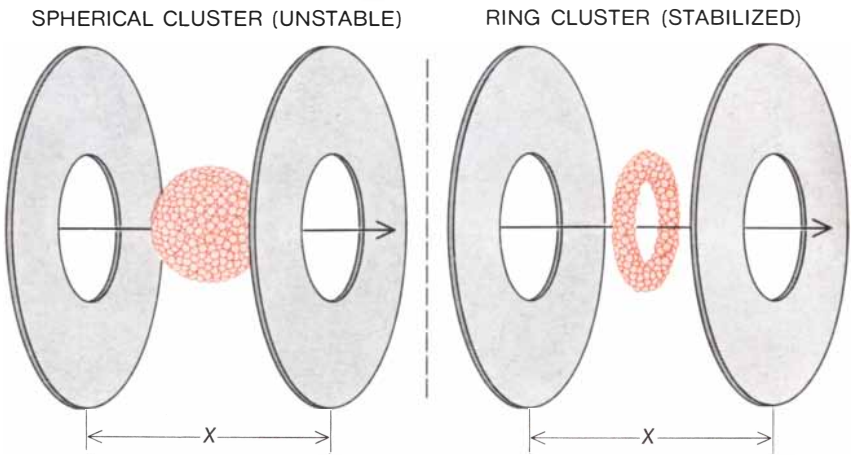
totally neutralizing the negative charges of the electron beam. At the Alfvén-Lawson limiting current the magnetic field at the edge of the beam is unleashed in full force, bending a highly relativistic-beam electron sharply around through 180 degrees within the diameter of the beam itself!

Experimental studies of these effects have become possible in the past few years with the development of very-high-current electron accelerators. Relativistic-electron currents of more than a million amperes have recently been generated. The self-destructive effects of the enormously strong collective fields associated with such beams have indeed been observed but it has also been found that they can be tamed. Beam currents 10 times higher than what was believed to be the limiting current have been found to propagate successfully. These achievements can be understood in terms of a more complex analysis, in which the current due to the motion of the slow electrons and ions in the beam causes additional neutralizing effects.

Working on such pinch phenomena with a 40,000-ampere beam of 1.3-MeV electrons at the Ion Physics Corporation, Stuart Graybill and John Uglum were surprised to find early in 1969 that copious quantities of very energetic ions were being generated when the beam entered a short tube of gas. By filling the tube with different gases they could produce a pulsed beam of monoenergetic protons, deuterons, helium ions or nitrogen ions. Further experiments that confirmed and extended this work were done at the Physics International Company by Gerold Yonas and his co-workers; they established that very strong electric fields (50 million volts per meter) must exist near the beam front and accelerate the ions. These observations are a clear and dramatic demonstration of ion acceleration to useful energies by collective fields that greatly exceed conventional laboratory electric fields.

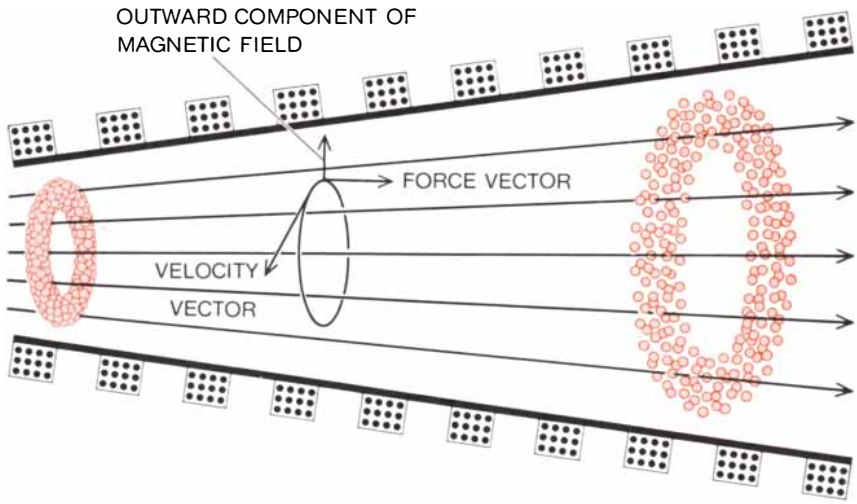
There is still no clear understanding of how this field arises or how it remains in phase with the ions. One suggestion, made by Norman Rostoker of Cornell University, is that the changing distribution of ionization leads to an electric-potential well that propagates rather slowly in the direction of the electron beam. Another idea, put forward by Sidney Putnam of Physics International, is concerned with the electric field that is induced where the beam pinches down (as a consequence of the increasing magnetic energy surrounding the beam); the field has a direction such as to decelerate the electron beam and accelerate the positive ions. Solution of the puzzle awaits more experiments. Certain experimental observations suggest a high degree of order in the mechanism, and it is conceivable that in a single beam the process could be repeated to generate much higher proton or ion energies. Some methods have been proposed to create a longitudinal density modulation (in effect a local potential well) by the application of external fields that make the modulation propagate in a controlled way.

It is a notable feature of collective-

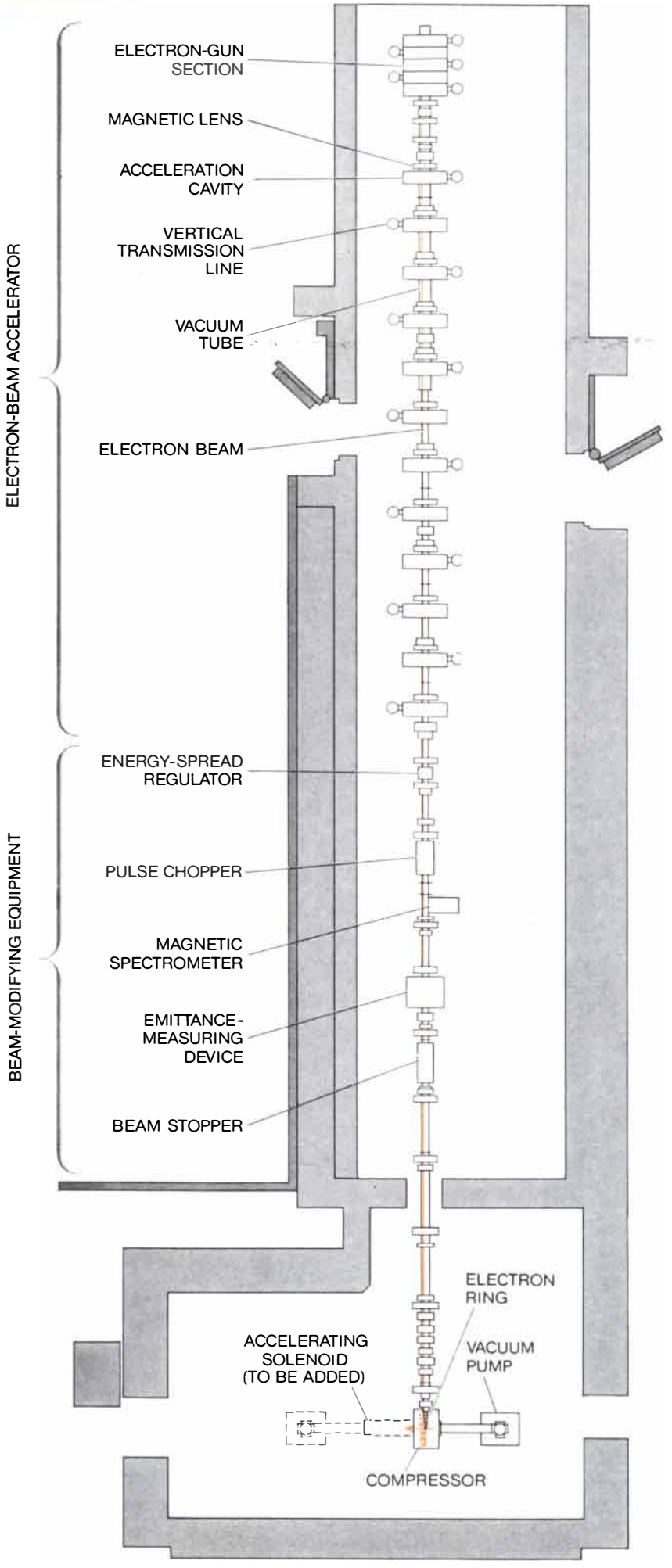


SPHERICAL CLUSTER of electrons can be accelerated in an electric field of x volts per meter generated by a pair of ring-shaped electrodes. The energy gain of each electron would be x electron volts per meter. The energy gain of a proton embedded in such a cluster, however, would be 1,836 times x electron volts per meter, since for a given speed change the energy change is directly proportional to the mass of the particle, and a proton is 1,836 times more massive than an electron. Unfortunately the large electric field that keeps the proton captured tends to blow the electrons in the cluster apart, and such a compact structure could not survive.

RING-SHAPED CLUSTER of electrons rotating at relativistic speeds in the same electric field would be self-stabilized, since the magnetic attraction of the parallel electron currents can be made to cancel practically all the electrostatic repulsion. (A few added positive ions do the rest.) When such a ring is accelerated at right angles to its plane, however, relativistic effects make the rotating electrons appear some 23 times heavier than they are when they are at rest. Hence a proton carried along inside a ring-shaped cluster of rotating relativistic electrons would undergo an energy gain of $(1{,}836/23)x$, or about $80x$, electron volts per meter.



UNEXPECTED ADVANTAGE in using the electron ring as a vehicle for accelerating positive particles is obtained when the ring is introduced into a magnetic solenoid whose axial field becomes weaker with distance. The combination of the rotational motion of the electrons and the small outward component of the magnetic field produces a force to accelerate the ring along the solenoid, thereby avoiding the need for electric accelerating systems. In the process the ring increases in size and the energy of the electrons decreases.



GENERAL LAYOUT of the electron-ring accelerator built by the author's group at the Lawrence Berkeley Laboratory is given here. The linear induction accelerator used to inject the electrons into the compressor generates an electron beam with an energy of 4.25 million electron volts (MeV) and an intensity of 1,000 amperes. The 17 acceleration cavities are powered by the vertical transmission lines. The equipment just in front of the wall modifies various properties of the beam before delivery to the compressor. A solenoid with a tapered magnetic field will be added soon for ion acceleration.

effect accelerators that they make it possible to accelerate simultaneously charges of several species. This cannot be done in conventional resonant accelerators. Their adjustments depend critically on the exact value of the charge-to-mass ratio of the particles that are to be accelerated.

The attractiveness of being able to accelerate a cluster of close-packed electrons can be judged from the following consideration. Imagine a spherical cluster of electrons, say one centimeter in diameter and consisting of 10^{13} electrons, into which a small number of protons are introduced, for instance by the ionization of neutral hydrogen gas. The protons find themselves trapped in a deep potential well and would have to overcome enormous fields—on the order of 600 megavolts per meter—to escape. If this cluster is accelerated in an electric field x volts per meter generated by either a pair of electrodes or a radio-frequency cavity, the electrons each gain an energy of x electron volts in a meter. Each proton carried along inside the cluster, however, gains energy at the rate of 1,836 times x. The reason is that for a given speed change the energy change is directly proportional to the mass of the particle, and a proton is 1,836 times more massive than an electron. If one were to accelerate a "bare" proton (one not embedded in the cluster), its energy gain would be simply x electron volts. Here the reward for utilizing collective effects instead of ordinary electric fields would be a gain in energy by the enormous factor of 1,836.

Unfortunately this example refers to an impractical case. The large electric field that keeps the proton captured also tends to blow the electrons in the cluster apart, and the compact structure could not survive. It is true that for a fast-moving cluster lateral stability could be achieved by adding enough ions to cause some pinching, but in the direction of motion the large electric

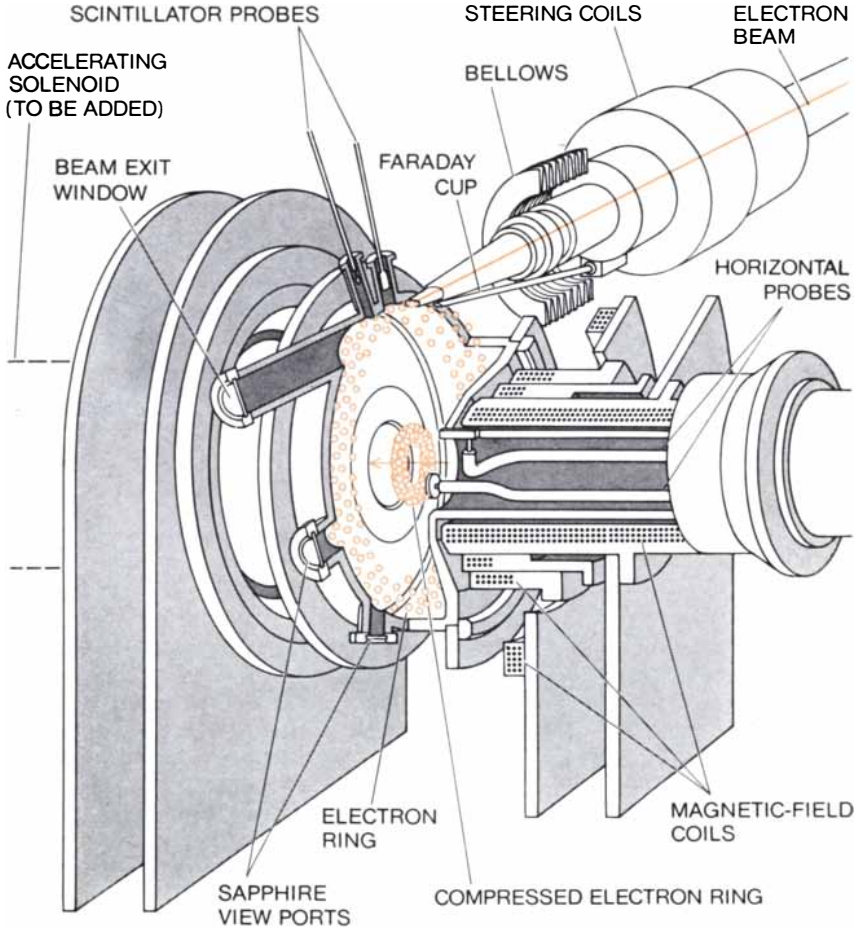
field acts with full potency to spread the cluster. One can now perceive the beauty of using a ring of rotating relativistic electrons as a self-stabilized cluster. When one considers a beam of 10-MeV or 20-MeV electrons curled into a ring by a suitable magnetic field, it follows from the earlier discussion of straight beams that the magnetic attraction of the parallel electron currents can be made to cancel something like 99.9 percent of the electrostatic repulsion. Thus a small amount of seeding with protons (more than .1 percent as many protons as electrons) will eliminate the last bit of repulsion and pinch the minor dimensions of the ring. The seeding leads to a self-stabilized cluster in which the deep potential well is now distributed in the torus occupied by the circulating electrons.

There is, however, a penalty for this type of stabilization. When the ring is accelerated at right angles to its plane, relativistic effects make the rotating electrons appear to be heavier than they are when they are at rest. Therefore the ring cluster responds to the externally applied field as if it were made up of "heavy" electrons with a mass higher by typically a factor of between 20 and 40. Accordingly the full factor of enhancement of 1,836 cannot be gained; analysis of all the details leads one to expect a peak enhancement closer to 80. This gain is still worth striving for. With applied fields of 10 million volts per meter, proton acceleration at rates of 80 times 10, or 800, MeV per meter could be reached, which would be a very large step beyond the best synchrotron.

There is an unexpected boon in utilizing the electron ring as an accelerating vehicle because it makes it possible to accelerate ions with rather simple magnetic fields alone; the complexity of introducing an electric accelerating system is avoided. When the ring is introduced into a magnetic solenoid whose axial field becomes weaker with distance (in other words, a solenoid in which the field lines flare slightly outward), the combination of the rotational motion of the electrons and the small outward component of magnetic field produces a force to accelerate the ring along the solenoid [*see bottom illustration on page 29*]. In the process the ring increases somewhat in size and the energy stored in the rotating electrons decreases as it is transferred into the longitudinal energy of the accelerating ring and ions. "Magnetic expansion" acceleration should be feasible for producing protons with an energy of about 1,000 MeV or heavy ions with about 500 MeV per nucleon. Accelerators in this range of energies can be important for pion physics, biological work, cancer therapy and the nuclear chemistry of heavy and superheavy elements.

The idea of harnessing rings of relativistic electrons to accelerate protons and positive ions was originally conceived by the Russian physicist Vladimir I. Veksler. He and his group at Dubna in the U.S.S.R. had worked for several years on the practical aspects of the creation of suitable electron rings before their results were revealed four years ago. Physicists in other laboratories were surprised and excited at the promise offered by such rings, whose properties and behavior were in large measure amenable to detailed theoretical calculation. Work on electron rings soon began elsewhere. In the U.S. two groups have since been organized, our group at the Lawrence Berkeley Laboratory and another group at the University of Maryland. In addition experiments have been performed at Garching and Karlsruhe in West Germany, and more recently smaller groups have begun work in France, Italy and Japan. Meanwhile in the U.S.S.R. another group has just begun a large-scale program on electron-ring accelerators at the Institute of Theoretical and Experimental Physics in Moscow.

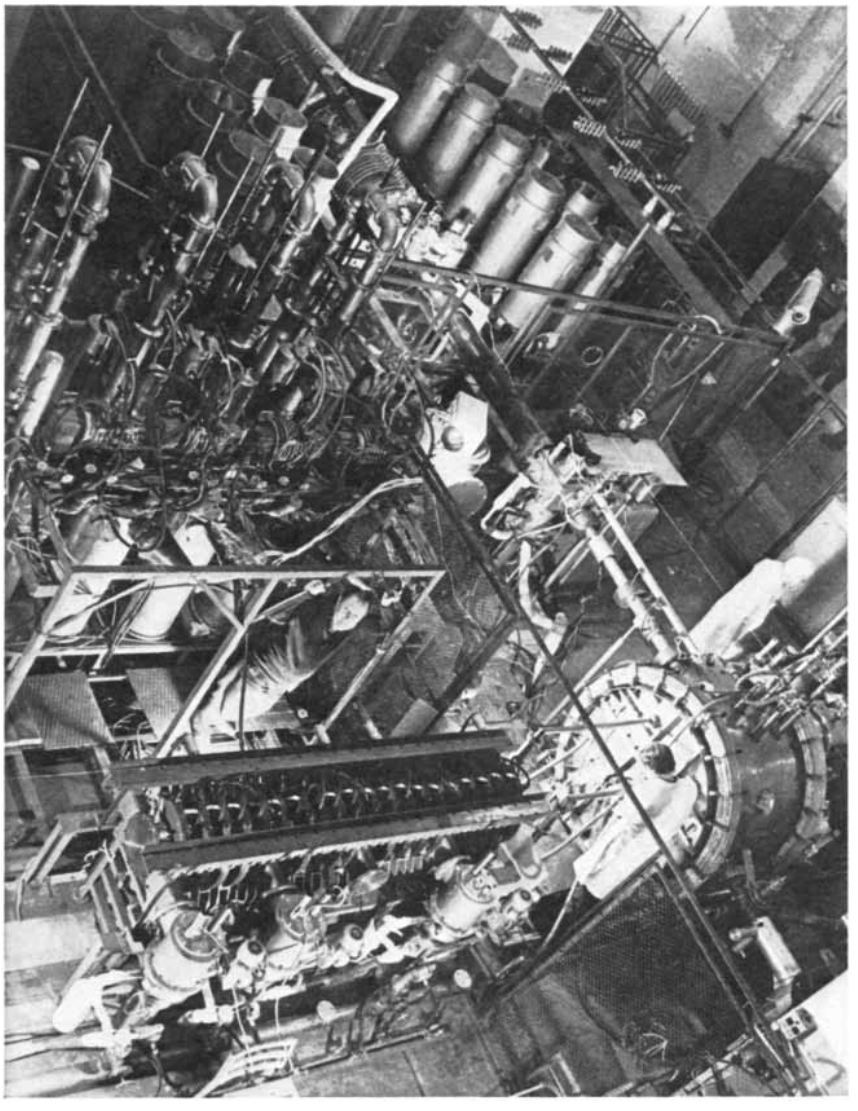
In principle one could imagine forming a ring with the right properties (a major radius of about four centimeters and a minor radius of about a millimeter) by injecting some 10^{13} electrons at about 15 MeV into a magnetic field that would curl them into a ring. In



CUTAWAY VIEW OF COMPRESSOR reveals the interior of the vacuum chamber where the electron ring is first formed (*large colored ring*) and then compressed (*small colored ring*). Various remotely controlled mechanisms for inserting probes into the interior of the chamber can also be seen. The magnetic field in the compressor is supplied by a nested set of copper coils. At the end of the compression stage a small puff of hydrogen (or some other gas) will be released into the compressor. The circulating electrons will ionize some of the gas, and the protons (or heavier ions) will remain trapped inside the potential well of the ring. The current will then be raised in the coils on one side and lowered in the coils on the other side, squeezing the ring sideways out of the compressor until it can enter a solenoid with a tapered field to provide some initial magnetic-expansion acceleration.

practice this cannot be achieved today. The following approach is commonly taken. A much larger ring, its dimensions approximately those of a bicycle inner tube, is first formed in a vacuum vessel called a compressor. There are several electron injectors in existence throughout the world that can provide the current (hundreds of amperes) and the energy (between 2 and 4 MeV) needed to make these rings. The magnetic field in the compressor is supplied by a nested set of copper coils. After the original large ring has been injected the magnetic field is rapidly raised in a period on the order of a tenth of a millisecond. This procedure leads to two effects. First, the circulating electrons gain greatly in energy through the transformer action of the changing magnetic field. Second, both the major radius and the minor radius of the ring shrink substantially. If the right parameters are chosen, the process of compression will transform the original large, fat, low-energy ring into the small, slender, high-energy ring that is to be the vehicle for accelerating the protons and heavier ions.

At the end of the compression stage a small puff of hydrogen (or some other gas) is released into the compressor. The circulating electrons ionize some of the gas, and the protons (or heavier ions) remain trapped inside the potential well of the filamentary ring, thereby supplying the needed self-stabilization. Now that the ring has been properly prepared, the next step is a delicate manipulation of the magnetic field to extract the ring from the compressor and launch it into the accelerating column. This manipulation is done by raising the current in the coils on one side of the compressor while lowering the current in the coils on the other side, the net result being to squeeze the ring sideways out of the compressor until it can enter a short solenoid with a tapered field to provide some initial magnetic-expansion acceleration. After this stage one has the choice of either continuing with added solenoids to give more expansion acceleration or introducing electrodes or radio-frequency cavities to accelerate the ring electrically.

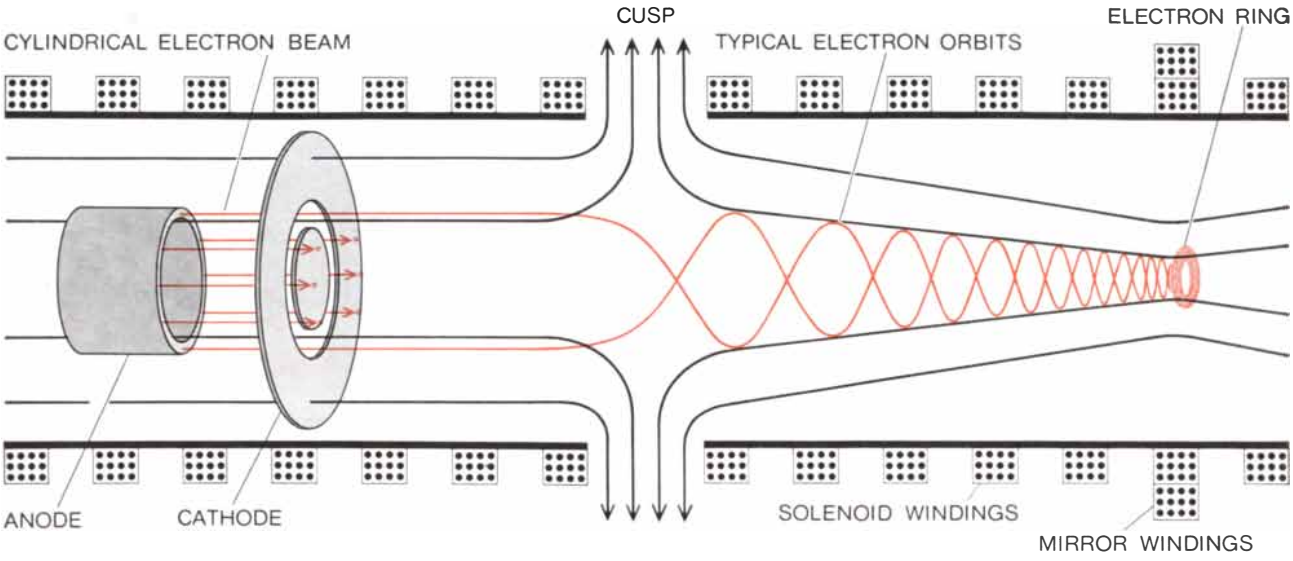


FIRST ELECTRON-RING ACCELERATOR, located at Dubna in the U.S.S.R., was reportedly used in a recent experiment to accelerate helium ions (alpha particles) to an energy of 30 MeV in a short tapered solenoid. The general layout of the machine is similar to that of the electron-ring accelerator at Berkeley, with the electron injector at top, the compressor at lower right and ring-acceleration equipment at bottom left. When this photograph was made, a system of radio-frequency cavities was being tested for ring acceleration downstream of the tapered solenoid. Experiments performed with this machine have demonstrated the feasibility of electron-ring acceleration at energies of interest for nuclear chemistry.

Experimental research on electron-ring accelerators has made significant progress in the past few years. Compressed rings with suitable properties for ion acceleration have been formed and studied at Dubna, Berkeley and Garching. In addition the Russian group has presented convincing evidence of having accelerated helium ions (alpha particles) to an energy of 30 MeV in a short tapered solenoid. This corresponds to achieving a very substantial collective accelerating field of 40 million volts per meter. Now they are attempting to transport the ring through a set of radio-frequency cavities and to accelerate it further with electric fields.

An early attempt by our group at Berkeley to extract and accelerate rings magnetically was thwarted by the collective instabilities of the ring. I mentioned these destructive collective effects above in connection with the comparatively weak beams in conventional accelerators; for the enormously intense currents in a typical electron ring they can be much more troublesome. In the past several months work at Berkeley has concentrated on the study and taming of these instabilities. Understanding and control have been achieved for the transverse instability, but the negative-mass effect still presents some mysteries. It is exceedingly important to understand quantitatively how to eliminate such instabilities, because they will ultimately set limits on how far one can go in achieving high-quality electron rings with deep potential wells.

The acceleration of alpha particles by the Dubna group is a demonstration of the feasibility of the concept of electron-ring acceleration at energies that are al-



PROPOSED ALTERNATIVE SCHEME for forming a ring of rotating relativistic electrons is currently being studied by investigators at the University of Maryland. A hollow cylindrical sheet of electrons is injected into a cusp-shaped magnetic field formed by placing end to end two solenoids with opposing fields. Individual electrons in the hollow beam (*colored arrows*) stream down the first solenoid parallel to the magnetic-field lines (*black arrows*) until they come to the cusp region, where they are strongly bent sideways and give up most of their forward motion in favor of rotational motion. An extra field winding (*right*) makes a magnetic mirror, which effectively provides a backstop against which the rotating electrons pile up to form a ring. Once a suitable ring has been stacked the magnetic mirror can be switched off to allow the ring to move farther down the solenoid into a region of accelerating field.

ready of interest for nuclear chemistry. Questions remaining to be answered by further experiments include: Are there unexpected troubles in the electric acceleration process? How far can one push to make rings with higher collective accelerating fields before new instabilities arising from ion-electron interactions become a problem? Answers should come fairly soon on these matters and should enable us to evaluate better the promise and limits of this novel method of harnessing collective effects for particle acceleration.

It is important in an electron-ring accelerator to make sure that the ring is not accelerated so vigorously that the ions are jerked out of the ring. Although the ions are held in the deep potential well of the ring by strong electric forces, too rapid a change in the guiding magnetic field can still cause the ions and electrons to part company. It is not difficult, however, to construct magnetic-field coils within the tolerances needed. A corollary is that at the end of the accelerator it is easy to separate the ion beam from the carrier electrons. The ring can be passed through a moderately strong magnetic lens that will disrupt the ring, focus the positive ions on a target and sweep the electrons out of the way.

If electron rings are used to accelerate protons to very high energies for the study of elementary-particle interactions, the secondary beams of mesons, hyperons or whatever from the target will have one distinctive feature. They will emerge in a bunch with a very short duration: less than a trillionth of a second. For certain types of experiment this is a disadvantage and can lead to substantially reduced collection rates. On the other hand, it can make other experiments simpler and can open up new possibilities such as making energy measurements on uncharged particles by time-of-flight techniques. In addition it should be straightforward to accelerate deuterons, other ions and possibly polarized protons to high energies.

There has been a proliferation of theoretical proposals for ring-compression schemes other than that based on the pulsed-magnetic-field apparatus described above. A particularly attractive idea, suggested in one form by Christofilos and in another by L. Jackson Laslett and Andrew M. Sessler, would be to utilize static-field compressors, in which compression could be achieved without resorting to a time-varying magnetic field, perhaps making it possible to produce rings at a very high rate of repetition.

Another novel scheme, on which experiments have been started at the University of Maryland, relies on injecting a hollow cylindrical sheet of electrons into a cusp-shaped magnetic field formed by placing end to end two solenoids with opposing fields [*see illustration above*]. Individual electrons in the hollow beam stream down the first solenoid parallel to the magnetic-field lines until they come to the cusp region, where they are strongly bent sideways and give up most of their forward motion in favor of rotational motion. An extra field winding makes a magnetic mirror, which effectively provides a backstop against which the electrons pile up to form a ring. In order to create a ring that is not too diffuse it is essential that ions be introduced during the process, so that the pinch effect plays a role in compressing the ring laterally. Once a suitable ring has been stacked the magnetic mirror can be switched off to allow the ring to move farther down the solenoid into a region of accelerating field.

The past few years have seen the first exciting and significant steps into the vast new territory of acceleration of particles by collective effects. The utility of the powerful electric and magnetic fields produced by intense electron beams in gases and by electron rings in a vacuum has been demonstrated in a conclusive way, and collective effects are probably not far from being harnessed for practical purposes. The continuing work on these two principal collective methods will surely lead to future efforts to turn collective effects to good advantage by exploiting one or another of the more exotic ideas mentioned only briefly in this article.

AN EARLIER AGRICULTURAL REVOLUTION

It has generally been assumed that man first domesticated plants and animals in the Middle East. Excavations in Southeast Asia now suggest that there the revolution began some 5,000 years earlier

by Wilhelm G. Solheim II

Agriculture is known to have been invented at least twice. Plants and animals were domesticated in the Old World and the process was repeated quite independently some millenniums later in the New World. Evidence of a third domestication has now been uncovered. The agricultural revolution, which was thought to have first occurred some 10,000 years ago among the emerging Neolithic societies of the Middle East, seems to have been achieved independently thousands of miles away in Southeast Asia. This separate agricultural revolution involved plants and animals for the most part unknown in the Middle East, and it may have begun as much as 5,000 years earlier.

The fact that some of the most technologically advanced cultures in the world in the period from about 13,000 B.C. to 4000 B.C. flourished not in the Middle East or the adjacent Mediterranean but in the northern reaches of mainland Southeast Asia is not easy to accept. Nonetheless, recent excavations in Thailand have convinced my colleagues and me that somewhere among the forest-clad mountains of the region man's first tentative efforts to exploit wild plants and animals opened the way first to horticulture and then to full-scale agriculture and animal husbandry.

In terms of prehistory mainland and island Southeast Asia are together the largest unknown region in the world. It is only in the past decade that they have been recognized as rewarding areas for archaeological investigation. My own work in the region began in the Philippines in 1949 and has since included investigations both in other island areas and on the mainland. The need for salvage archaeology in advance of construction work in the lower basin of the Mekong River resulted in the excavation of two sites in northern Thailand by my colleagues and me in 1965–1966. The first of these sites, a prehistoric mound called Non Nok Tha, is close to the area that was flooded by the waters of the Nam Phong Reservoir. Our work there, sponsored jointly by the Thai Government Department of Fine Arts, the University of Otago in New Zealand and my department at the University of Hawaii, began during the 1965–1966 dry season with the support of the National Science Foundation. Simultaneously my department sponsored a second excavation at a promising site in the northwest corner of Thailand. This was Spirit Cave; its discoverer and excavator, Chester F. Gorman, was a student of mine at the University of Hawaii.

The mound called Non Nok Tha was tilled jointly by a few of the farmers of a nearby hamlet. They had almost covered it with plantings of bananas, chili peppers and mulberry bushes, but there were two cleared areas on the mound where the rice from the surrounding paddy fields was threshed at harvesttime. Digging at the edge of a mulberry planting, we found that the uppermost strata of the mound contained a few iron implements and numerous remains of individuals who had been cremated before burial. Below these levels were earlier graves whose occupants had not been cremated; the burials included various grave furnishings. Among them were stone molds for casting bronze axe blades, some of the axes themselves, other bronze objects and a number of polished stone tools. The graves also held pottery that was well made but generally lacked decoration. In the lowest levels of the site we found still more polished stone tools and a few specimens of decorated pottery but no bronze at all.

These were significant discoveries. First, never before had a site been found in Southeast Asia that contained evidence of a substantial interval when bronze was known but iron was not. Second, the pottery from the lowest strata included two vessels decorated with a pattern that I had first seen in the Philippines and had later observed on pottery from excavations elsewhere in Southeast Asia.

The Non Nok Tha site proved to be so rich in burials that our work went more slowly than we had expected. Eventually we pieced together more than 100 pottery vessels for the Thai National Museum, sent samples off for carbon-14 analysis and packed up a number of other finds for laboratory study back in Hawaii. Several of the site's deepest burials we left untouched but mentally reserved for later excavation. We were convinced that the lowest levels of the mound were quite old, going back perhaps as far as 1000 B.C., but we could not be certain until the carbon-14 results were known.

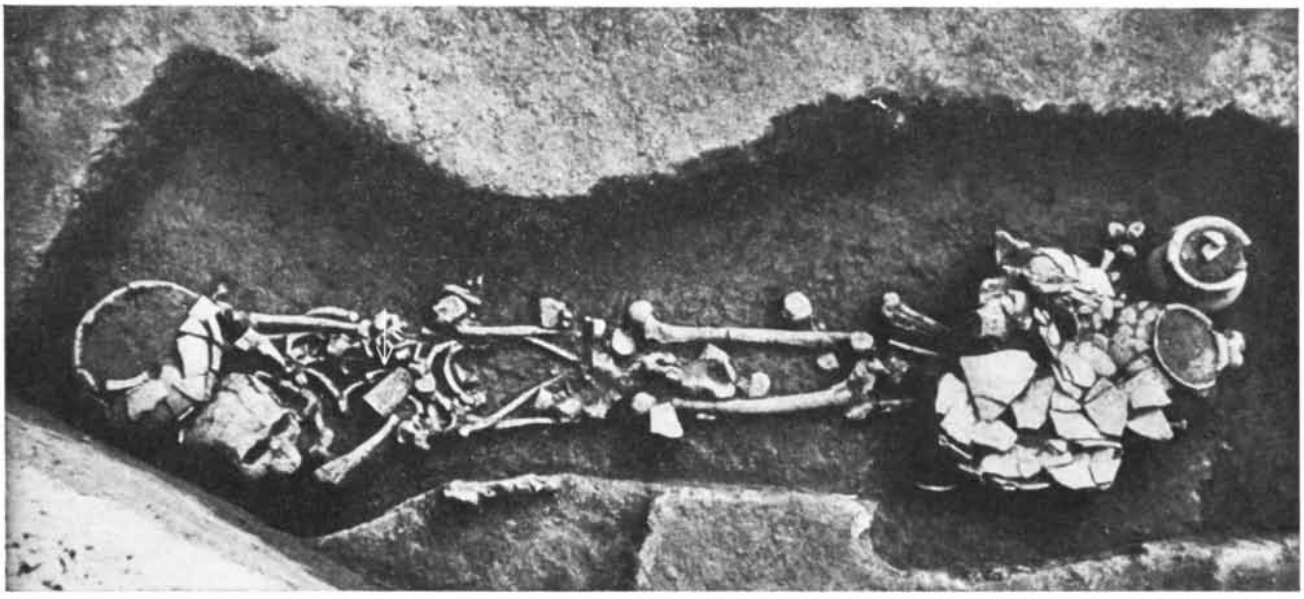
As we unpacked, cleaned and examined the specimens in Hawaii we continued to find the unexpected. For example, we found among the potsherds from the lowest levels some that bore the imprint of cereal grains and husks. It appeared that the cereal might be rice, and we sent a number of samples to an authority on cereal grains, Hitoshi Kihara of the Kihara Institute for Biological Research in Japan. Kihara concluded that the cereal had been *Oryza sativa*, the common species of rice that is grown throughout Asia today.

We had also found animal bones in many of the burials at Non Nok Tha. They gave the impression that portions had been cut from large animals and placed in the graves. Charles Higham of the University of Otago offered to identify the bones. He found them indistin-

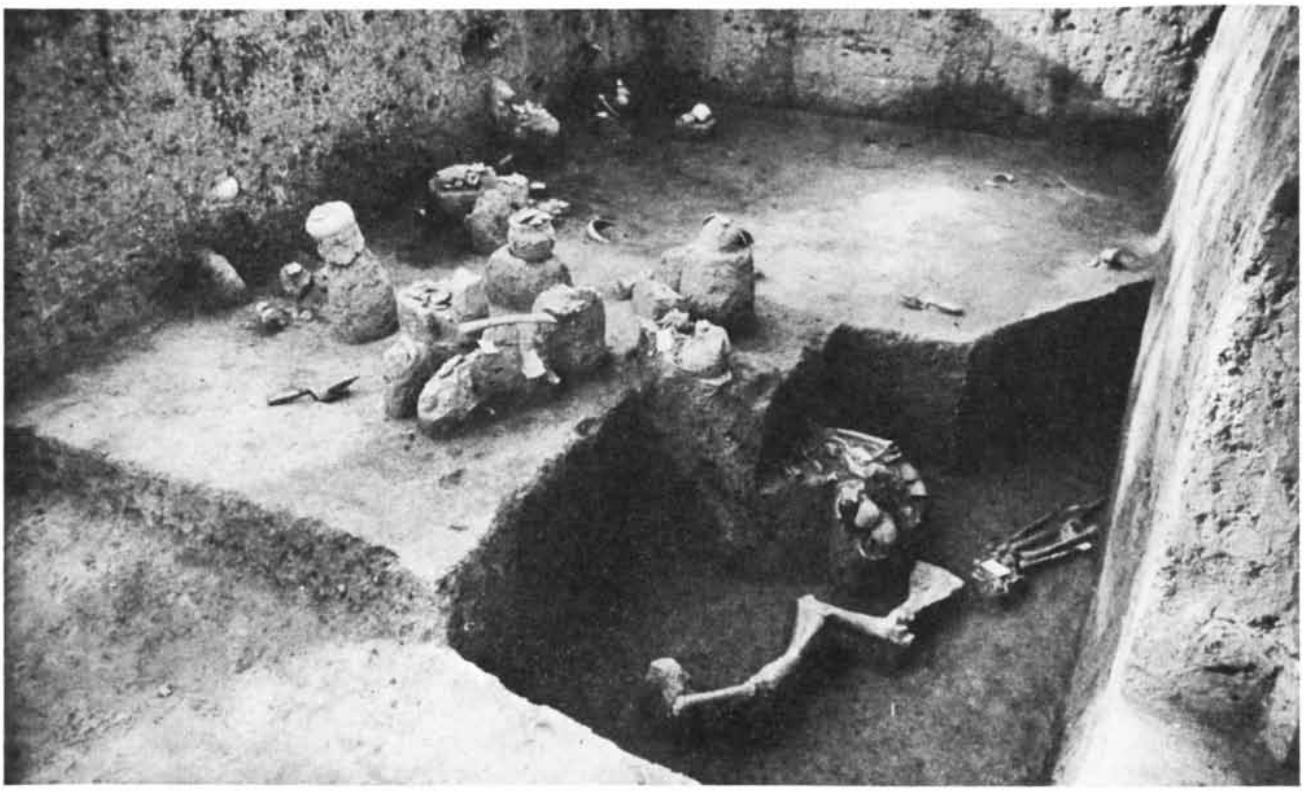
guishable from the bones of *Bos indicus*, the common species of humped cattle from India.

At last we received the carbon-14 results. They showed that the mound was much older than we had anticipated. The lowest levels we had excavated at the site had been deposited before 3000 B.C., some of them possibly before 4000 B.C. None of the uppermost strata, however, was more than 1,000 years old.

Early in 1968 Donn T. Bayard of the University of Otago was able to excavate in a new area at Non Nok Tha. He opened several deep graves; one contained a remarkably well-preserved human skeleton and, resting on the skeleton's chest, a socketed tool made of copper. This body must have been buried sometime during the fourth millennium B.C., which makes the copper implement made of copper (*arrow*). The copper contains traces of arsenic and phosphorus, suggesting that, if the tool was not actually a casting, the metal had at least been heat-treated by roasting the ore to prepare it for cold hammering. No older socketed tool is known.



HUMAN BURIAL of the fourth millennium B.C. was unearthed in 1968 at Non Nok Tha, a rich archaeological site in northern Thailand, by Donn T. Bayard of the University of Otago in New Zealand. In addition to a skeleton the burial contained a socketed implement made of copper (*arrow*). The copper contains traces of arsenic and phosphorus, suggesting that, if the tool was not actually a casting, the metal had at least been heat-treated by roasting the ore to prepare it for cold hammering. No older socketed tool is known.



ANIMAL LIMB BONES (*right foreground*) are typical of the many aggregates of animal bone, some of them more than 5,000 years old, found in burials at Non Nok Tha. They are probably the bones of the familiar species of Indian humped cattle, *Bos indicus*.

plement the oldest-known socketed metal tool in the world.

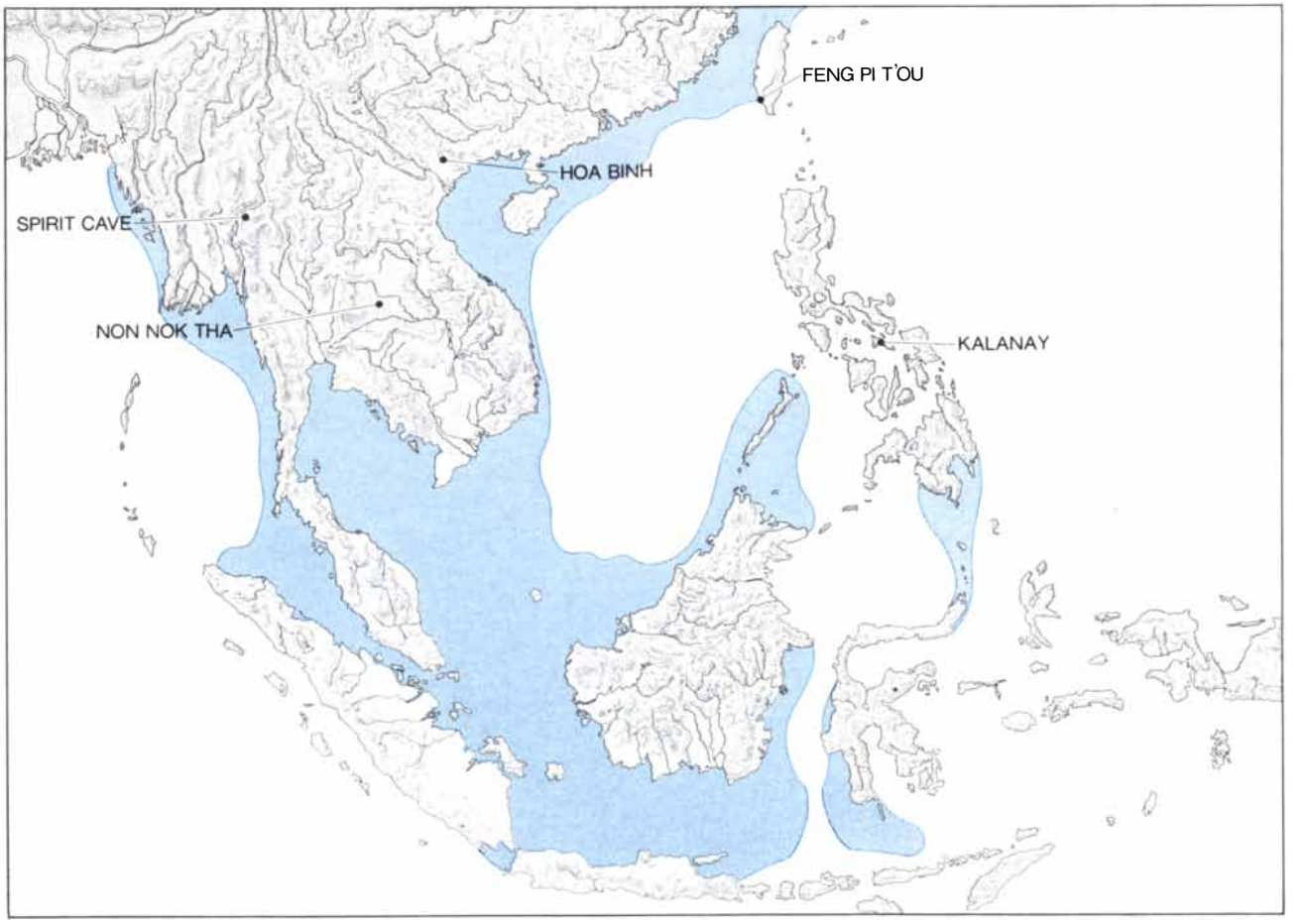
In summary, the findings at Non Nok Tha included several surprises. They showed that rice was an established cereal and that humped cattle were commonplace in mainland Southeast Asia perhaps as early as the fifth millennium B.C. Bronze metallurgy had flourished in the region for a considerable interval before the appearance of iron, and a unique form of copper tool had appeared on the scene even earlier. None of these developments had taken place in a vacuum: some of the earliest pottery in this corner of the mainland bore designs that are found elsewhere in mainland and island Southeast Asia. Even bigger surprises, however, were in store at Spirit Cave.

Gorman's site lies high up on the face of a limestone cliff that overlooks a small stream not far from the border between Thailand and Burma. The cave was last occupied, Gorman found, about 5600 B.C. The levels that underlie the top level span more than 4,000 years, and the earliest strata were formed about 10,000 B.C. The artifacts from all these occupation levels (except for those in the levels of the last 1,200 years or so) proved to be very much the same: simple stone tools representative of a Southeast Asian hunters' and gatherers' culture. The culture is called Hoabinhian because the cave sites where its simple artifacts were first unearthed during the 1920's are in the mountains near the town of Hoa Binh in North Vietnam. Since the initial discovery, many other sites containing Hoabinhian artifacts have been found in the mountainous terrain of northern Southeast Asia, almost always in small caves not far from streams. At first sight Gorman's cave looked like one more Hoabinhian find, remarkable only for being farther west than any previously discovered.

Two findings soon demonstrated the importance of Spirit Cave. The first was the result of Gorman's careful search for plant material. By sifting all the soil as it was removed from his test trenches, Gorman recovered the fragmentary remains of 10 separate plant genera, among them pepper, butternut, almond, candlenut and betel nut. Of even greater significance was evidence of a species of cucumber, a bottle gourd, the Chinese water chestnut and certain legumes: the pea (*Pisum*), either the bean or the broad bean (*Phaseolus* or *Vicia*) and possibly also the soybean (*Glycine*). Some of these plant remains were present in all levels at the site.

It seems probable that several of the plants were cultivated by the inhabitants of Spirit Cave. The relatively small number of specimens that Gorman recovered in 1966, however, does not make it possible to draw a clear-cut conclusion on this crucial point. Last year Gorman did further work at Spirit Cave and recovered additional specimens. They are now being analyzed by Douglas Yen of the Bernice P. Bishop Museum in Honolulu, so that the question may soon be settled.

Let us assume for the sake of argument that Yen's finding will be negative



MAJOR EARLY SITES in Southeast Asia include Hoa Binh, in North Vietnam, where remains of a hunters' and gatherers' culture were found during the 1930's, and another Hoabinhian site, Spirit Cave in Thailand, where the earliest-known evidence of farming may have been found. Pottery at Spirit Cave, among the oldest in the world, resembles later wares found on Formosa. A second site in Thailand, Non Nok Tha, not only yielded a uniquely early kind of metal tool but also contained pots with designs like those found on Malayan and Philippine pottery. Lower seas in the past greatly enlarged the land area; color shows one estimate of the increase.

and that the plants unearthed at Spirit Cave are merely wild species that had been gathered from the surrounding countryside to supplement the hunters' diet. What should we then conclude? Such a finding would be evidence of an advanced stage in the utilization of wild plants in Southeast Asia, a stage that in my opinion is at least as early as any equivalent stage known in the Middle East. On the other hand, what if Yen's finding is positive? This would be evidence that the inhabitants of Spirit Cave were engaging in horticulture at least 2,000 years before the date that has been suggested for the first domestication of plants in the Middle East.

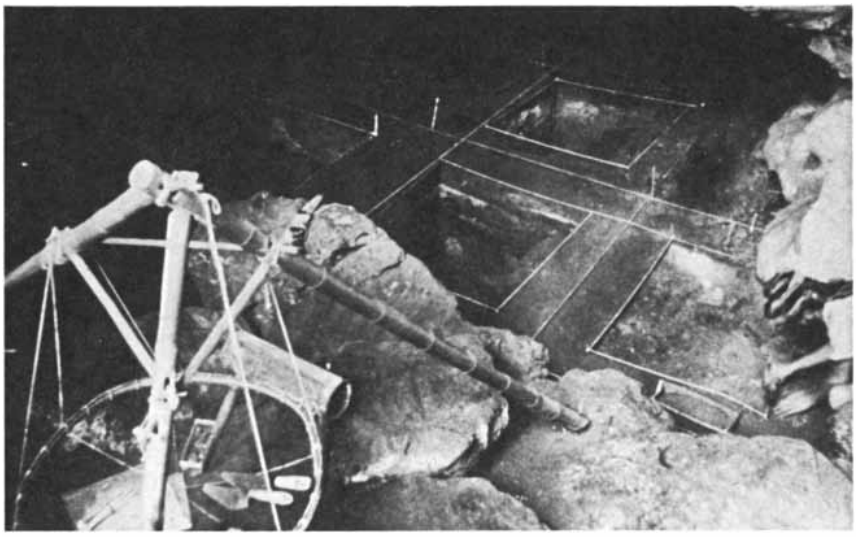
Gorman's next major finding was evidence that a clearly non-Hoabinhian influence had reached Spirit Cave about 6800 B.C. This was indicated by the presence of new and distinctive artifacts among the simple Hoabinhian tools, including rectangular stone adzes with partially polished surfaces and knives made by grinding both sides of a flat piece of slate to form a wedge-shaped cutting edge. Tools such as these are unknown anywhere else in mainland Southeast Asia at such an early date.

Even more important was the presence in the upper strata of pottery fragments that showed a variety of finishes. Most of the potsherds bore imprints that were probably produced by striking the soft clay with a cord-wrapped paddle before the pot was fired. Other potsherds show evidence of polishing, still others had been incised by cutting into the soft clay with a comblike tool, and a few are coated with what appears to be a resin glaze. Except for certain extremely early examples of pottery that have been unearthed in Japan, these fragments from Spirit Cave are the oldest pottery in the world. Moreover, the variety of decoration and finish they display testifies to the existence of a sophisticated potter's tradition with a considerable past.

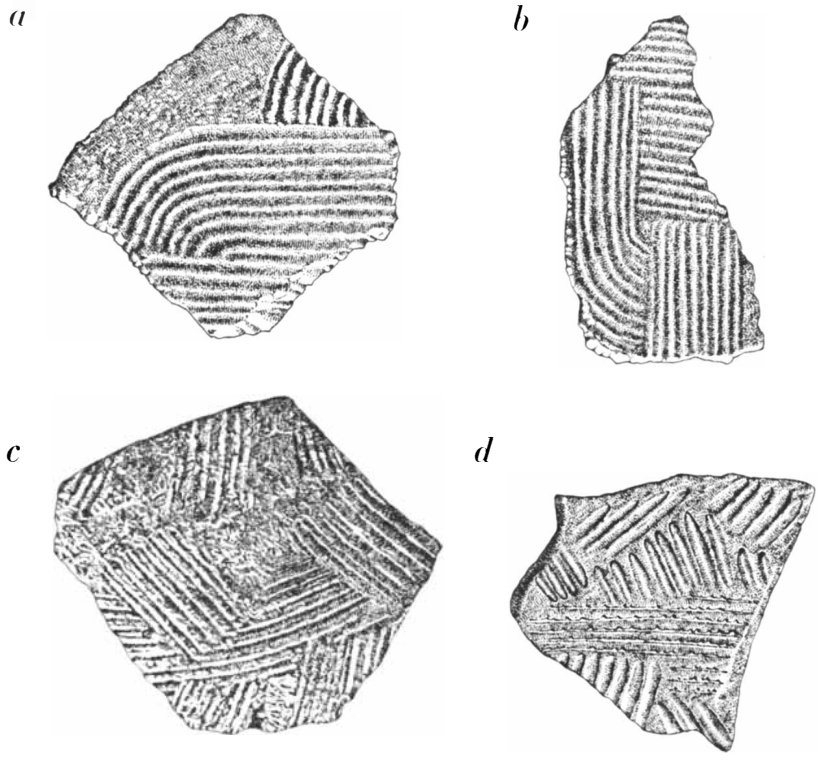
In addition to collecting more plant material last year, Gorman found examples of a new class of artifacts to add to the inventory from Spirit Cave. These are small pellets of baked clay. In excavating Non Nok Tha we had found similar pellets that were more recent. Our workmen had suggested that they were projectiles for use with a "pellet bow," a weapon that children in northern Thailand still use to hunt birds. The device is like a conventional bow except that a slingshot-like pouch for the pellet is attached to the bowstring.

It is a peculiarity of Southeast Asian archaeology that stone projectile points, which are among the most abundant artifacts elsewhere in the world, are rarely found in the region. The reason is presumably that the early hunters in the area made their projectiles of wood, as the Southeast Asians who hunt with the bow and arrow still do; objects made of wood, of course, quickly rot and leave no trace in the archaeological record. If the clay pellets at Spirit Cave were indeed made for pellet bows, then such weapons and possibly conventional bows and arrows could have been in use there 9,000 or more years ago.

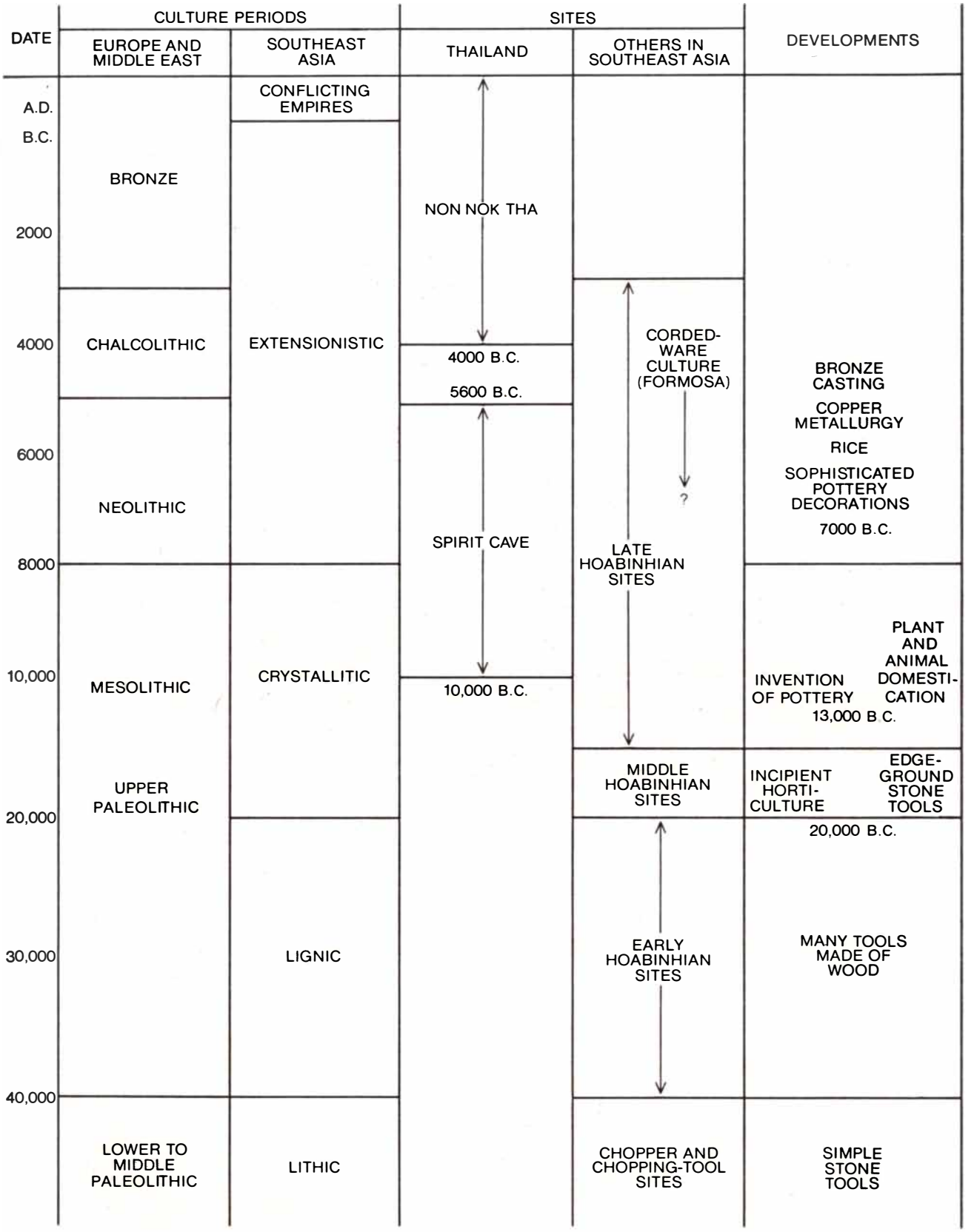
Taking into account both the range



SPIRIT CAVE, located near the Burma border in northwestern Thailand, is being excavated by Chester F. Gorman of the University of Hawaii. From the start, during 1965–1966, Gorman has found the remains of many kinds of edible plants, including nuts, water chestnuts, a species of cucumber and two or three legumes. Some of the material is 12,000 years old. If cultivated and not picked wild, the plants represent man's earliest farming endeavor.



POTTERY FRAGMENTS from Spirit Cave, some nearly 9,000 years old, are decorated with sophisticated incised designs (*a*, *b*), suggesting that the potter's art had evolved at a still earlier date elsewhere in Southeast Asia. The designs resemble some on pottery more than 4,000 years old (*c*, *d*), found on Formosa in 1964 by Kwang-chih Chang of Yale University.



NEW FRAMEWORK for the prehistory of Southeast Asia has been developed on the basis of recent excavations. The traditional subdivisions of the stone and metal ages of Europe, which cannot be applied in Southeast Asia, are shown at left. Beside them are the five new divisions of the past in Southeast Asia, ranging from prehistoric up to historic times. The names of the earliest periods indicate the kind of material most used for tools: Lithic suggests stone tools; Lignic, wooden ones. The later period names indicate the principal trends during each period. The local cultures of the region took shape during the Crystallitic period. Major population shifts away from mountain habitats occurred in the Extensionistic period. The Period of Conflicting Empires witnessed the rise of states after the start of the Christian Era. Elsewhere the chart shows the relation of five periods to specific sites and basic developments.

of the carbon-14 dates and the rich inventory of remains unearthed at Spirit Cave and Non Nok Tha, it is apparent that a radical revision is needed in our concepts of the prehistory of Southeast Asia. The theoretical outline of a new regional prehistory that follows is based on such a revision. My reconstruction is largely hypothetical; it will need to be tested by further excavation both in mainland and island areas. Much of it is based on the findings at the two Thailand sites, but I have included additional data from my own work and the work of others in the region. Because the outline uses some unfamiliar terms, I shall first briefly review the conventional terminology of prehistory in order to place these new terms in perspective.

For more than a century archaeologists have referred to three stages of the Stone Age: the Paleolithic ("old stone"), the Mesolithic ("middle stone") and the Neolithic ("new stone"). The terms were first defined to fit the prehistory of Europe, but they have since been applied so widely that it is often unthinkingly assumed that cultures all over the world have passed through these three stages. It is quite clear, however, that in many regions the European terms do not correctly describe the sequence of events; this is particularly true of Southeast Asia. The evolution of cultures there has been distinctly different from that in Europe and the Middle East. I am therefore suggesting the substitution of new terms for Southeast Asia.

For the earliest period in Southeast Asia I propose the term Lithic. The term refers to the early human use of chipped and flaked stone for tools. This period, roughly equivalent to the early and middle Paleolithic of Europe, is, I suggest, the only one of the three Stone Age stages that all mankind has shared. In Southeast Asia the technology of chipping and flaking stone developed slowly and never did reach the level of extremely fine workmanship found in many other parts of the world. A possible reason for this seeming backwardness is proposed below. In any event I have arbitrarily set the end of the Lithic period around 40,000 B.C.

For the next period I propose the name Lignic, a term derived from the Latin word for wood. It is part of my hypothesis that during this period tools made of wood—particularly those made of bamboo—became more important to the peoples of Southeast Asia than tools made of stone. Such a change in the preferred material for tools would help to account for the slow progress in stone-working technology during much of the region's prehistory.

Since wood rarely lasts in the ground, we can hardly expect to excavate the ancient wood tools that would prove my hypothesis. It is possible, however, to examine the stone tools for signs of the kind of wear suggestive of woodworking. Gorman has done this with the Hoabinhian tools from Spirit Cave, and he has found several that show edge wear indicating their use to shape both large pieces of wood and wood shafts of small diameter. It may be significant in this connection that a high percentage of the charcoal in the hearths at Spirit Cave consists of charred bamboo.

The period when wood tools proliferated I have arbitrarily defined as beginning about 40,000 B.C. and ending about 20,000 B.C. In terms of the currently known prehistoric cultures of Southeast Asia this dating would equate the Lignic period with early Hoabinhian.

For the third period in the region I propose the name Crystallitic. The term has nothing to do with raw materials but is intended to suggest that during this interval distinct local cultures began to take shape, or crystallize, in Southeast Asia. Before the period ended, sometime around 8000 B.C., I believe there had developed many elements of culture that are still found in the region today. In terms of known prehistoric sites the Crystallitic period equates with middle and late Hoabinhian.

I suggest that it was during Crystallitic times that the technique of shaping stone tools by grinding and polishing was first developed. At the start this method was applied only to a tool's cutting edge; later it was gradually extended over the entire surface. Stone tools that are ground and polished rather than chipped and flaked are typical of the Neolithic stage of the Stone Age in Europe and the Middle East; they make their first appearance about 8000 B.C. in the Middle East. I believe the same technology was being pioneered in Southeast Asia much earlier.

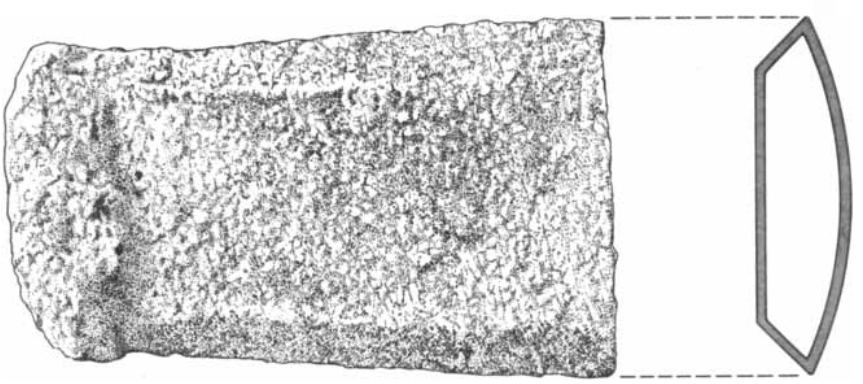
The Crystallitic period is also the interval when plants were domesticated. I suggest that what is called middle Hoabinhian was a culture or cultures whose adherents were experimenting with many different kinds of wild plants for many different reasons. At some point, probably about 13,000 B.C. somewhere in the northern reaches of Southeast Asia, such experiments culminated in the domestication of some of these plants and the consequent appearance of horticulture as a new means of food procurement.

Having advanced this far in the Crystallitic period, we make contact at last with some of our newfound evidence: the contents of the lowest levels at Spirit Cave. I suggest that it was the late Hoabinhian culture, as it is represented in these levels, that achieved the transformation from horticulture to generalized plant and animal domestication and that also achieved the invention of pottery. In different parts of the region different plants would have been selected for cultivation. The same was probably true of the animals involved: the pig, the chicken and possibly even the dog. All these different elements of late Hoabinhian culture spread and accumulated as time passed. When the lowest levels at Spirit Cave were being formed, about 10,000 B.C., the people there appear to have possessed such elements of late Hoabinhian culture as an advanced knowledge of horticulture and perhaps domesticated pigs. They had no pottery, however, nor would they acquire any until an entire spectrum of sophisticated wares arrived at Spirit Cave from elsewhere in Southeast Asia some three millenniums later.

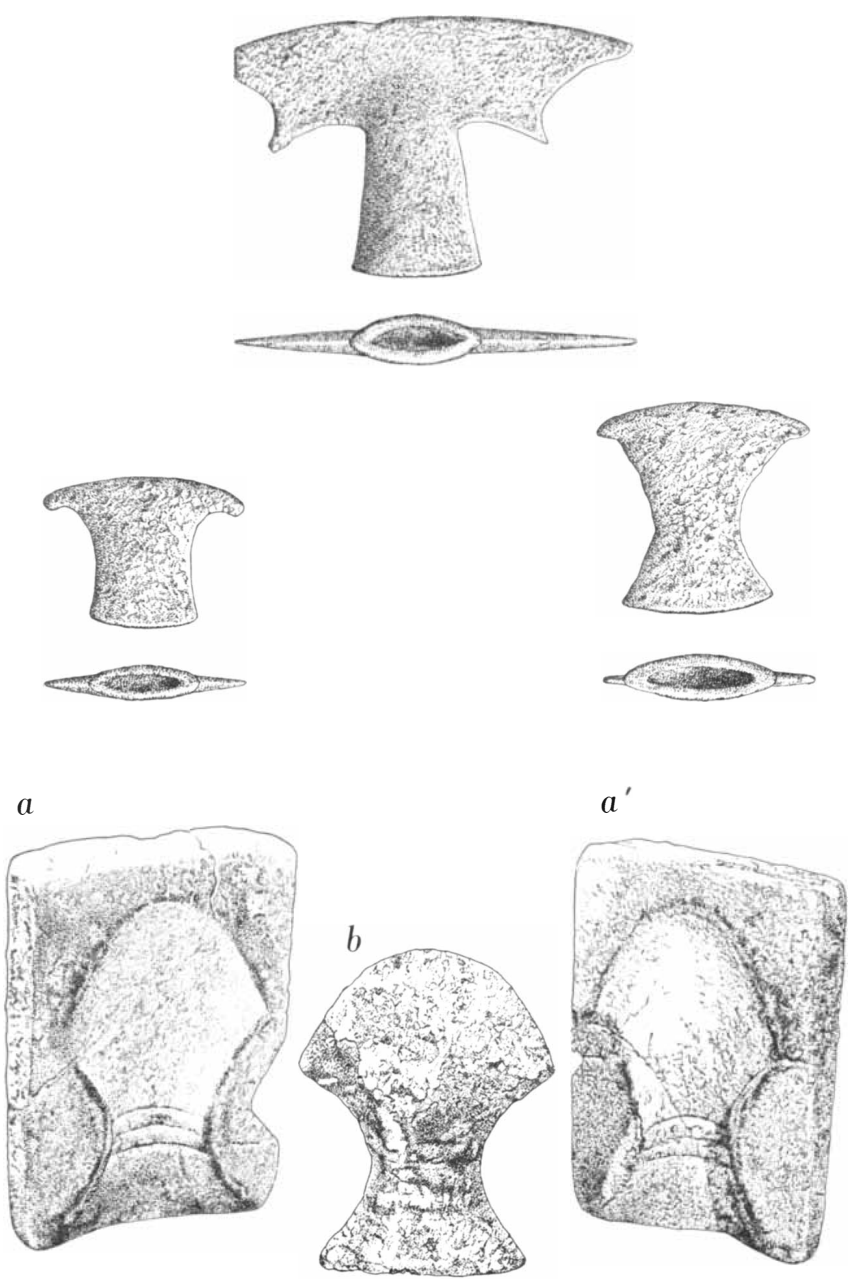
For the fourth period in Southeast Asia I propose the name Extensionistic. The term refers to the major trend during the interval: the movement of peoples out of the mountains where they had previously lived. The Extensionistic trend, which began around 8000 B.C. and ended at about the beginning of the Christian Era, led the mountain peoples not only into the many other hospitable habitats of the mainland but also beyond them; the mountain peoples traveled by overland routes or by water in virtually every direction. I believe future investigation will show that local cultures that were distinctly different from the late Hoabinhian evolved at the start of this period.

At present the only known candidate for one of these newly evolved cultures is the complex of traits that appeared in the upper levels at Spirit Cave, with its rectangular adzes, its slate knives and its sophisticated cord-marked pottery. That complex, however, is not eligible for status as an independent culture; one cannot separate its elements from their late Hoabinhian associations. Is there a similar complex elsewhere, free of a Hoabinhian matrix?

There is on the island of Formosa (Taiwan). Its artifacts include rectangular adzes, polished slate points and cord-marked pottery, some of it decorated with comb incisions like the ware at Spirit Cave. This Formosan complex



SIMPLE COPPER TOOL, perhaps used as an adze or an axe, is the one found in 1968 in a deep burial at Non Nok Tha. The socket at one end (*see section*) makes the tool unique.

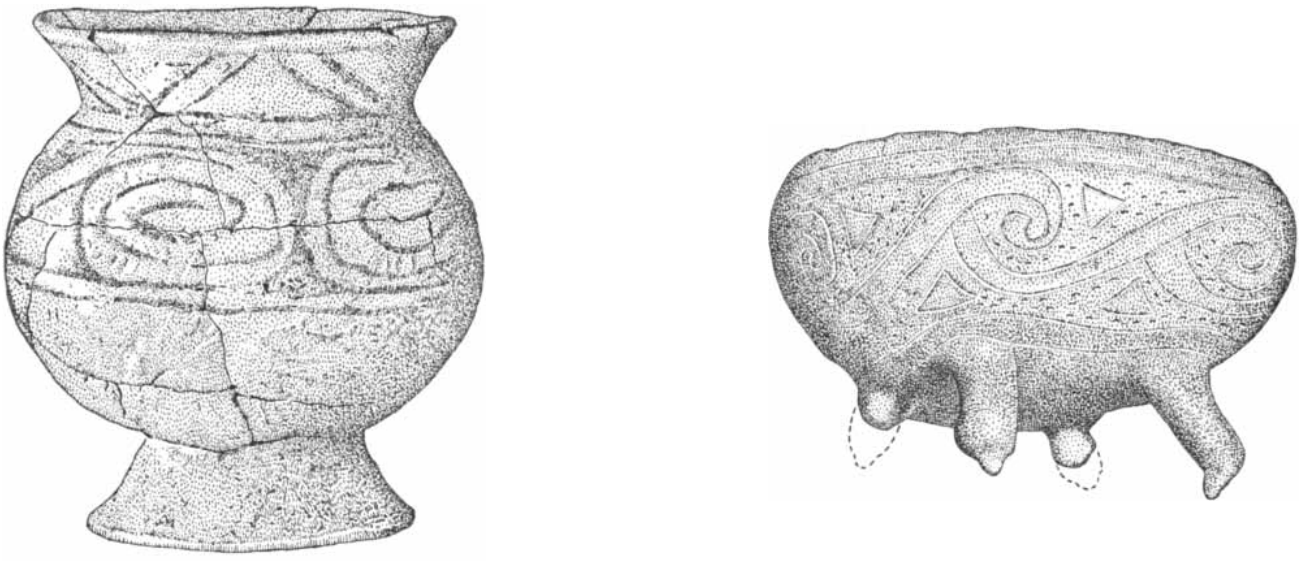


WORK IN BRONZE at Non Nok Tha included casting axe heads in stone molds. Illustrated are two halves of a mold and a matching axe head (*a*, *a′*, *b*), and profile and top views of three typical axe heads, showing the sockets for inserting the wooden axe handles.

has been named the Corded-Ware Culture by its discoverer, Kwang-chih Chang of Yale University. Little is known about its age except that, at the two open-air sites where it has been unearthed, the culture ceased to exist sometime before 2500 B.C. Considering the winds of change that marked the Extensionistic period, it is worth emphasizing two points here. First, Formosa is a long way from Spirit Cave. Second, the remains of the Corded-Ware Culture on Formosa are found not in a conventional Hoabinhian cave setting but out in the open.

What were the consequences of the population movements during this period? I suggest that the earliest movements brought the people no farther from the mountains than into the adjacent piedmont. Even this movement led to enough of an environmental change, however, to make the hunting of wild animals and the gathering of wild plants activities of diminishing significance and to make farming much more important. The transition from dependence on wild food resources to dependence on domesticated ones would of course have been gradual. Indeed, the transition is not complete in Southeast Asia even today. People in towns and even in cities still collect wild produce, and many such items are found in the markets. Wild animals are hunted and trapped wherever possible and make an important addition to the limited human intake of protein.

The beginning of the Extensionistic period coincided with the end of the Pleistocene epoch. The lower level of the sea during the Pleistocene ice ages gave Southeast Asia about twice the land area it has today. The seacoast ran along the edge of what is now the Sunda Shelf, and the islands of western Indonesia and the Philippines were connected to the mainland. Few Hoabinhian sites have been found adjacent to the present coastline; it seems logical to expect that many of the Extensionistic settlements were located in river valleys and on shores that are now submerged. The slowly rising sea that accompanied the last retreat of the continental glaciers must have forced some peoples to retreat to the mainland and have left others marooned on the islands. The inhabitants of the Sunda Shelf who retreated to the coast of South China and North Vietnam were probably the ancestors of the people who took readily to the sea starting about 4000 B.C. and ultimately pushed as far as Madagascar off the coast of Africa and Easter Island off the coast of South America.



DECORATED POTS, uncommon at Non Nok Tha, were confined to the lower levels of the mound, where the burials are about 5,000 years old. One decorative motif, which combines painted triangles with a curvilinear scroll (*left*), is repeated in the incised decorations on pottery found at offshore sites in the region such as the cave site near Kalanay on the Philippine island of Masbate (*right*).

What material changes would have been evident as various local cultures evolved during this period? Again the material from Spirit Cave and Non Nok Tha is suggestive. The slate knives from the upper levels at Spirit Cave closely resemble the kind used today to harvest rice in parts of Indonesia. Is it possible that rice was one of the plants the final residents of Spirit Cave collected and perhaps cultivated? The cereal was certainly known at Non Nok Tha perhaps no more than 1,500 years later, as the pottery impressions of rice grains and rice husks have demonstrated.

Metallurgical analysis of the copper tool that Bayard unearthed at Non Nok Tha shows that it contains traces of phosphorus and arsenic. This fact suggests that the tool was not merely pounded into shape from a large nugget but that the copper was smelted or otherwise heat-treated, and it implies some degree of metallurgical sophistication among the fourth-millennium inhabitants of the site. Not long afterward the people of Non Nok Tha were casting bronze in the form of socketed tools much like the earlier copper one. No trace of copper, tin or lead ores—the three metals used in this region to make bronze—has been found in association with the bronze-casting equipment at Non Nok Tha. Either the ores were smelted at some other part of the site or the refined metals, or the alloyed bronze itself, reached Non Nok Tha through trade channels.

We are led to the conclusion that the relatively advanced metallurgy practiced at Non Nok Tha during the fourth and third millenniums B.C. had its roots in earlier metallurgical developments elsewhere in the region. If one gives due weight to the evidence that socketed metal tools were unknown outside the region until about 2000 B.C., a further conclusion seems inescapable: The development of metallurgy in Southeast Asia was probably independent of and unrelated to the development of metallurgy in the Middle East.

A major puzzle remains. Given the spread of plant and animal domestication, the advances in metallurgy and the development of trade, it is surprising that the entire Extensionistic period was unmarked by the kind of social evolution that accompanied the same events in the Middle East. We find neither the rise of cities nor the growth of centralized political power. Even as late as the second and first millenniums B.C. fortifications are unknown anywhere in Southeast Asia, which strongly suggests that organized warfare was also unknown. With one possible exception the various cultures of the region seem to have shared much the same kind of economic base and to have enjoyed contact with one another but to have remained politically independent. The exception is the culture of the region of what is now North Vietnam and the adjacent parts of China; during the second millennium B.C. a centralized authority that was quite independent of the imperialistic dynasties in northern China may have arisen in this area. More archaeological investigation is needed, however, before the actual extent of the development can be determined.

The first years of the Christian Era coincided with the start of the most recent period in Southeast Asia, which I call the Period of Conflicting Empires. It was then that, as a result of Indian political and religious influences, the first centralized states finally made their appearance in the region, just as they had in India itself. Once established, a number of petty states flourished in succession for some 1,500 years; they were largely parasitic on the population of the region. Beginning in the 16th century and continuing into the 20th, European imperialism gradually supplanted the native variety; the change in rulers made no appreciable difference to the inhabitants. The occupation of Spirit Cave had of course ended long before the Period of Conflicting Empires began; except for the appearance of iron and the water buffalo the archaeological record at Non Nok Tha was scarcely affected by the events of the period.

Following the withdrawal of the European colonialists after World War II, the Period of Conflicting Empires also came to an end. Southeast Asia today is in an interval of readjustment that has been complicated equally by the collapse of empires and of philosophies. The least one may hope for a region where early man took so many remarkable strides along the road to civilization is a return to the live-and-let-live pattern of cultural independence that characterized much of its prehistory.

Tides and the Earth-Moon System

Tidal friction in shallow waters has controlled the evolution of the earth-moon system for aeons. It is causing the length of the day to increase and the moon to recede from the earth

by Peter Goldreich

Anyone who has spent any time at the seashore is aware of the tides, the periodic advance and retreat of the sea from the land. It is well known that the tides are driven by the gravitational pull of the moon and the sun on the earth. Not so well known is the important role the tides have played in the evolution of the earth-moon system. As the tidal currents flow over the bottom, a turbulent boundary layer is created in which the mechanical energy of the water's movement is dissipated as heat of friction. This mechanical energy is derived from the rotation of the earth and the motion of the moon in its orbit. The degradation of mechanical energy into heat results in an inexorable increase in both the length of the earth day and the distance of the moon from the earth. The best modern estimates of the rate at which this evolution is proceeding imply that the earth-moon system has been fundamentally altered by the action of tidal friction. Currently the length of the day is increasing by approximately two milliseconds per century and the moon is spiraling away from the earth at about three centimeters per year.

The moon raises tides both in the oceans and in the solid body of the earth. (It also raises tides in the atmosphere, but the effect of these tides on the earth-moon system is so small that they will not be discussed here.) For the moment consider the idealized case of the tides raised by a satellite traveling in a circular orbit around the equator of a solid, elastic planet (a planet that might be made out of rubber). The gravitational force of the satellite would distort the planet into an ellipsoid with its long axis through the equator pointing toward the satellite. The distortion arises because the satellite's gravitational attraction is greater on the near side of the planet than on the far side. Although the gravitational force diminishes as the inverse square of the distance, the height of the tidal bulges varies as the inverse third power of the separation between the satellite and the planet, reflecting the fact that the tide is due to the *difference* in the satellite's attraction at opposite sides of the planet. An observer standing at one point on the surface of the planet experiences two equal tidal maximums during each revolution that the planet makes under the satellite. The planet and the satellite orbit around their mutual center of mass. The centrifugal acceleration that results from this whirling varies with distance from the center of mass across the planet. This effect, together with the differential force of the satellite's gravity across the planet, is responsible for the raising of two tidal bulges.

FULL MOON SETS OVER OCEAN at low tide just before dawn at the John Fitzgerald Marine Reserve 20 miles south of San Francisco. The tide pool in foreground was left at high tide six hours before the photograph was taken, evidence of the moon's gravitational pull on the ocean causing tides.

The restoring forces that limit the height of the tides are due to the planet's elasticity and self-gravity. In its response to the periodic tidal stretching force (periodic as it is felt at a point fixed to the planet) the planet behaves rather like a weight that is hung from a spring and subjected to a periodic force. If the weight is set in motion and then left alone, it will bob up and down at its natural frequency, which depends only on the mass of the weight and the stiffness of the spring. If the weight is not allowed to oscillate freely but is driven up and down in a periodic fashion, the response of the system depends on whether the frequency of the applied force is higher or lower than the natural frequency of the system. If the forcing frequency is lower than the system's natural frequency, the displacement of the mass is in phase with the force, that is, the maximums and minimums of the displacement occur simultaneously with the maximums and minimums of the force. If the forcing frequency is higher than the natural frequency, the displacement is half a cycle out of phase with the applied force.

The solid body of the earth is more complex than such a simple harmonic oscillator and is capable of simultaneous oscillation in modes of many different frequencies. The longest oscillation period, which is determined by the time it takes seismic waves to travel through the earth, is about 55 minutes. In recent years detailed studies of the earth's free oscillations have been made following their excitation by great earthquakes [see "Resonant Vibrations of the Earth," by Frank Press; SCIENTIFIC AMERICAN, November, 1965]. Because the tidal period of slightly less than 12 hours is longer than the free-oscillation periods, high tides in the solid earth are located along an extension of the earth-moon center line.

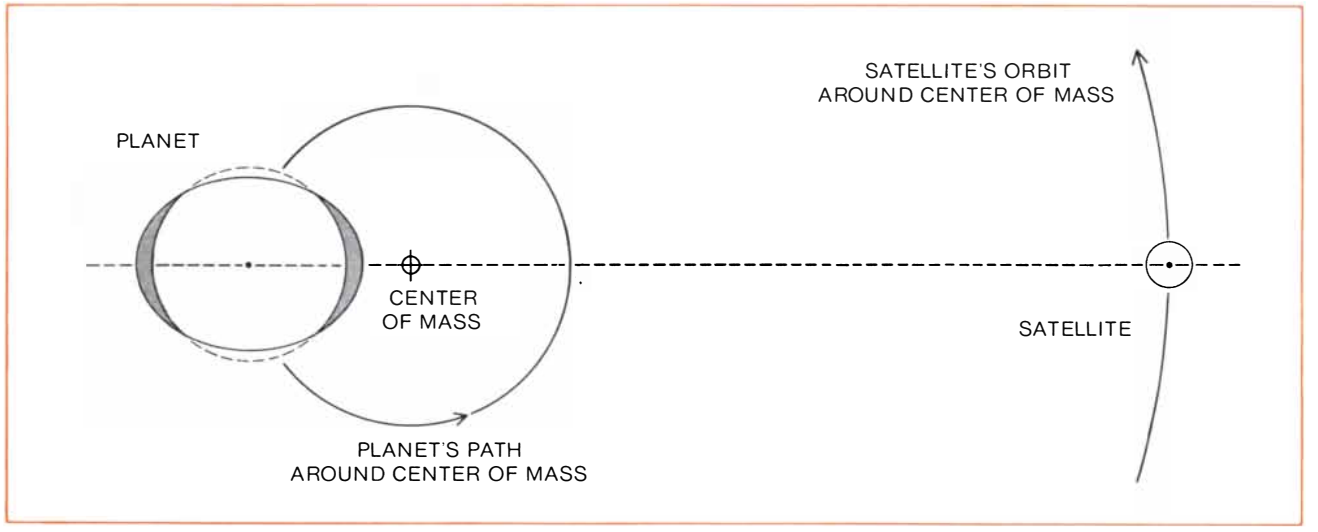
The range of free-oscillation periods of the oceans is set by the time it takes tidal waves to propagate across them. (These are not the sometimes disastrous "tidal waves" generated by earthquakes.) The speed of tidal waves can be computed by multiplying the earth's surface gravitational acceleration by the ocean depth and taking the square root of this quantity. For a typical ocean depth of four kilometers the speed is .2 kilometer per second. Thus a tidal wave can cross 10,000 kilometers of ocean in

about half a day, and the oscillation periods of the oceans are comparable to the tidal forcing period. This coincidence, and the irregular shape of the ocean basins, conspire to make the ocean tides exceedingly complicated. Accordingly we shall continue to restrict our attention to solid-body tides whenever possible.

In a real planet the dissipation of energy by friction modifies the periodic tidal distortion. Once again we shall find it helpful to exploit the analogy between a planet and a weight-spring system. The presence of friction causes the displacement of a spring to lag in phase behind the applied force. By the same token high tides at points on opposite sides of the earth occur some time after these points are aligned along the direction of the center of mass of the planet-satellite system. If the planet's rotation period is shorter than the satellite's orbital period (as is the case with the earth and the moon), the delay will result in the displacement of the tidal maximums in the direction of the planet's rotation. The asymmetrical position of the tidal bulges with respect to the line of centers gives rise to a net torque, or twisting force, on the satellite. This torque increases the satellite's orbital angular momentum. For every action there is a reaction; a torque of equal magnitude but opposite direction acts on the planet and brakes its spin. A net torque results because the satellite is attracted more strongly by the closer tidal bulge that is leading it in longitude. The tidal torque varies as the inverse sixth power of the planet-satellite separation.

The tidal torque transfers angular momentum and energy from the planet's spin into the satellite's orbit. It requires less energy to give the moon a certain amount of angular momentum than it takes to give the earth the same amount. Therefore, although the total angular

TIDAL DISTORTION at the equator of a perfectly elastic planet is caused by a satellite traveling in a circular orbit. The gravitational force of the satellite would deform the planet into an ellipsoid with its long axis *A*–*B* pointing toward the satellite. View is from the top of the system; *P* indicates the pole of the planet. The planet rotates in the same direction as the satellite orbits.



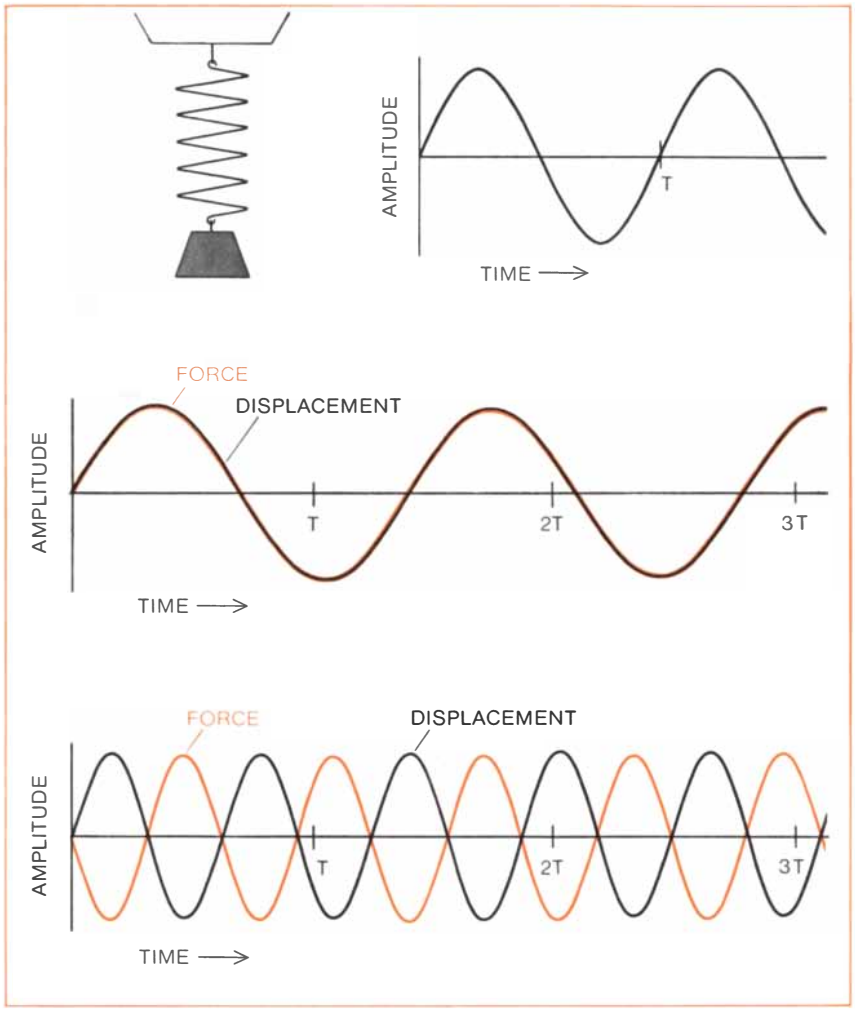
TWO EQUAL TIDAL BULGES are felt by an observer standing at one point on the planet's surface. The bulges are raised by the differential force of the satellite's gravity across the planet coupled with the centrifugal acceleration that results from the system's whirling around its center of mass. The centrifugal acceleration increases with distance from the center of mass across the planet. The center of mass of the earth-moon system is actually about 1,000 miles below the surface of the earth, not outside as shown.

momentum of the planet-satellite system remains invariant, its total mechanical energy decreases as the result of frictional losses in the planet. This loss of energy through friction plus the transfer of angular momentum means that the moon is gaining angular momentum at the expense of the earth's spin. The direction of flow of angular momentum is entirely determined by the relative lengths of the month and the day.

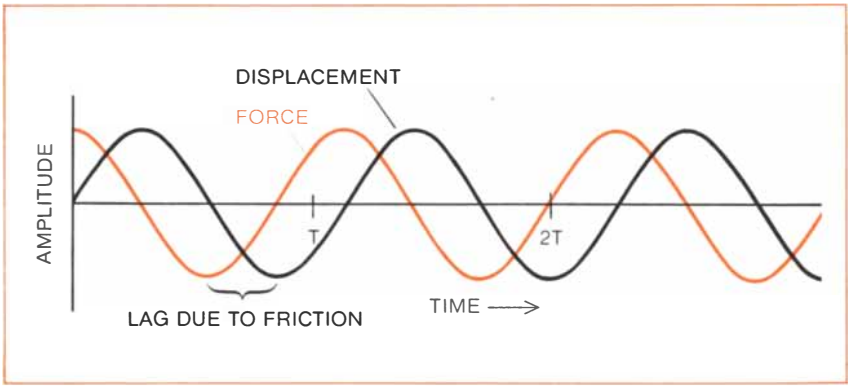
This picture of tidal evolution is not significantly changed by the inclusion of the ocean tides. It is true that simple schematic drawings do not accurately depict the shape of tidally distorted oceans. Nevertheless, the qualitative evolution I have outlined of the transfer of angular momentum from the planet's spin into the satellite's orbit still applies.

The eccentricity, or oval shape, of a satellite's orbit is affected by tidal friction. Imagine a satellite moving on an elliptical orbit around a planet and that the transfer of angular momentum from the planet's rotation into the satellite's orbit takes place in just two impulses, a large one at pericenter (nearest approach) and a smaller one at apocenter (farthest separation). The impulsive addition of angular momentum at pericenter does not alter the pericenter position but does increase the distance to apocenter. Similarly, the impulsive addition of angular momentum at apocenter does not affect the position of apocenter but does act to increase the pericenter distance. Because the tidal torque varies as the inverse sixth power of the planet-satellite separation, more angular momentum is added near pericenter than near apocenter, and the orbital eccentricity increases.

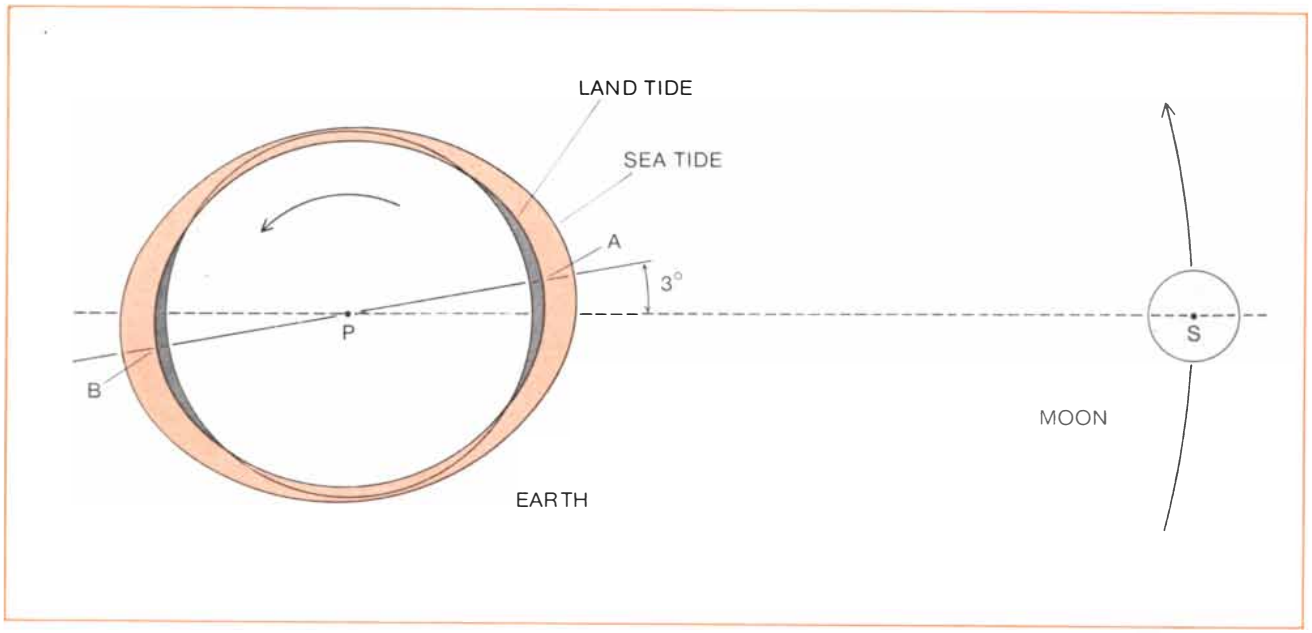
The tides raised by the planet in the satellite tend to decrease the orbital eccentricity. In all cases where this effect is significant these same tides will have "despun" the satellite to synchronous rotation, that is, the satellite's period of rotation around its own axis equals its period of revolution around the planet and it always presents one face to the planet. In this case the tidal bulges on the satellite are purely radial, along the planet-satellite line of centers, and do not produce a tidal torque. Tidal energy is still dissipated by friction as the height of the tides in the body of the satellite oscillates in response to the variation of the planet-satellite separation. The energy dissipated as heat in the satellite is derived from the satellite's orbital energy, and since for a fixed angular momentum the lowest-energy orbit is a circle, these tides will tend to make the satellite's orbit rounder.



WEIGHT-SPRING OSCILLATOR is a model of the periodic tidal stretching force the planet feels as it is pulled by the satellite. If the weight-spring system is set in motion and then left alone, it will bob up and down at its natural frequency (*top*). If the weight is periodically driven up and down by a forcing frequency (*color*) lower than the system's natural frequency, the displacement (*black*) of the weight is in phase with the force (*middle*). If the forcing frequency is higher than the oscillator's natural frequency, then the displacement of the weight is a half-cycle out of phase with the applied force (*bottom*).

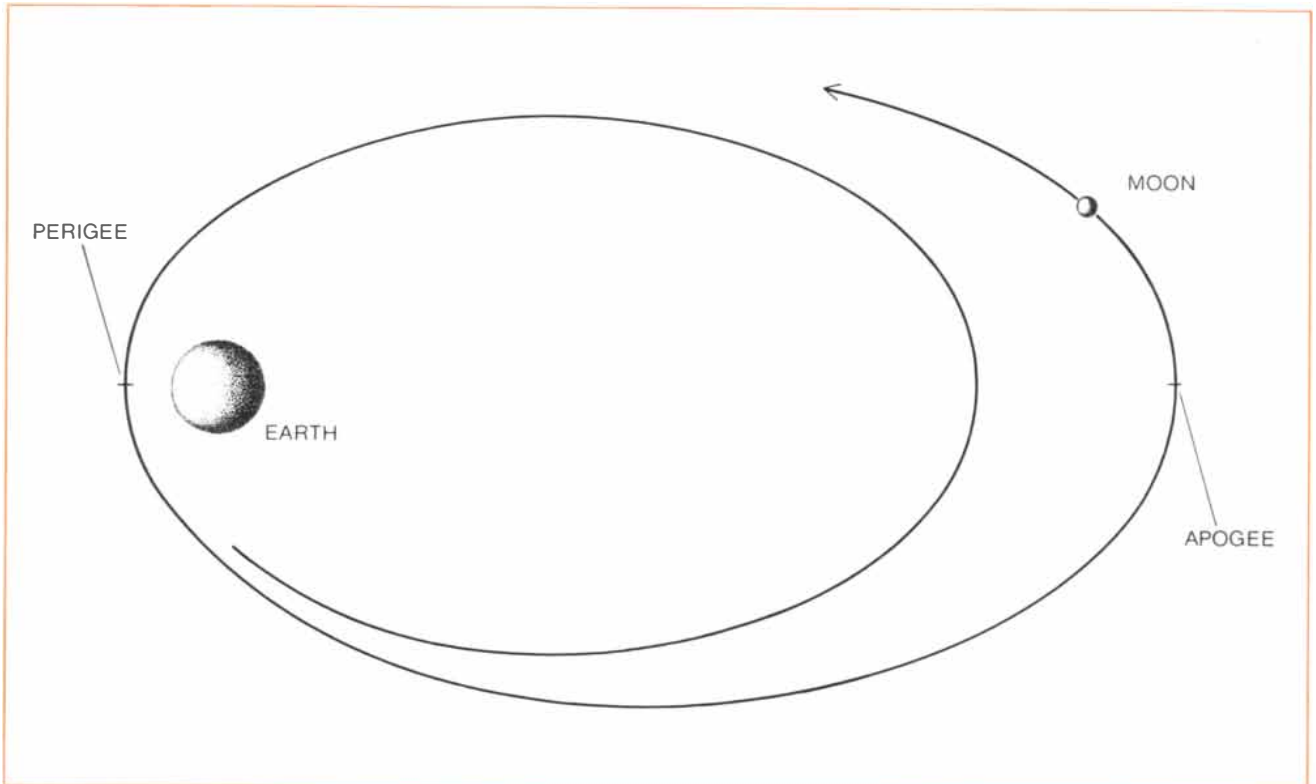


FRICTION CAUSES A DELAY in the response of the weight-spring system forced below its natural frequency. This delay means that the displacement of the weight (*black*) lags in phase behind the forcing frequency (*color*), in analogy to the way that the tidal bulges on opposite sides of the earth occur sometimes after these points align with the moon.



TIDAL BULGES LAG three degrees in phase behind the line of centers *P–S* between the earth and the moon because of friction. In the case of the earth, whose rotation period is shorter than the moon's orbital period, the lagging tides are carried ahead of the moon in longitude by the rotation of the earth. Land tides on the earth are black and sea tides color. View is from the pole.

Because tidal torque varies as the inverse sixth power of distance, its effect can be felt only over rather small separations, thus limiting its importance in the dynamics of the solar system. No planetary orbit has been appreciably altered by the tides the planet raises on the sun. The tides the sun raises on the planets have significantly despun the two innermost planets, Mercury and Venus. They have had a more modest effect on the spin of the earth and no appreciable effect on the rest of the planets. Tidal friction is more important in satellite systems than in the planetary system because the distances involved are smaller. Thus the tides raised by the earth on the moon, and by Jupiter and Saturn on their large inner satellites, have brought these satellites into synchronous rotation. The evolution by tidal friction of



MOON SPIRALS AWAY FROM EARTH as the earth transfers some of its angular momentum to the moon. The moon's orbit is an ellipse. The impulsive addition of angular momentum at perigee (moon's closest approach to earth) changes the orbit to a larger, more eccentric ellipse. The moon will return to the original perigee if no further changes of angular momentum take place.

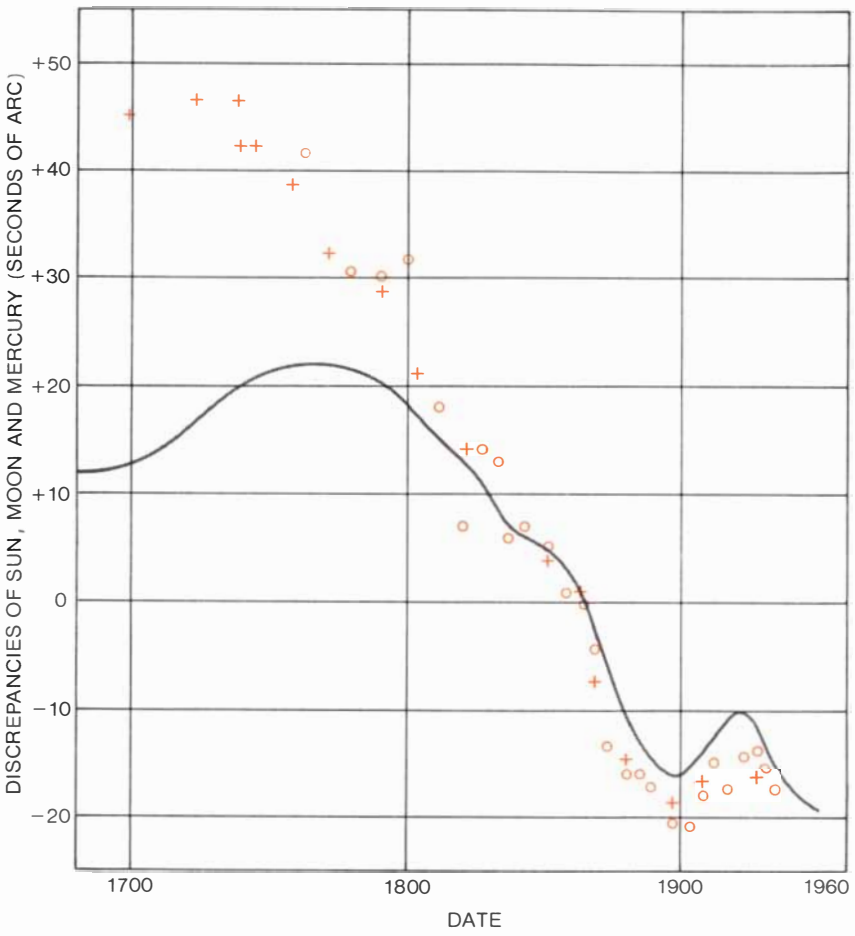
satellite orbits, other than the moon's, has not been established observationally, although there are several theoretical arguments that suggest the possibility of some tidally induced changes.

The synchronous rotation of the moon is the most easily observable consequence of tidal friction in the solar system. In 1754 Immanuel Kant first proposed tidal friction as the explanation of the moon's rotation. At its present distance from the earth the despinning of the moon from a much faster initial rotation rate would have taken no longer than 10 million years, a mere instant of geologic time. The stability of the moon's synchronous rotation is provided by its irregular figure, which is slightly elongated in the direction of the earth. This stability prevents the synchronous lock from being broken by the collision of comets and asteroids with the moon.

Direct measurement of the rate at which the lunar orbit is expanding is not yet possible, although lunar laser ranging provides some hope for the future [see "The Lunar Laser Reflector," by James E. Faller and E. Joseph Wampler; SCIENTIFIC AMERICAN, March, 1970]. Instead two less direct methods based on the relation between the radius and the period of the lunar orbit have been used.

The first of these methods relies on "modern" (over the past 250 years) observations of the longitudes of the sun, Mercury and the moon as seen against the background of the distant stars. The observed longitudes of the sun and Mercury are compared with the longitudes that are predicted based on a constant rotation rate for the earth. Discrepancies between the two sets of longitudes allow a determination of the irregularities in the earth's rotation rate, some of which are the result of nontidal causes such as changes in the earth's moment of inertia and the transfer of angular momentum between the core and the mantle by the redistribution of mass inside the earth. When the effects of variations in the earth's rotation rate are subtracted from the discrepancy in the longitude of the moon, a unique determination of the rate of increase of the lunar orbital period is obtained.

The second method of measuring the rate at which the lunar orbit is expanding involves studying the records of ancient solar eclipses. Such records were exploited for the first time by Edmund Halley in 1693. The method has the advantage of employing a longer time base than the one founded solely on modern observations, but it suffers from ambiguities and inaccuracies in the ancient records. In a recent study Robert R. Newton of Johns Hopkins University concluded that some records of ancient eclipses were biased by their authors' desire to enhance their historical or literary impact. We shall rely on the results obtained from the "modern" observations, although they are currently subject to some dispute.



DISCREPANCY IN THE MOON'S LONGITUDE is shown by the observations of occultations of stars by the moon (*gray*). The moon has a tendency to be behind its position as predicted assuming a constant rotation rate for the earth, that is, the moon's orbital period is lengthening. Discrepancies in the positions of the sun (○) and Mercury (+) occur between their predicted longitudes and their observed longitudes. The measurements allow a determination of the irregularities in the earth's rotation rate, some of which are due to nontidal causes. When these irregularities are subtracted from the discrepancy in the longitude of the moon, the rate of increase of the lunar orbital period is determined.

As I have mentioned, these observations imply an expansion of the moon's orbit by about three centimeters per year and an increase of approximately 20 microseconds per year in the length of the earth day. The associated value of the angle of the tidal lag with respect to the earth-moon line of centers is roughly three degrees. The rate of tidal-energy dissipation is about 2.6×10^{12} watts, which is close to the world's power consumption and about 15 percent of the total heat flow from the earth's interior.

If we assume that the present value of the tidal-lag angle is characteristic of its value in the past, we can extrapolate the lunar orbit back in time and deduce when, if ever, the moon was close to the earth. This procedure yields the startling result that the moon was near the earth less than two billion years ago. It is generally believed that the planets, including the earth and the moon, are 4.6 billion years old. One would think that, if the moon was ever near the earth, it would have been so early in the history of the earth-moon system. Is there any way out of the discrepancy in the two dates? The most plausible way would be to discard the assumption that the tidal phase lag has remained constant over geologic time. Certainly it would seem foolhardy to extrapolate back two billion years relying on a parameter that has been determined from data extending back no more than 4,000 years, the time scale covered by the ancient solar-eclipse

records. Our confidence in the applicability of the present tidal phase lag has been bolstered by a remarkable discovery in paleontology. In 1963 John W. Wells of Cornell University suggested that the yearly growth bands on fossil corals are made up of daily growth ridges. From counts of the number of daily increments per yearly band on fossil corals the number of days per year has been determined all the way back to the mid-Devonian period of some 380 million years ago [see "Corals as Paleontological Clocks," by S. K. Runcorn; SCIENTIFIC AMERICAN, October, 1966]. The data that now exist are consistent with a constant phase lag of three degrees over the past 380 million years. As remarkable as the coral data are, they only take us back to a time when the length of the day was 22 hours and the moon's distance from the earth was 58 earth radii instead of the present 60.25. To assess what is involved in a further backward extrapolation it is necessary to locate the region where the tidal energy is being dissipated.

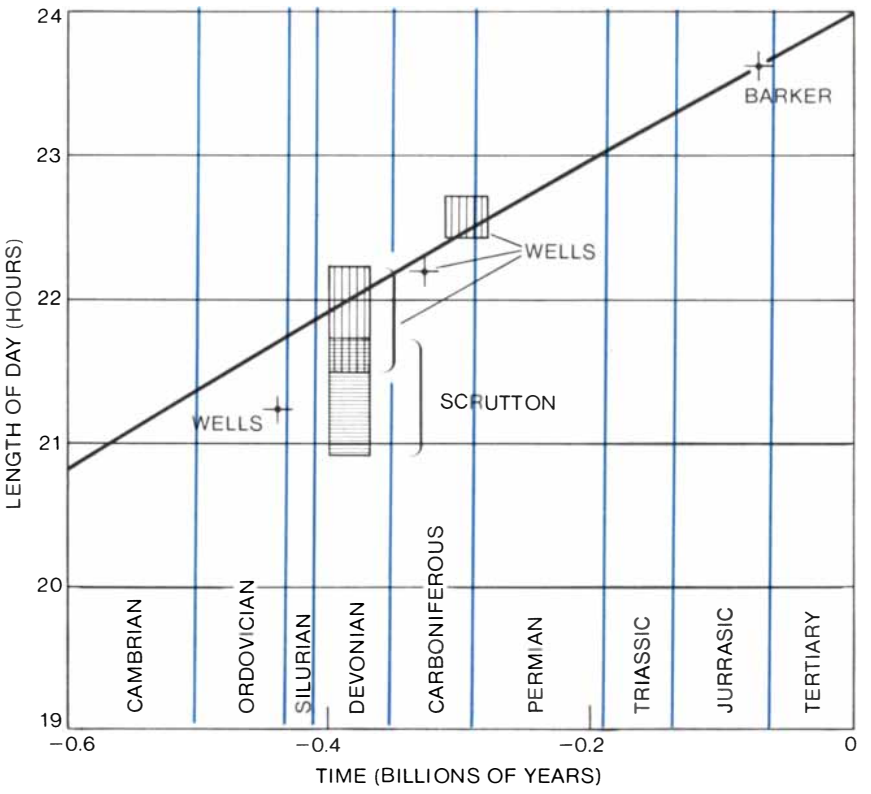
Although the astronomical data can reveal the total rate at which tidal energy is being lost in the earth, they provide no information on the location of the dissipation. The first question is: Is the tidal energy being dissipated mainly in the oceans or in the solid body of the earth? Three methods have been used to estimate the total rate of tidal-energy loss in the oceans and they all suggest that most, if not all, of the tidal dissipation in the earth is taking place in the shallow seas. The crudest method consists in consulting tidal charts that show the tidal distortion of the oceans. From the shape of the tidal distortions the torque on the moon can be directly calculated.

Two more accurate methods were suggested by Sir Geoffrey Taylor in 1919. Both methods rely on the fact that the tidal currents in the deep oceans are so slow that hardly any energy should be dissipated there. Thus the dissipation of tidal energy should be confined to the shallow seas where the tidal currents are amplified to much greater velocities. Taylor's first method is to estimate the tidal currents in the shallow seas and then directly calculate the frictional dissipation of energy in the turbulent boundary layers along the ocean bottom. This method is very sensitive to errors in the tidal velocities because the rate of energy dissipation is proportional to the cube of the velocity. At one time this method was used to predict that more than half of the total dissipation took place in the Bering Sea. Later estimates revised the adopted tidal currents downward. It is now maintained that the Bering Sea is responsible for only 10 percent of the total tidal-energy loss. Taylor's second method is currently the most favored one. It requires knowledge of both the tidal height and the flux of water at the entrance to a shallow sea. From the tidal height the pressure at any depth can be calculated, and the net work done on the water entering and leaving the shallow sea can be estimated. A compilation of calculations of this type by Gaylord R. Miller, then at the University of California at San Diego, shows that the shallow seas account for roughly two-thirds of the astronomically determined rate of energy dissipation.

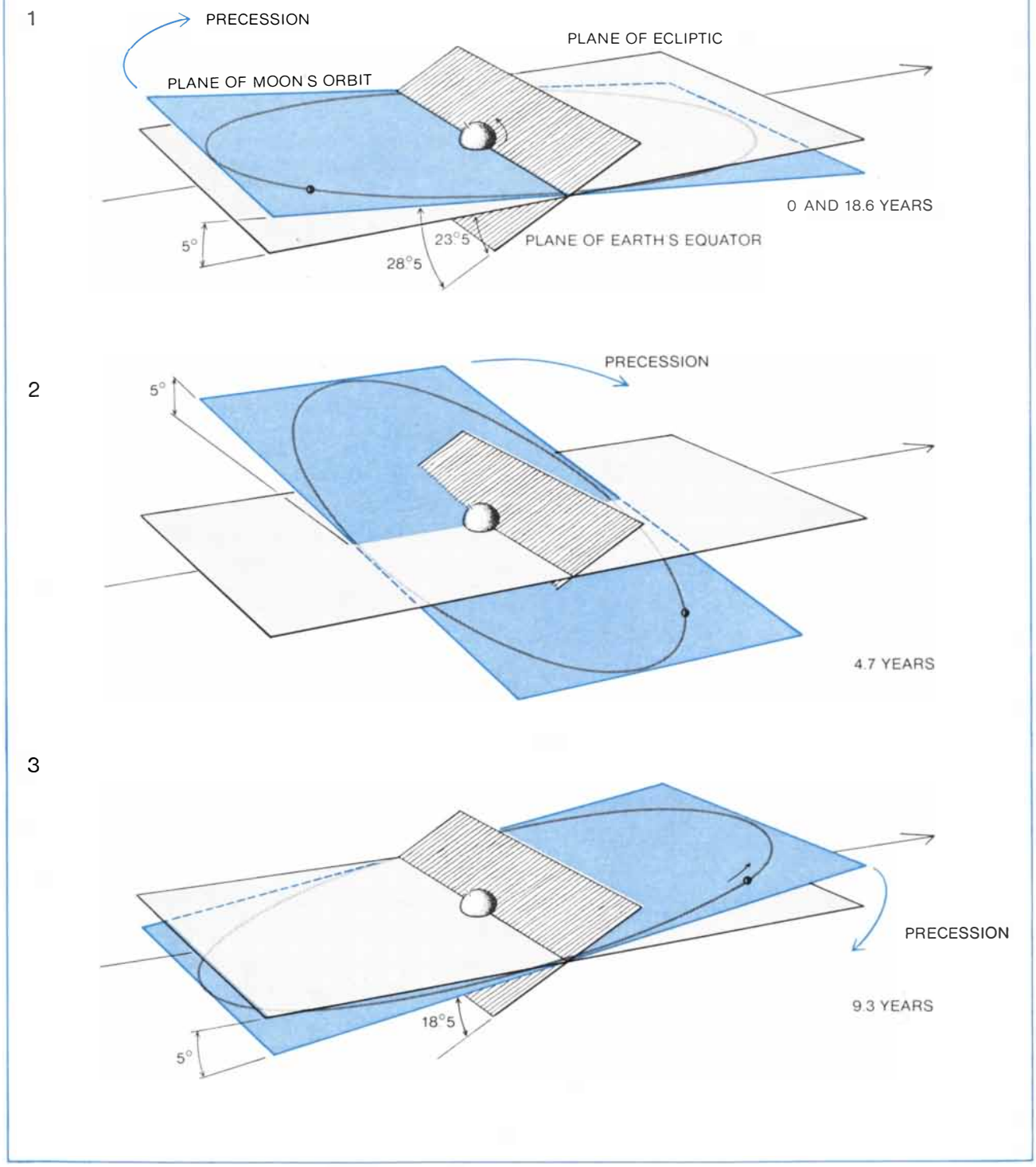
If we accept that tidal energy is dissipated mainly in the shallow seas, then there is little basis for assuming that the tidal phase lag was constant during the past two billion years. Evidence of continental drift and sea-floor spreading points to a very different configuration of oceans and continents as recently as 500 million years ago. This might well account for the factor of two or three by which the present phase lag must exceed its past average in order to reconcile the rate of tidal evolution with the age of the earth-moon system.

The dynamics of the earth-moon system is complicated by the influence of the sun. At present the plane of the moon's orbit is inclined to the ecliptic (the plane of the earth's orbit) by five degrees; the earth's Equator is inclined to the ecliptic by 23.5 degrees. The sun forces the moon's orbital plane to rotate in space while maintaining a constant inclination to the ecliptic; this precession of the lunar orbit has a period of 18.6 years. The combined attraction of the moon and the sun for the earth's rotational bulge forces the earth's axis to precess with the much longer period of 27,000 years. The attraction of the moon by the earth's rotational oblateness is insignificant, and the lunar orbit is controlled by the sun. Thus the inclination of the moon's orbit to the earth's Equator varies between 18.5 and 28.5 degrees with a period of 18.6 years.

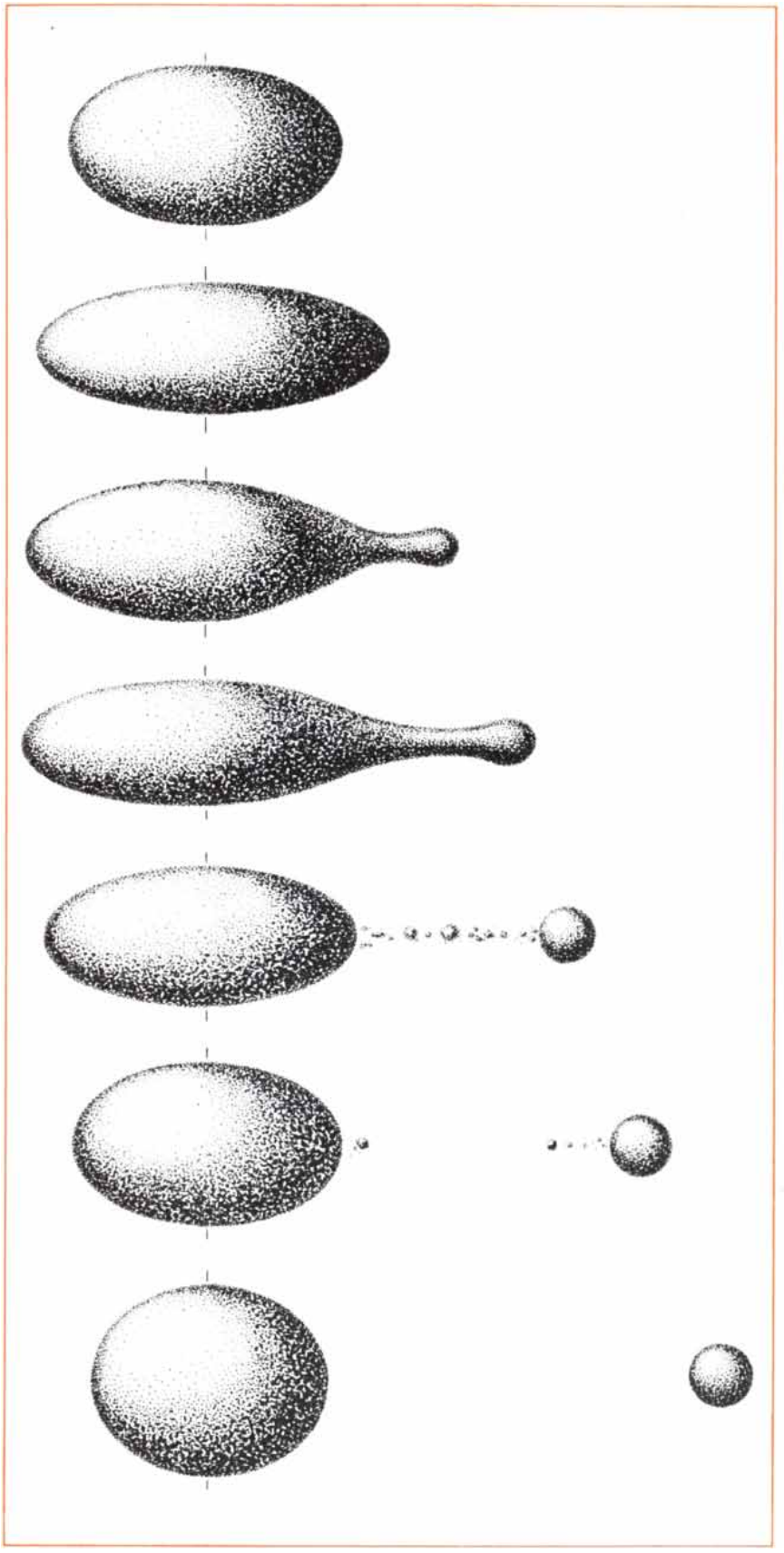
In the past the moon was closer to the earth, and the earth's equatorial bulge was greater. Both factors enhanced the strength of the interaction of the earth's oblateness and the lunar orbit. Simple calculations show that if the moon were ever closer than 10 earth radii, this interaction would have dominated the interactions of the lunar orbit and the



DAY WAS LESS THAN 21 HOURS LONG 600 million years ago according to the counts of daily growth bands on corals. Names of investigators appear next to their results. Length of the day is extrapolated back through the ages on the assumption that the present tidal phase lag of three degrees has remained constant throughout the past half-billion years.



PLANE OF MOON'S ORBIT PRECESSES with a period of 18.6 years because of the gravitational force of the sun. The moon's orbital plane (*dark color*) is inclined to the ecliptic (plane of the earth's orbit) by five degrees; the plane of the earth's Equator (*hatched area*) is inclined to the ecliptic by 23.5 degrees. The sun forces the moon's orbital plane to rotate in space while maintaining a constant inclination (five degrees) to the ecliptic. Once every 18.6 years the plane of the earth's Equator will tilt above the ecliptic in the same direction (*long arrow*) that the plane of the moon's orbit will tilt below (*top*). The angle between them is 28.5 degrees, the most it can ever be. In 4.7 years the plane of the moon's orbit will have precessed a quarter of the way around and will be at right angles to its former position (*middle*). In 9.3 years the plane of the moon's orbit will have gone halfway around the precessional cycle (*bottom*). The plane of the earth's Equator and the plane of the moon's orbit will both be tipped above the ecliptic in the same direction, and the angle between the two planes will be at its minimum amount of 18.5 degrees. In 18.6 years the moon's orbital plane will have continued the rest of the way around and returned to its initial position, where it will repeat the precessional cycle.



FISSION HYPOTHESIS supposes that the earth was initially rotating with a period close to 2.6 hours. As iron was drawn from the mantle into the core the earth's moment of inertia decreased, increasing its rotation rate. The body became unstable, changing from an oblate spheroid (*top*) to a pear-shaped figure with its long axis in the plane of its rotation (*middle*). Eventually the neck of the pear split off and formed the moon (*bottom*).

sun. At less than 10 earth radii the inclination of the moon's orbit to the equatorial plane of the earth would have remained constant. As the lunar orbit and the earth's Equator precessed, the angle the Equator makes with the ecliptic and the inclination of the lunar orbit to the plane of the ecliptic would have varied periodically.

Armed with an understanding of tidal friction and earth-moon-sun rigid-body dynamics, it is possible to appreciate the calculated history of the earth-moon system. Calculations similar to the ones I carried out several years ago have been performed by many others, starting with the pioneering work of George Darwin (the son of Charles Darwin) in the late 19th century. More recent investigators include Horst Gerstenkorn, William M. Kaula, Gordon J. F. MacDonald, S. Fred Singer and Louis B. Slichter. Such calculations show that the moon never moved in an equatorial orbit. Its orbital inclination to the earth's Equator would have exceeded 10 degrees if the moon had ever been closer than 10 earth radii in the past. At the closest approach of the moon the length of the day would have been about five hours.

We are on less certain ground when we consider the variation of the eccentricity of the lunar orbit in the past. The value of the tidal phase lag in the moon is not yet known. (Recall that the tides raised by the earth in the moon act to decrease the eccentricity.) Even if tidal dissipation in the moon is about the same as the dissipation in the earth, the dominant effect on the eccentricity of the lunar orbit would be the one due to the tides raised in the earth. In fact, the absence of oceans on the moon suggests that its tidal phase lag is smaller than the earth's. Thus it seems highly likely that the moon's orbital eccentricity is increasing at the present time, and that it was therefore even smaller in the recent past.

The results of the calculations I have been discussing are not in themselves very controversial. Speculation on the origin of the moon is largely concerned with deciding where along the calculated evolutionary path the moon actually originated. If there were some readily discernible landmark to which the evolutionary path pointed, there would undoubtedly be more agreement on the origin of the moon. In the absence of any landmarks a bewildering variety of hypotheses have been put forward. No one article can do justice to all of them, but we can examine the general classes of proposals to see how well they satisfy the constraints imposed by the evolutionary calculations. Theories of lunar origin

divide naturally into three categories: fission, capture and binary formation.

The idea that the moon was split off from the earth was originally put forward by George Darwin. The modern version of this hypothesis relies on the idea of rotational instability of the primordial earth; this instability caused the earth to split apart into the present earth-moon system. This class of theories is faced with enormous difficulties. Foremost among them is the fact that the earth's day would have been about five hours when the moon had just split off. An initial rotation period close to 2.6 hours is required to induce rotational instability. Proponents of the fission hypothesis must account for the disappearance of about half of the initial angular momentum of the earth-moon system. This has not proved to be an easy task. The fission hypothesis also cannot readily explain the current five-degree inclination of the lunar orbit to the ecliptic plane. The evolutionary calculations imply that the lunar orbit would have been inclined by at least 10 degrees to the earth's Equator when the two bodies were close together. Fission would result in an initially equatorial orbit. Computations show that an initially equatorial orbit would have evolved into an orbit that lay in the ecliptic plane at the present earth-moon separation.

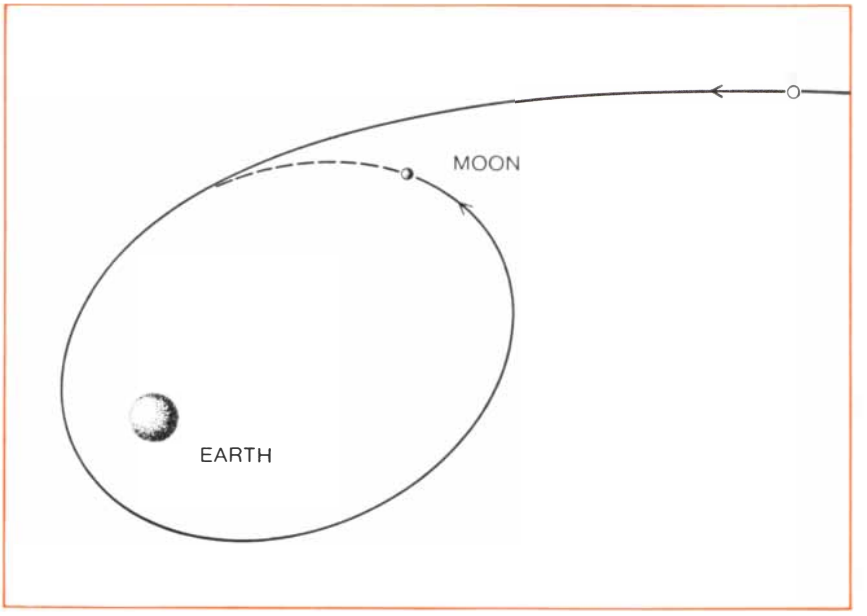
Capture theories of lunar origin also have their problems. These theories propose that the moon was once an independent planet that was caught in the earth's gravitational field and swung into a lunar orbit. Problems with capture theories are in large part due to the implausibility of the capture process, which requires the velocity of an initially unbound moon to be reduced by some kind of dissipation process. Tidal friction and collisions with material already in earth orbit have been proposed as dissipative sources leading to capture. Tidal-friction losses during a single close encounter between the earth and the moon would be very small, and energy losses due to collisions with orbiting material would require a significant fraction of a lunar mass to be in orbit around the earth.

This brings us to the one really new hypothesis of lunar origin that has appeared during the past half-century. It was conceived by Gerstenkorn, who found by tracing the orbit of the moon back past the time of minimum earth-moon separation that still earlier the moon would have moved in a highly inclined eccentric orbit. Gerstenkorn reasoned that the moon had initially been captured in such an orbit, and that tidal friction had subsequently decreased the orbital eccentricity and inclination and then increased the radius to its present value. As ingenious as Gerstenkorn's idea is, it has not met with wide acceptance. One of the obstacles it faces is the improbability of the capture process, which Gerstenkorn attributes to tidal dissipation in the body of the moon. Capture by this process requires that the moon approached the earth with a relative velocity of less than one kilometer per second aimed at a distance of no more than two earth radii from the surface of the earth. The source of concern is that at closest approach the moon would have been within the Roche limit of 2.86 earth radii. Within this distance the earth's tidal gravitational force exceeds the lunar self-gravity and would tear the moon apart. It is not clear whether or not the moon's elastic strength would have been sufficient to prevent its destruction during the short time it would have spent within the Roche distance.

The hypothesis that the moon formed from material orbiting the earth faces no real dynamical problem. There is little doubt, however, that the bulk composition of the moon, which has a density of 3.3 grams per cubic centimeter, is significantly different from the bulk composition of the earth, whose mean density is 5.5 grams per cubic centimeter. Advocates of the fission theory have appealed to the similarity between the density of the earth's mantle and that of the moon as support for their hypothesis. Capture enthusiasts have used the density difference to argue that the moon did not form from the same batch of material as the earth. In spite of the difficulties associated with the chemical differences between the earth and the moon (which may largely involve the much lower abundance of iron in the moon), I favor the hypothesis that the moon formed from material in orbit around the earth. As yet there is no compelling reason to prefer any particular distance from the earth as the site of formation.

After all this discussion of the history of the earth-moon system it is appropriate to describe briefly what the future holds in store. It is worth mentioning that all future predictions are based on the assumption that the sun does not suffer a catastrophe such as a supernova explosion before these changes have taken place.

Calculations show that the moon will continue to recede from the earth until it reaches a distance of about 75 earth radii. At that stage the length of the earth day and the lunar month will be equal. The most likely possibility is that, once the day and the month are equal, the synchronous rotation of the earth with respect to the moon's revolution would be stabilized for all future time by the shape of the earth's figure, which would be slightly elongated in the direction of the moon. The earth would point its axis of least inertia toward the moon, just as the moon now keeps its axis of least inertia toward the earth. This relation would be maintained as the



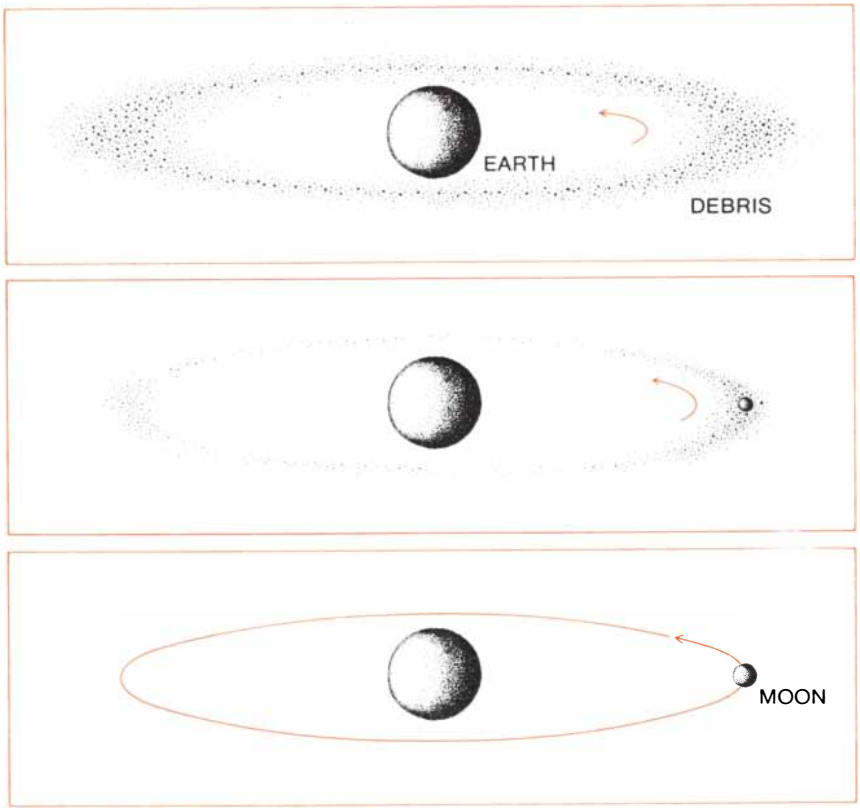
CAPTURE HYPOTHESIS proposes that the moon was once drifting free in space and was caught in the earth's gravitational field and swung into lunar orbit. Its excess energy would have been dissipated by tidal friction and collisions with debris orbiting the earth.

solar tides continued to attempt to brake the earth's rotation. Most of the angular momentum of the earth-moon system would be held in the moon's orbit. The earth feels a torque from the moon as well as from the sun. As the sun drew angular momentum from the earth-moon system by braking the earth's spin, the moon would impart some of its angular momentum to the earth, enough to compensate for the sun's braking plus a little more. The result is that the earth's rotation would speed up again and the moon would approach the earth, thus decreasing its moment of inertia around the center of mass of the earth-moon system.

The possibility that the moon may once have been as close as a few earth radii deserves some additional comment. We can estimate that tides several kilometers in height would have sloshed about as the earth rotated with a five-hour day under a moon whose orbital period would have been only slightly longer. The energy dissipated by tidal friction would have exceeded the solar heating of the earth. Much of the ocean water would have evaporated, and there is the additional possibility that the earth's mantle would have melted to a considerable degree. All this activity would have been confined to the short time that the moon remained near the earth—perhaps a few thousand years at distances of less than six earth radii. If this event was sufficiently early in the earth's history, subsequent evolution of the earth's surface could have erased all traces of it from the geological record. It might be expected that the moon, by virtue of its more stable surface, would be the place to look for the evidence of such a catastrophic event.

The returns from the two lunar landing missions *Apollo 11* and *Apollo 12* have already enormously increased our understanding of the origin of the moon. It is fair to say, however, that we cannot yet distinguish between the three rival theories of lunar origin. Nevertheless, there are clues in the data obtained so far that bear rather directly on all these hypotheses.

Analysis of the rocks returned by *Apollo 11* indicates that they crystallized 3.6 billion years ago; similar analysis of the rocks from *Apollo 12* shows that they crystallized only 3.25 billion years ago. These crystallization ages are substantially different, ruling out the simultaneous filling of maria ("seas") all over the moon. The filling of the maria must not have been associated with any unique event in lunar history. It had been speculated that the flooding of the lunar maria resulted from melting induced by tidal friction in the moon following its capture by the earth. This idea can now be laid to rest.



ACCRETION HYPOTHESIS suggests that the moon was formed after the earth from material of somewhat different composition in earth orbit. As yet there is no compelling reason to prefer any particular distance from the earth as the site of the moon's formation.

Before the Apollo missions relative ages had been assigned to various areas on the moon by counting craters within those areas. It was reasoned that the more craters there were in an area, the older that area was, since it must have been exposed for a longer period of time to accumulate larger numbers of craters. Because the flux of impacting bodies in the past was entirely unknown, absolute ages could not be deduced. The radioactive dating of the returned samples has for the first time provided absolute ages for portions of the lunar surface. These age determinations lead to the conclusion that the flux of impacting bodies diminished by about three orders of magnitude (1,000 times) during the first billion years after the formation of the moon. This conclusion is forced on us by the fact that the flux of impacts on the highlands is about three orders of magnitude greater than the flux responsible for the craters on the lower Mare Tranquillitatis (the *Apollo 11* site). Since the age of the highlands must be less than 4.6 billion years (the probable age of the moon), they are at most one billion years older than Mare Tranquillitatis. The high initial bombardment rate is almost certainly connected with the process by which the moon was assembled out of smaller bodies.

Analysis of the returned samples indicated that the mare material was chemically fractionated at about the time the moon formed. We do not yet know to what extent this fractionation took place in space before the moon assembled and how much of it was due to subsequent internal fractionation in the body of the moon. The problem is probably connected to the difference in composition between the earth and the moon. Its solution would be a significant step forward. We have also learned that the composition of the mare rocks and soil cannot be representative of the bulk lunar composition [see "The Lunar Soil," by John A. Wood; SCIENTIFIC AMERICAN, August, 1970].

The return of lunar-highland samples is eagerly awaited. Perhaps these, the oldest rocks on the moon, will reveal the crucial evidence we need to distinguish between the rival theories of lunar origin. It is more likely that we shall still be speculating about the origin of the moon a century hence.

We want to be useful ...and even interesting

Kodak

Knowledge by eye

So does your library

Our biggest-ticket item is the KODAK KOM-90 Microfilmer. It translates from magnetic tape to microfilm at 120,000 characters a second. Fast way to milk a computer and micropublish the output in manageable form. University of Pittsburgh has one.

A project is afoot there to find out what a community of scholars wants of its library. Sure, scholars claim to want to savor of the corpus of all recorded knowledge and thought, in order to remedy its deficiencies. But what do they *really* want nowadays?

Schemes for automating the corpus have not been lacking. The more ambitious, the better. All come up against the patron's urge to crawl out from under the edge of an all-encompassing scheme, proving it is impossible to anticipate how *I* think. Institutions likewise treasure their individuality.

The people at Pittsburgh are examining how an individual institution's processing of the MARC (*M*achine-*R*eadable *C*ataloging) tapes from the Library of Congress can best serve its community. The planners of MARC wisely stuck to extant technology. Visitors from Kodak bring to Pittsburgh inside knowledge of where science might build technology if customers want it.

For contact with the project, write Knowledge Availability Systems Center, University of Pittsburgh, Pittsburgh, Pa. 15213. To bone up on MARC before troubling Pittsburgh, send $2.25 to the Superintendent of Documents, Washington, D.C. 20402 for Document LC 1.2:M 18/5. To buy or rent microfilming equipment, look up Kodak under that heading in the Yellow Pages.

Vision extended by ZIP code

Now that enough ecologists and other biologists are extending their vision into the near infrared by means of 35mm KODAK EKTACHROME Infrared Film in 20-exposure magazines, processing comes easy. A KODAK Prepaid Processing Mailer PK20, sold by almost any photo dealer, now does it. Back come slides from the nearest Kodak Processing Station, ready for the projector, just as though *everybody* expected to see green as blue, red as green, and infrared as red.

. . . but you may get trapped

So, having noted the above, one loads up a 35mm camera, mounts a KODAK WRATTEN Filter No. 12 in front of the lens to keep out blue light, and hires a light plane.

It's a beginning. Whether aloft, in the field, or in the lab, the most you have a right to expect from the first trial of infrared color film is encouragement. You are trying to put one of your private and more subjective sensory channels at the disposal of science. For science's sake, objectivity must be sought. Between the colors perceived and the hard facts conveyed through infrared reflectance, tight linkage is needed. Kodak color products and services that cover the visible spectrum sell well without a claim to fidelity. In the infrared the claim would be meaningless. Reproducibility is what matters.

Unless that first set of slides back from the processing station has answered all questions and exhausted the line of inquiry, more photography is called for. Perhaps quite a bit more before one learns the look that one is looking for. One would like to be able to count on constancy of rendition.

A very small difference in the proportions of three dyes at a given point in the picture can make a very big difference in the look. That it works at all is a miracle. Yet daily in darkrooms the world around, dyes are made to form in something like the right proportions, through manual methods, in small tanks, by people whose principal interest is *not* photographic processing. It stands to reason that automatic equipment in continuous operation under control of specialists might yield more reproducible results. It further stands to reason that for even better reproducibility the specialists should concentrate on that, rather than on providing enjoyment to millions of families the way Kodak processing specialists do. An important inquiry where photography is important may deserve its own specialists.

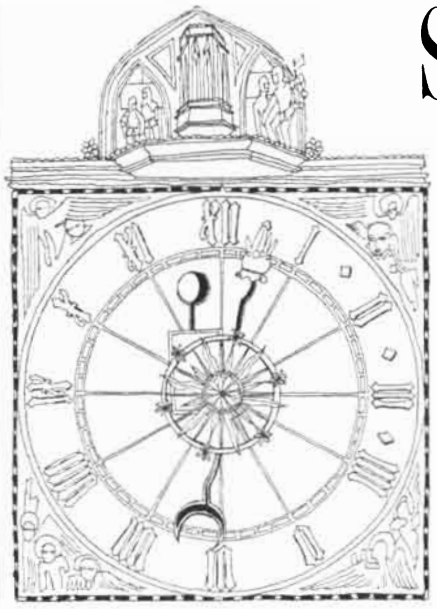
The trend has been strong in recent years toward high-volume processing. The chemical engineers who design the machines and take pride in their efficiency want high operating temperatures. This influences the design of the films.

Therefore where color photography—infrared or not—finds important use as a tool of investigation more than just communication, it is well to maintain frequent contact with Kodak, Scientific Photography Markets, Rochester, N.Y. 14650.

We also feel obligated to point out that direct infrared photography *without color* is less demanding while still intriguing. What would the student conclude from this comparison?

PANCHROMATIC

INFRARED



The Chemistry of Disarmament

The main obstacle now confronting the negotiators at the ongoing conference on chemical disarmament in Geneva is the question of what type of verification procedures should be included in a treaty banning the production, development and stockpiling of chemical-warfare agents. In an effort to facilitate these discussions 20 scientists from 13 countries, including the U.S. and the U.S.S.R., met in Stockholm in September under the auspices of the Stockholm International Peace Research Institute (SIPRI) to consider possible inspection techniques for verifying the nonproduction of organophosphorus compounds, a class of chemical compounds to which some of the most powerful chemical-warfare agents namely the nerve gases, belong. The final report of the symposium, containing the main joint conclusions of the participants and summaries of the papers presented, has recently been released.

As the convener of the SIPRI symposium, the Swedish physiologist S. J. Lundin, points out in the report, the fact that the symposium was confined to the organophosphorus compounds "was not intended to imply that CW [chemical warfare] agents other than the nerve gases are unimportant." Rather these compounds were chosen not only because they include the most important CW agents—the nerve agents—but also "to facilitate the discussion of a feasible model of a verification system."

Lundin's own contribution to the background papers describes the physiological effects of the nerve agents. In brief, he writes, "very low concentrations of organophosphorus compounds inhibit the activity of cholinesterase, an enzyme responsible for important physiological processes in the nervous system. Cholinesterase also occurs in both the plasma and the red cells of the blood. The inhibitory effect of the organophosphorus compounds is the basis for their use both as insecticides and as chemical-warfare agents."

The inhibition of cholinesterase, Lundin adds, can also be utilized to detect the presence of an organophosphorus compound in the environment or in an organism. Such detection methods may be the basis for limited on-site inspections of possible production and storage facilities for the compounds.

In general the symposium report emphasizes the need for both national (unilateral) and international verification organizations. Most verification activities regarding chemical disarmament, it is assumed, "will encounter only normal industrial operations, and the deterrence concept implies that verification methods do not need to be 100 percent efficient."

In particular the symposium focused on the question of how a possible international organization could process national data on the production and flow of elemental phosphorus. For example, diversions from established trends in the worldwide flow of phosphorus might be taken as indicating possible clandestine production of nerve agents. In this connection, the report points out, "the work of devising a practicable system based on the processing of national data on elemental phosphorus production should be facilitated by the fact that there are only about 35 phosphorus-producing plants in the world today."

With regard to the problem of differentiating nerve agents from those organophosphorus compounds that have peaceful uses only, it was concluded that "it might be possible to obtain practicable differentiations. These might be based on the principal structural formulae of the compounds together with data on their physical, chemical and toxicological properties."

The possibility that chemical plants, particularly those producing organophosphorus insecticides, could be converted to the production of nerve agents

or other CW agents was also considered. In this event, however, both the size of the plant and economic considerations were judged probable limiting factors.

There was general agreement among the participants in the SIPRI symposium that "should any type of 'on-site' inspection eventually be agreed upon, adequate analytical methods that could be successfully applied already exist. However, further work is needed to perfect and apply the existing methods, bearing in mind the probable future developments of CW agents that are not known today." Finally, it was proposed that "a working group should urgently study the problem of which data on the production and flow of elemental phosphorus, and possibly some other phosphorus compounds, a national control organization might select and produce for further processing by a possible international verification organization."

In conclusion the SIPRI report states that "the formulation of accurate technical methods of verifying the nonproduction of nerve agents—thereby solving the problem of those chemical agents which are of the greatest concern for reasons of national security—might well pave the way for the formulation of control systems for other CW agents."

Intercontinental Welfare

The U.S. and the U.S.S.R. have begun a program of cooperative research on heart disease, cancer and problems of environmental health. The program is an expansion of exchanges in the health area that have been going on since 1956. As outlined in an exchange of letters by Elliot L. Richardson, Secretary of Health, Education, and Welfare, and Boris V. Petrovsky, Minister of Health of the U.S.S.R., the expanded program involves setting up a Joint Committee for Health Cooperation. Moreover, according to Richardson's letter, the "agencies and persons responsible for the cooperation will establish direct contact, forming when necessary subcommittees or panels of experts from each country in each particular area."

In a press conference on the announcement of the agreement Richardson said the joint committee will develop plans "whereby Soviet and American scientists can identify areas in which their

interests converge and in effect work together on a basis that helps to assure that there will be not only an exchange of results among them but also collaboration in dividing areas in order to avoid duplication." It is envisioned that as time goes on the program will be expanded to take in problems in addition to heart disease, cancer and environmental health.

Toward a 1,024-Joule Hamburger

Weight-watchers confront the prospect, if the U.S. adopts the metric system, of fixing in their minds the fact that a wedge of apple pie provides 1,463 joules of energy and that three ounces of regular hamburger contain the energy equivalent of 1,024 joules. The reason is that calories are not counted in the metric system; as the Royal Society's committee on nutritional science put it, "the joule should be adopted as the unit of energy in all nutritional work and the calories should fall into disuse." The joule, which is the metric unit of energy, is defined as the work done in moving one kilogram of mass through one meter of space in one second of time. A food's energy value in joules can be closely approximated by multiplying the caloric value (which is usually expressed in kilocalories) by 4.18.

The prospect of such a change in terminology is a matter of concern among nutritionists. In one of several discussions of the subject the *Journal of the American Dietetic Association* noted that "the changeover may very well have traumatic overtones, both to professionals and laymen." Inviting comments, the journal received and printed one from Thomas Moore of the University of Cambridge, who drew attention to the fact that the joule refers mainly to mechanical work, whereas in biology the expenditure of energy occurs mainly in the form of heat. He proposed a compromise: "A more realistic, and less laborious, change might be to redefine the calorie as a unit of food, rather than of heat. On this basis the 'new calorie' could be conveniently defined as that amount of any food which, when used in the animal body exclusively for the generation of energy, produces 4,180 joules."

Stout Hearts

It has been five years since the human heart was first transplanted. What can now be said of the procedure and its value? An indication is provided by the *British Medical Journal,* which discusses the long-term results obtained by Norman E. Shumway of the Stanford University School of Medicine. The verdict is that, all things considered, year-old transplanted hearts perform remarkably well.

Within two years after the first heart transplant in 1967 more than 100 patients throughout the world elected to receive a transplanted heart. Even though the operation was often successful and the transplanted heart initially functioned well, it soon became apparent that the body's efforts to reject the implanted foreign tissue led to progressive deterioration and early death. Shumway, who developed the basic operation for transplanting the heart, reports on 26 patients who had heart transplants at Stanford during 1970. Thirteen of the patients died in the hospital. What is encouraging is that 75 percent of the patients who were alive two months after the operation were still alive after one year. Thirteen patients left the hospital and seven were still alive at the end of two years. Although the initial mortality is high, the patients faced an even higher risk of death without the operation.

The transplanted hearts in 10 of the patients were studied, and in each case the heart was found to be totally without functioning nerves. The heart nevertheless responded to exercise. During the first three minutes of exercise there was only a small increase in heart rate but the diastolic pressure of the left ventricle went up, thus increasing the flow of blood. With continuing exercise the heart rate also went up, presumably in response to catecholamines released into the blood by the central nervous system. In the normal heart catecholamines are released from nerve endings within the organ. The two mechanisms for increasing blood flow operate simultaneously in the normal heart. Because the transplanted heart is denervated, the two effects are sequential. The output of the transplanted heart rose steadily with exercise at a rate parallel to but lower than the rate for a normal heart.

The *British Medical Journal* notes that work on the problem of suppressing the body's efforts to reject foreign tissue has been intensive, and that there are signs that better control of the immunological problems will be achieved and that the high mortality rate from acute rejection (too little suppression) or from infection (often a sequel of too much suppression) will be lowered substantially.

Country Blues

A standard image of rural life includes a tidy, white-painted farmhouse owned by a prosperous white farmer whose apple-cheeked wife and children are notably healthier and better nourished than their city cousins. The image is wrong, it was concluded at a recent symposium conducted under the auspices of the National Research Council. With respect to prosperity, whereas only 30 percent of the U.S. population are rural residents, 40 percent of the nation's poor are. Per capita income in the 2,250 rural counties in the U.S. is below the national average; these same counties have nearly half of all U.S. families with incomes of less than $3,000 a year. Moreover, only one rural family in five is actually engaged in farming. Of these farm families some 1.6 million—almost 70 percent of all Americans who work on the land—receive only 18 percent of the $3 billion in annual Government subsidies that were first conceived as a weapon against rural poverty. The payments to more than a million of the families average less than $250.

As for the tidy farmhouse, 4.8 million of the nation's 8.5 million substandard dwellings are in rural areas. One farm family in four lives in substandard housing, compared with one urban family in 12. More than four million U.S. dwellings, 90 percent of them in rural areas, are without piped water. Some 6.9 million dwellings have neither a bathtub nor a shower; 5.4 million of them are rural. One-third of the nation's farmhouses lack flush toilets.

With respect to nutrition, the picture is a little brighter. Some of the poorest farm families maintain better standards of nutrition than either their poor nonfarming neighbors or the urban poor. This is most notable in the Northeast, where only 17 percent of farm poor are badly nourished, compared with 35 percent among their neighbors and 35 percent among the nation's urban poor, and 43 percent of the poor farmers in the South. Just how apple-cheeked everyone is, however, remains open to question. The rates of infant and postnatal mortality are higher in rural than in urban areas, the number of draftees rejected on health grounds is much higher, and some 19 percent of the adults on farms suffer from such chronic ailments as hypertension, arthritis and heart disease to an extent that limits their capacity to work. One change for the better in health care has taken place in recent decades: whereas only 57 percent of rural births were attended by physicians in 1944, the figure is now above 90 percent.

In summary, the conferees found that the cliché of the happy farmer fits less than half of all farm families and less than a tenth of the nonfarming rural pop-

ulation. The collective plight of the remainder—resident blacks in the South, Mexicans and Indians in the West, Puerto Ricans in all areas and a multitude of migrant rural families—accounts for the generally substandard level of U.S. rural life.

Through a Gas Darkly

Aided by a new television-camera tube developed at Princeton University, the 200-inch Hale telescope on Palomar Mountain in California has revealed hitherto unobtainable details of gaseous cloud banks in space, some of which may be associated with a quasar thought to be the second most distant object yet discovered. The new findings, according to Donald C. Morton of Princeton, could help to resolve the long-standing question of whether quasars are enormously energetic objects at huge distances or much less energetic objects at relatively modest distances.

The quasar, which is designated PHL 957, is thought to be nine billion light-years away. The modified SEC (secondary-electron conduction) Vidicon television tube has made it possible to observe in detail the layering of clouds between the quasar and the earth. According to Maarten Schmidt of the Hale Observatories, the group working with the instrument has found six definite layers and several possible ones.

There are at least two possible explanations for the layering. One is that the clouds are shells or puffs of gas recently ejected by the quasar. Another is that the clouds are random collections of intergalactic dust and gas not associated with the quasar that happen to be scattered along the line of sight. If it can be demonstrated that at least some of the clouds are at intermediate distances, it would provide support for the conclusion that quasars are at huge distances.

The clouds make themselves visible as dark absorption lines against the bright emission spectrum of PHL 957. With the high sensitivity of the new television system, many of the lines can be studied in much more detail than was previously possible. Several sets of lines show a different amount of red shift, indicating that the absorbing clouds are at different distances. One or two sets appear to be red-shifted virtually as much as the emission lines of the quasar itself, which suggests that some of the clouds are at about the same distance as the quasar and are probably associated with it. The question now is whether the remaining clouds were ejected by the quasar or simply happen to lie between the earth and the quasar. The television-tube observations are continuing.

Poor Richard's Pots

What are the odds that an archaeologist digging in Philadelphia would find artifacts authentically associated with Benjamin Franklin? They must be very long indeed, but they have been beaten by a graduate student at the University of Pennsylvania. Digging down through the compacted layers of debris in a deep brick-lined pit at Franklin Court on Market Street, Barbara Liggett and her colleagues (who are excavating the 18th-century site on behalf of the National Park Service) found a series of strata containing quantities of broken glass and pottery. The pottery included red and yellow earthenware bowls, pots, mugs, trivets and other kitchen and table articles. Some of the earthenware had collapsed when it was fired; in association with these pots were fragments of "kiln furniture," the clay stands that potters of the period used to support their pots in the kiln at firing time.

How does Franklin enter on the scene? When the young printer first arrived in Philadelphia in 1723, he took up lodgings on Market Street with a family named Read and soon became engaged to a daughter named Deborah. Franklin left Philadelphia for London in 1724; during his absence Deborah married a ne'er-do-well potter, one Rogers, whose kiln was near the Read house in the 1720's. Like Franklin, Rogers lived for a time in the Read house, but it was not long before he deserted his bride and disappeared from Philadelphia.

Returning from London in 1726, Franklin pursued his interrupted courtship of Deborah. The two were married in 1730 and lived in rented quarters for the next 35 years. During that time various affairs kept Franklin busy at home and abroad. It was not until 1765 that he began to build in the middle of the block near the Read house, enabling Deborah to move at last into a home of her own.

Only then, apparently, did the thrifty Deborah dispose of the pottery that her first husband had given her. The wasted ware and the kiln furniture found in the Market Street pit match similar articles made by Rogers that have been found near the site of his kiln. Liggett believes that some of the unique earthenware in the pit is also the work of Rogers, and that the ware was dumped there some 40 years afterward when Mrs. Franklin cleared out the Read family home before moving into her new house.

The Structure and History of an Ancient Protein

To oxidize food molecules all organisms from yeasts to man require a variant of cytochrome c. Differences in this protein from species to species provide a 1.2-billion-year record of molecular evolution

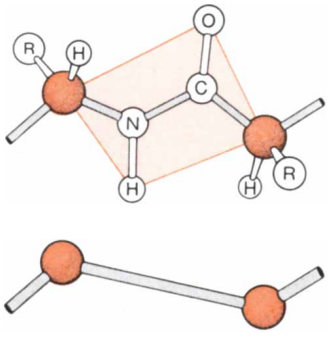
by Richard E. Dickerson

Between 1.5 and two billion years ago a profound change took place in some of the single-celled organisms then populating our planet, a change that in time would contribute to the rise of many-celled organisms. The machinery evolved for extracting far more energy from foods than before by combining food molecules with oxygen. One of the central components of the new metabolic machinery was cytochrome *c*, a protein whose descendants can be found today in every living cell that has a nucleus. By studying the cytochrome *c* extracted from various organisms it has been possible to determine how fast the protein has evolved since plants and animals diverged into two distinct kingdoms and in fact to provide an approximate date of 1.2 billion years ago for the event. For example, the cytochrome *c* molecules in men and chimpanzees are exactly the same: in the cells of both the molecule consists of 104 amino acid units strung together in exactly the same order and folded into the same three-dimensional structure. On the other hand, the cytochrome *c* in man has diverged from the cytochrome *c* in the red bread mold *Neurospora crassa* in 44 out of 104 places, yet the three-dimensional structures of the two cytochrome *c* molecules are essentially alike. We think we can now explain how it is that so many of the 104 amino acid units in cytochrome *c* are interchangeable and also why certain units cannot be changed at all without destroying the protein's activity.

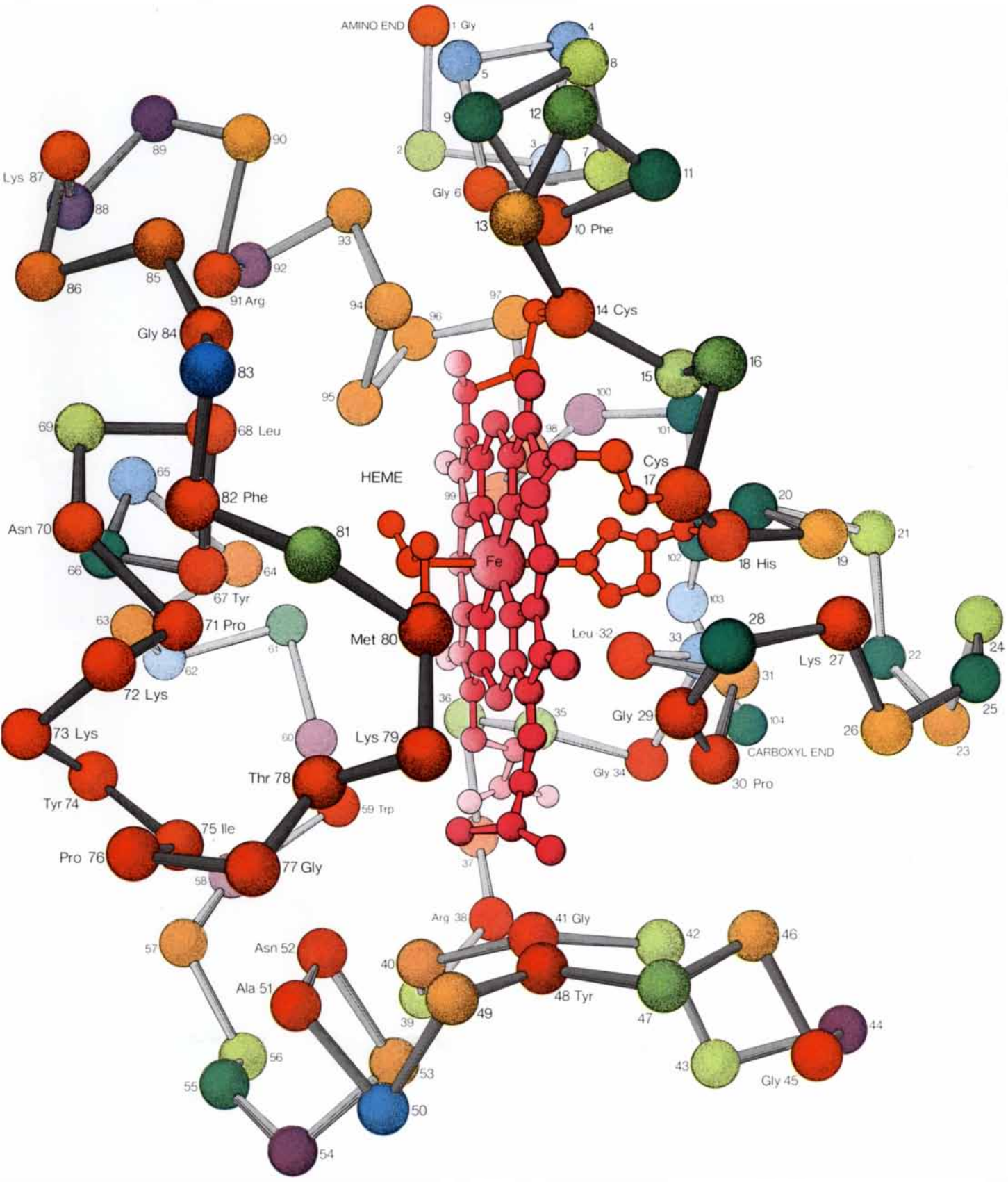
Let us try to visualize the earth before cytochrome *c* first appeared. The first living organisms on the planet were little more than scavengers, extracting energy-rich organic compounds (including their neighbors) from the water around them and releasing low-energy breakdown products. We still have the "fossils" of this life-style in the universal process of anaerobic (oxygenless) fermentation, as when a yeast extracts energy from sugar and releases ethyl alcohol, or when an athlete who exercises too rapidly converts glucose to lactic acid and gets muscle cramps. Anaerobic fermentation is part of the common biochemical heritage of all living things.

The upper limit on how much life the planet could support with only fermentation as an energy source was determined by the rate at which high-energy compounds were synthesized by nonbiological agencies: ultraviolet radiation, lightning discharges, radioactivity or heat. When some organisms developed the ability to tap sunlight for energy, photosynthesis was born and the life-carrying capacity of the earth increased enormously. This was the age of the bac-

SKELETON OF CYTOCHROME *c* MOLECULE is depicted in the illustration by Irving Geis on the opposite page. A variant of this protein molecule is found in the cells of every living organism that utilizes oxygen for respiration. The illustration shows in simplified form how 104 amino acid units are linked in a continuous chain that grips and surrounds a heme group, a complex rosette with an atom of iron (*Fe*) at its center. The picture is color-coded to indicate how much variation has been tolerated by evolution at each of the 104 amino acid sites in the molecule. Some species lack the 104th amino acid, and all species except vertebrates have as many as eight extra amino acids at the beginning of the chain (*see table on pages 60 and 61*). The amino acids that are most invariant throughout evolution, and presumably the most important, are shown in red and orange; the more variable sites appear in yellow-green, blue-green, blue and purple. The indispensable heme group is crimson. Each amino acid is represented only by its "alpha" carbon atom: the atom that carries a side chain unique for each of the 20 amino acids. The upper drawing at left below shows how two amino acids link up through an amide group (*colored panel*); the side chains connected to the alpha carbons (*color*) are represented by the balls labeled *R*. The lower drawing at left below shows the scheme used in the cytochrome *c* skeleton on the opposite page; all amide linkages (–CO–NH–) are omitted and the only side groups shown are those that are attached to the heme. The amino acids at the 35 invariant sites of the cytochrome *c* molecule (*red*) are designated in abbreviated form (*see key at right below*).



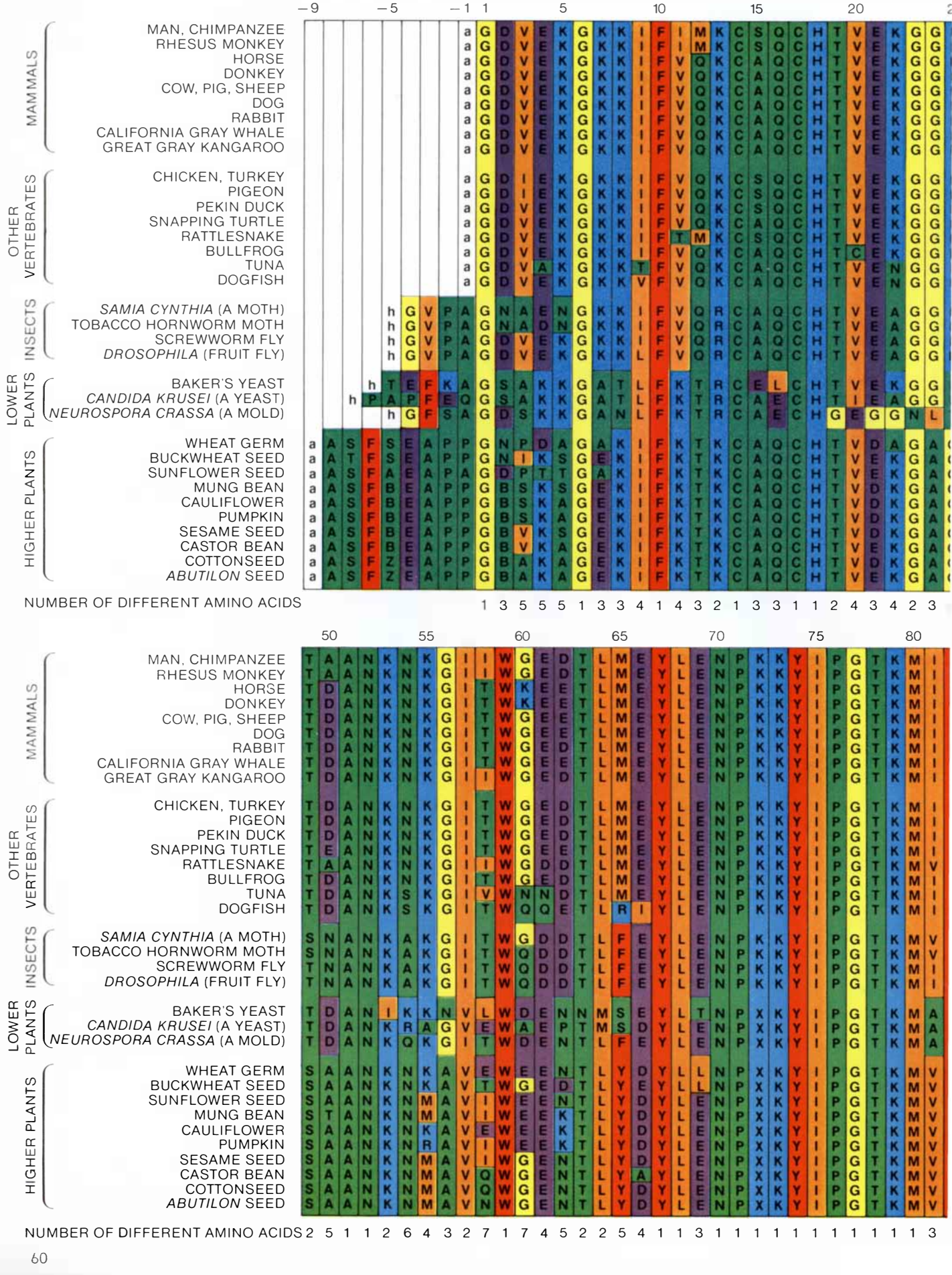
Ala	Alanine	Leu	Leucine
Asp	Aspartic acid	Lys	Lysine
Asn	Asparagine	Met	Methionine
Arg	Arginine	Phe	Phenylalanine
Cys	Cysteine	Pro	Proline
Gly	Glycine	Ser	Serine
Glu	Glutamic acid	Thr	Threonine
Gln	Glutamine	Trp	Tryptophan
His	Histidine	Tyr	Tyrosine
Ile	Isoleucine	Val	Valine

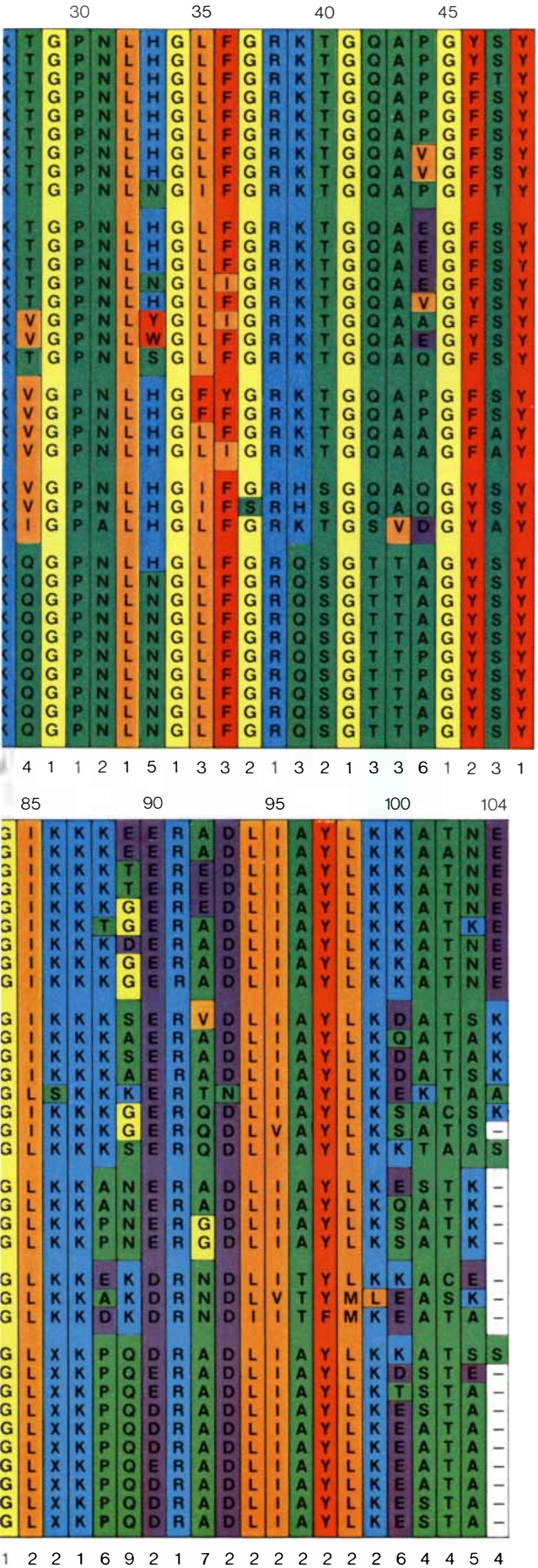


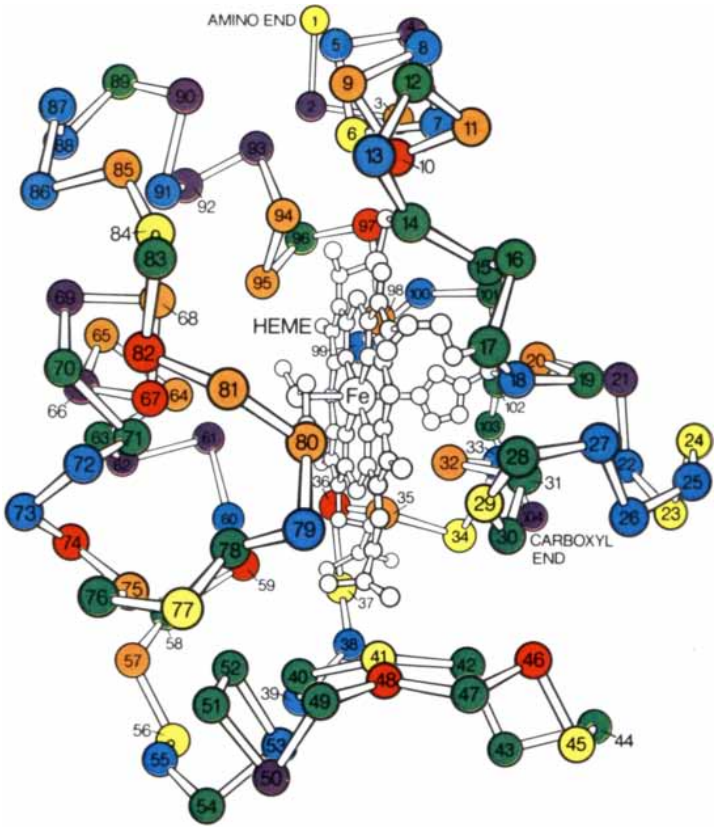
NUMBER OF DIFFERENT AMINO ACIDS FOUND AT A GIVEN POSITION IN 38 SPECIES

1	2	3	4	5	6 TO 9	
35	23	17	12	9	8	HEME

NUMBER OF SUCH AMINO ACID SITES IN THE MOLECULE







HYDROPHOBIC, AROMATIC RINGS

Phe Trp Tyr

HYDROPHOBIC, NOT AROMATIC

I Ile Leu M Met V Val

HYDROPHILIC, BASIC

His Lys Arg

methylated Lys

HYDROPHILIC, ACIDIC

Asp Glu

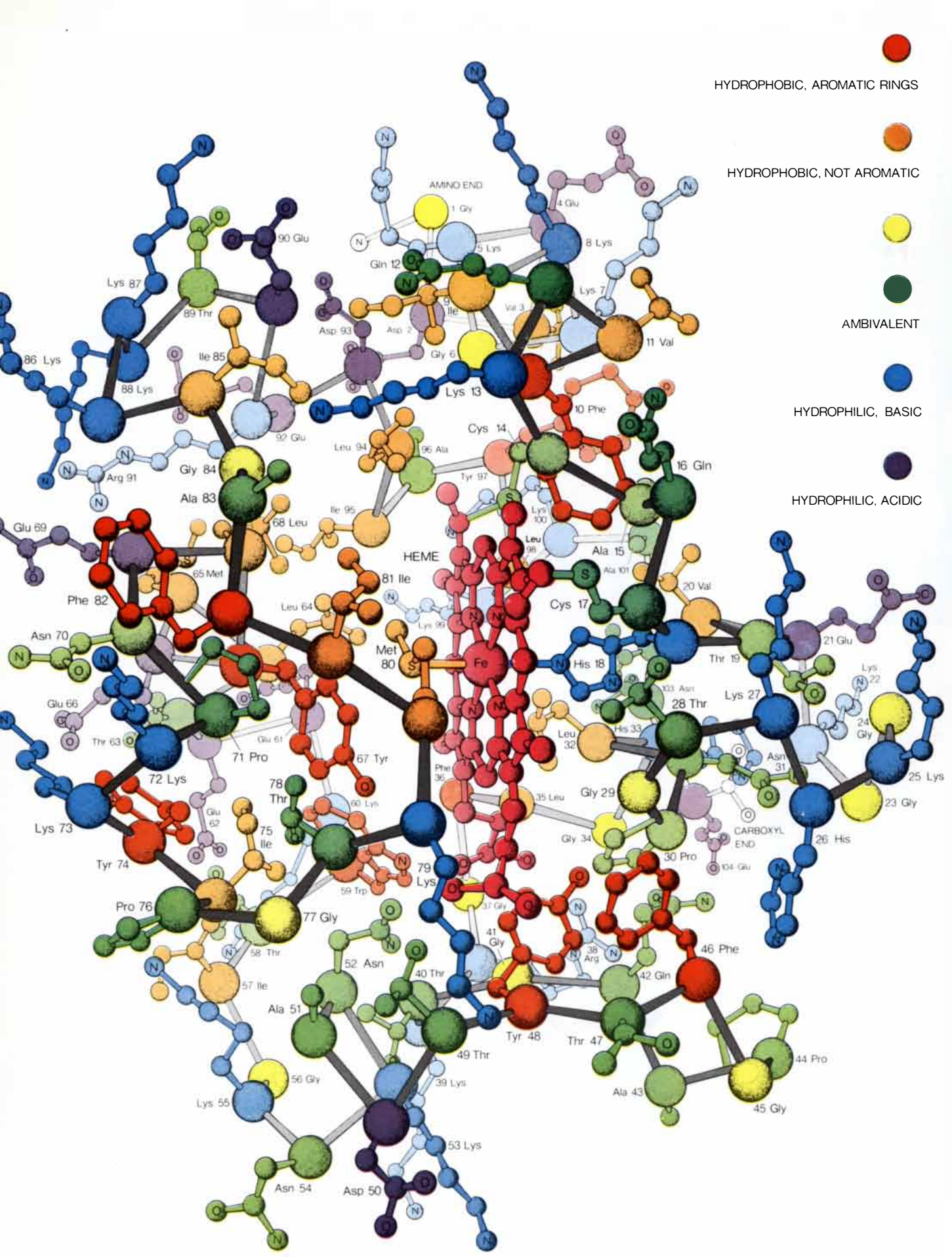
AMBIVALENT (HYDROPHOBIC BUT SMALL, OR POLAR BUT UNCHARGED)

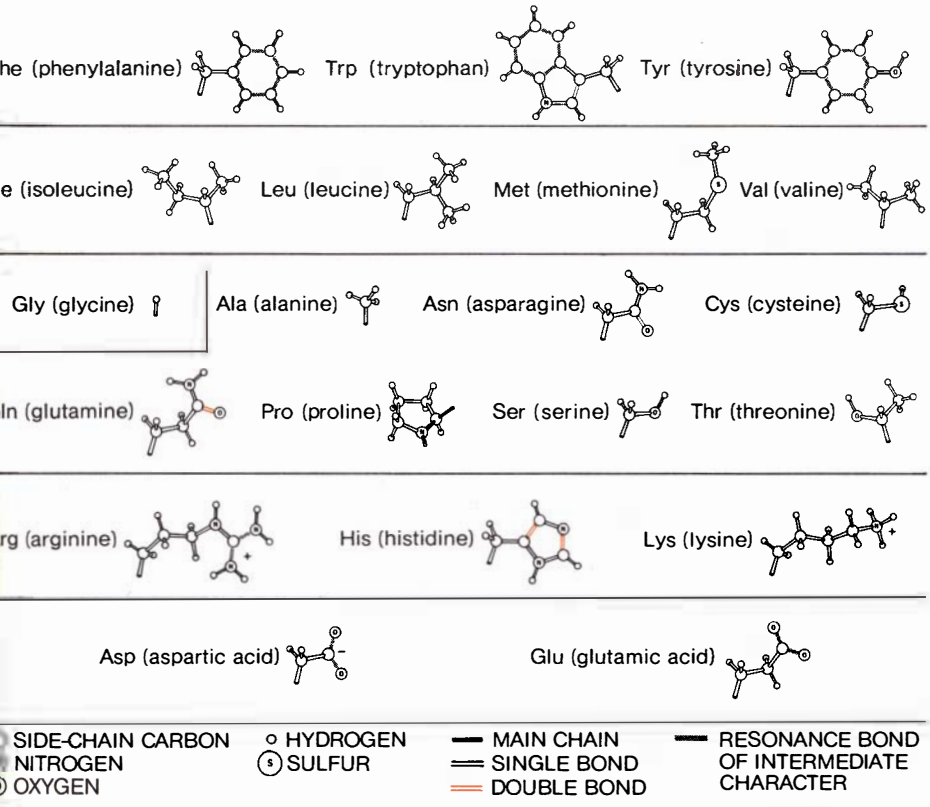
Ala Asn or Asp Cys Asn P Pro Q Gln

Ser Thr Gln or Glu

NO SIDE CHAIN (HYDROGEN ATOM) G Gly

COMPOSITION OF CYTOCHROME *c* IN 38 SPECIES is presented in the table at left. No other protein has been so fully analyzed for so many different organisms. The color code used here differs from the one used in the molecular skeleton on page 59. On these two pages color is employed to classify amino acids according to their chemical properties (*see key directly above*). Thus the three "oily" (hydrophobic) amino acids with aromatic benzene rings in their side chains (phenylalanine, tryptophan and tyrosine) are shown in red. Four other amino acids that are hydrophobic but nonaromatic are shown in orange. At the other extreme, amino acids that are hydrophilic, or water-loving, are shown in blue or violet. Amino acids that can be found in either aqueous or nonaqueous environments (and hence are ambivalent) are green or yellow. Polar amino acids can have asymmetric distributions of positive and negative charge. The detailed structure of the side chains of the amino acids can be found on page 63. It is easy to pick out from the table at left the amino acid sites where evolution has allowed no change or has allowed substitution only by chemically similar amino acids; these sites are identified by vertical bands of a single color. A letter *a* at the beginning of the chain indicates that a methyl group (CH_3) is attached to the amino end of the molecular chain. A letter *h* indicates that the methyl group is absent.





CYTOCHROME *c* MOLECULE WITH SIDE CHAINS appears in the illustration on the opposite page. The picture shows the structure of horse-heart cytochrome *c* in the oxidized state as determined through X-ray crystallography hy the author and his colleagues at the California Institute of Technology. Alpha-carbon atoms are numbered and the amide groups (–CO–NH–) connecting the alpha carbons are represented only by a solid bond, as in the preceding drawings. For clarity three side chains at the "back" of the molecule have been left out: leucine 35, phenylalanine 36 and leucine 98. The color coding follows the coding in the illustration on pages 60 and 61. One can see from the three-dimensional structure that side chains in the interior of the molecule, around the heme group (*crimson*), tend to be hydrophobic (*red and orange*), whereas amino acids with hydrophilic side chains (*blue and violet*) are found on the outside, where they are ordinarily in contact with water. A major exception to this rule is the hydrophobic side chain of phenylalanine 82, which sits on the surface of the molecule at the left of the heme. The region between the hydrophilic chains, and above isoleucine 81, is a cavity that is apparently open to solvent molecules. Lysine 13, above this cavity, is known to interact with a large oxidase complex when cytochrome *c* is oxidized. The structures of the side chains of 20 amino acids appear above.

teria and the blue-green algae [see "The Oldest Fossils," by Elso S. Barghoorn; SCIENTIFIC AMERICAN, May, 1971].

The more advanced forms of photosynthesis released a corrosive and poisonous gas into the atmosphere: oxygen. Some bacteria responded by retreating to oxygen-free corners of the planet, where their descendants are found today. Other bacteria and blue-green algae developed ways to neutralize gaseous oxygen by combining it with their own waste products. The next step was to harness the energy released by oxidation of these waste compounds. (If you are going to burn your garbage, you might as well keep warm by the fire.) This was the beginning of oxidation, or respiration, the second big breakthrough in increasing the supply of energy available to life on the earth.

When a yeast cell oxidizes sugars all the way to carbon dioxide and water instead of stopping short at ethyl alcohol, it gets 19 times as much energy per gram of fuel. When oxygen combines with lactic acid in the athlete's muscles and the cramps are dissipated, he receives a correspondingly greater energy return from his glucose. Any improvement in metabolism that multiplies the supply of energy available by such a large factor would be expected to have a revolutionary effect on the development of life. We now believe the specialization of cells and the appearance of multicelled plants and animals could only have come about in the presence of such a large new supply of energy.

Bacteria and blue-green algae are prokaryotes (prenuclear cells); their genetic material, DNA, is not confined within an organized nucleus, and their respiratory and photosynthetic machinery (if it is present) is similarly dispersed. Green algae and all the higher plants and animals are eukaryotes (cells with "good" nuclei); their DNA is organized within a nucleus, and their respiration is carried out in the organelles called mitochondria. In eukaryote plants photosynthesis is conducted in still other organelles called chloroplasts. Mitochondria are the powerhouse of all eukaryote cells. Their role is to break down the energy-rich molecules obtained from foods, combine them with oxygen and store the energy produced by harnessing it to synthesize molecules of adenosine triphosphate (ATP). The mitochondria of all eukaryotes are alike in their chemistry, as if once the optimum chemical mechanism had been arrived at it was never changed.

Biological oxidation involves at least a score of special enzymes that act first as acceptors and then as donors of the electrons or hydrogen atoms removed from food molecules. In the last part of the process one finds a series of cytochrome molecules (identified by various subscript letters), all of which incorporate a heme group containing iron, the same heme group found in hemoglobin. Electrons are passed down a chain of cytochrome molecules: from cytochrome b to cytochrome c_1, from cytochrome c_1 to cytochrome c, to cytochromes a and a_3 and finally to oxygen atoms, where they are combined with hydrogen ions to produce water. This is a stepwise process designed to release energy in small parcels rather than all at once. In the transfer of electrons from cytochrome b to cytochrome c_1 and again in the transfer between the cytochromes a, a_3 and oxygen, energy is channeled off to synthesize ATP, which acts as a general-purpose energy source for cell metabolism.

Most of the cytochromes are bound tightly to the mitochondrial membrane, but one of them, cytochrome c, can easily be solubilized in aqueous mediums and can be isolated in pure form. The other components can be isolated as multienzyme complexes: b and c_1 as a cytochrome reductase complex, and a and a_3 as a cytochrome oxidase. The reductase donates electrons to cytochrome c; the oxidase accepts them again. To illustrate how similar all eukaryotes really are to one another, it has been found that cytochrome c from any species of plant, animal or eukaryotic microorganism can react in the test tube with the cytochrome oxidase from any other species. Worm or primate, whale

or wheat are all alike under the mitochondrial membrane.

The Evolution of Cytochrome *c*

Since cytochrome *c* is so ancient and at the same time so small and easily purified, it has received much attention from protein chemists interested in the evolutionary process. The complete amino acid sequence of cytochrome *c* has been determined for more than 40 species of eukaryotic life. Thirty-eight of these sequences are compared in the illustration on pages 60 and 61. We have more information on the evolution of this molecule than on the evolution of any other protein.

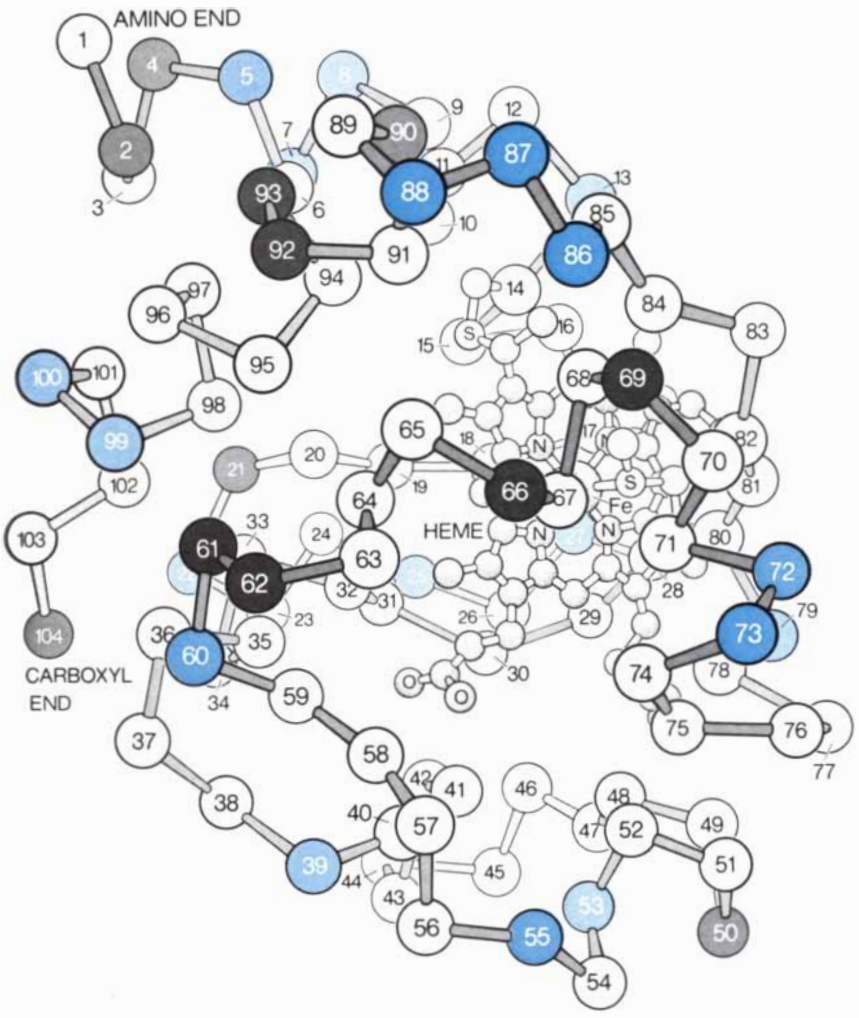
Emanuel Margoliash of Northwestern University and Emil Smith of the University of California at Los Angeles were among the first to notice that the amino acid sequences from various species are different and that the degree of difference corresponds quite well with the distance that separates the two species on the evolutionary tree. Detailed computer analyses of these differences by Margoliash, by Walter Fitch of the University of Wisconsin and by others have led to the construction of elaborate family trees of living organisms entirely without recourse to the traditional anatomical data. The family trees agree remarkably well with those obtained from classical morphology; it is obvious that comparison of amino acid sequence is a powerful tool for studying the process of evolution.

Another result may at first be surprising. Cytochrome *c* is still evolving slowly and is doing so at a rate that is approximately constant for all species, when the rate is averaged over geological time periods. This kind of analysis of molecular evolution was first carried out on hemoglobin a decade ago by Linus Pauling and Emile Zuckerkandl at the California Institute of Technology. If we compare hemoglobin and cytochrome *c*, we find that cytochrome *c* is changing much more slowly. Why should this be? The protein chains are synthesized from instructions that are embodied in DNA, and it is in the DNA that mutations take place. Do mutations occur more often in the DNA that makes hemoglobin than in the DNA that makes cytochrome *c*? There is no reason to think so. The explanation therefore must lie in the natural-selection, or screening, process that tests whether or not mutant molecules can do their job.

Before discussing the various "formulas" that have passed the test of making a successful cytochrome *c*, I shall describe briefly the structure of proteins. All protein molecules are built up by linking amino acids end to end. Each of the 20 different amino acids has a carboxyl group (–COOH) at one end and an amino group ($-NH_2$) at the other. To link the carboxyl group of one amino acid with the amino group of another amino acid a molecule of water must be removed, producing an amide linkage (–CO–NH–). Because only a part (although, to be sure, the distinctive part) of an amino acid enters a protein chain, the chemist refers to it as a "residue." Thus he speaks of a glycine residue or a phenylalanine residue at such-and-such a position in a protein chain.

The carbon adjacent to the amide linkage is called the alpha carbon. It is important because each amino acid has a distinctive side chain at this position. The side chain may be nothing more than a single atom of hydrogen (as it is in the case of the amino acid glycine) or it may consist of a number of atoms, including a six-carbon "aromatic" ring (as it does in the case of phenylalanine, tryptophan and tyrosine).

The 20 amino acids can be grouped into three broad classes, depending on the character of their side chains [*see illustration on preceding page*]. Five are hydrophilic, or water-loving, and tend to acquire either a positive or a negative charge when placed in aqueous solution; three of the five are basic in character



PLOT OF DISTRIBUTION OF ELECTRIC CHARGES on the back of horse-heart cytochrome *c* reveals that most of the 19 hydrophilic lysines (*color*), which carry positive charges (and hence are basic), are distributed on the two flanks of the molecule. Nine of 12 negatively charged (acidic) side chains (*gray*) are clustered in one zone in the upper center of the molecule. The electrically negative character of this zone has been maintained throughout evolution, although the specific locations of the acidic side groups vary. No organism, from wheat germ to man, has a cytochrome *c* with fewer than six acidic amino acids in this zone and no organism has more than five acidic amino acids everywhere else on the molecule. Furthermore, these extreme values are not found in the same species. It is highly likely that these charged zones participate in binding cytochrome *c* to other large molecules.

(arginine, histidine and lysine) and the other two are acidic (aspartic acid and glutamic acid). Seven are not readily soluble in water and hence are termed hydrophobic; they include the three amino acids mentioned above that have rings in their side chains plus leucine, isoleucine, methionine and valine. The remaining eight amino acids react ambivalently to water: alanine, asparagine, cysteine, glutamine, glycine, proline, serine and threonine.

Now let us see how much the successful formulas for cytochrome *c* differ from species to species. The cytochrome *c* molecules of men and horses differ by 12 out of 104 amino acids. The cytochrome *c*'s of the higher vertebrates—mammals, birds and reptiles—differ from the cytochrome *c*'s of fishes by an average of 19 amino acids. The cytochrome *c*'s of vertebrates and insects differ by an average of 27 amino acids; moreover, the cytochrome *c* molecules of insects and plants have a few more amino acid residues at the beginning of the chain than the equivalent molecules of vertebrates. The greatest disparity between two cytochrome *c*'s is the one between man and the bread mold *Neurospora;* they differ at more than 40 percent of their amino acid positions. How can two molecules with such large differences in amino acid composition perform identical chemical functions?

We begin to see an answer when we look at where these changes are. Some parts of the amino acid sequence, as indicated on pages 60 and 61, never vary. Thirty-five of the 104 amino acid positions in cytochrome *c* are completely invariant in all known species, including a long sequence from residue 70 through residue 80. The 35 invariant sites are occupied by 15 different amino acids; they are shown in red in the structural drawing on page 59. Another 23 sites are occupied by only one of two different but closely similar amino acids. There are 18 different sets of interchangeable pairs at these 23 sites; they are shown in orange in the illustration. At 17 sites natural selection has evidently accepted only sets of three different amino acids; these 17 interchangeable triplets are colored yellow-green.

It was already known from sequence studies, before the X-ray structural analysis, that where such substitutions are allowed the interchangeable amino acids almost always have the same chemical character. In general all must be either hydrophilic or hydrophobic or else neutral with respect to water. Such interchanges are called conservative substitutions because they conserve the overall chemical nature of that part of the protein molecule.

In only a few places along the chain can radical changes be tolerated. Residue 89, for example, can be acidic (aspartic acid or glutamic acid), basic (lysine), polar but uncharged (serine, threonine, asparagine and glutamine), weakly hydrophobic (alanine) or devoid of a side chain (glycine). Almost the only type of side chain that appears to be forbidden at this point in the molecule is a large hydrophobic one. Such "indifferent" regions are rare, however, and cytochrome *c* overall is an evolutionarily conservative molecule.

We have no reason to think the gene for cytochrome *c* mutates more slowly than the gene for hemoglobin, or that the invariant, conservative and radical regions of the sequence reflect any difference in mutational rate within the cytochrome *c* gene. The mutations are presumably random, and what we see in these species comparisons are the molecules that are left after the rigid test of survivability has been applied. Invariant regions evidently are invariant because any mutational changes there are lethal and are weeded out. Conservative changes can be tolerated elsewhere as long as they preserve the essential chemical properties of the molecule at that point. Radical changes presumably indicate portions of the molecule that do not matter for the operation of the protein.

This is as far as we can go from sequence comparisons alone. The explanation of variability in terms of the essential or nonessential character of different parts of the molecule is plausible, yet science has always been plagued by plausible but incorrect hypotheses. To progress any further we need to know how the amino acid sequence is folded to make an operating molecule. In short, we need the three-dimensional structure of the protein.

The Molecule in Three Dimensions

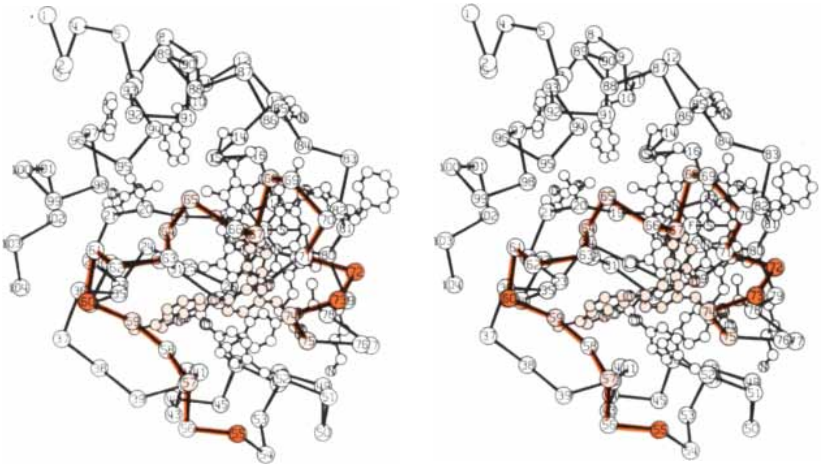
With the active collaboration of Margoliash, who was then working at the Abbott Laboratories in North Chicago, I began the X-ray-crystallographic analysis of horse-heart cytochrome *c* at Cal Tech in 1963, with the sponsorship of the National Science Foundation and the National Institutes of Health. As cytochrome *c* transfers electrons in the mitochondrion, it oscillates between an oxidized form (ferricytochrome) and a reduced form (ferrocytochrome); the iron atom in the heme group is alternately in the +3 and +2 oxidation state. We decided to begin our analysis with the oxidized form, a decision that was largely tactical since both oxidation states would ultimately be needed if we were to try to decipher the electron-transfer process.

In X-ray crystallography one directs a beam of X rays at a purified crystal of the substance under study and records the diffraction pattern produced as the beam strikes the sample from different angles. X rays entering the sample are themselves deflected at various angles by the distribution of electron charges within the crystal. Highly sophisticated computer programs have been devised for deducing from tens of thousands of items of X-ray-diffraction data the three-dimensional distribution of electronic charge. From this distribution one can infer, in turn, the distribution of the amino acid side chains in the protein molecule.

We obtained our first low-resolution map of the oxidized form of horse-heart cytochrome *c* five years ago and the first high-resolution map three years ago. These maps have been used to construct detailed three-dimensional models of the protein. One can also feed the three-dimensional coordinates into a computer and obtain simple ball-and-stick drawings that can be viewed stereoptically, enabling one to visualize the folded chain of the protein in three dimensions [*see illustrations on next two pages*]. Just a year ago we calculated the first high-resolution map for the reduced form of cytochrome *c*. We are now improving this model and comparing the two oxidation states.

Several striking features of the amino acid sequences of cytochrome *c* were in the back of our minds as we worked out the first high-resolution structure. We knew that the most strongly conserved sites throughout evolution were those occupied by three distinctive types of residue: the positively charged (basic) residues of lysine; the three hydrophobic and aromatic residues of phenylalanine, tryptophan and tyrosine, and the four hydrophobic but nonaromatic residues of leucine, isoleucine, methionine and valine. These sites can now be located with the help of the illustration on page 62, whose color coding differs from the coding of the illustration on page 59. Here hydrophobic residues are shown in warm colors (red and orange) whereas neutral residues and hydrophilic residues, both basic and acidic, are shown in cool colors (green, yellow, blue and violet).

It had been known from the chemical analysis of the molecule's amino acid se-



STEREOSCOPIC PAIR OF LEFT SIDE of oxidized cytochrome *c* molecule, drawn by computer, shows only a few key side chains for clarity. The main chain (*color*) from sites 55 to 75 defines a loop, the "left channel," which is filled with strongly hydrophobic side chains; their alpha carbons are in light color. Three of these side chains include aromatic rings: tryptophan 59, tyrosine 67 and tyrosine 74, also shown in light color. Alpha carbons with hydrophilic, positively charged side chains around the left channel are shown in dark color. This pair and one on opposite page can be viewed with standard stereoscopic viewer.

quence that the basic and hydrophobic groups tend to appear in clusters along the chain. For example, basic residues are found in the regions of sites 22 through 27, sites 38 and 39, sites 53 through 55 and sites 86 through 91. Hydrophobic residues are found in regions 9 through 11, 32 through 37, 80 through 85 and 94 through 98. The residues at sites 14 and 17 (cysteine) and site 18 (histidine) are invariant, which is understandable since they form bonds to the heme group. Less understandably, the long stretch from site 70 to site 80 is equally invariant. Before the structural evidence was available it had been suspected that methionine, at site 80, might be bonded to the iron atom on the other side of the heme from the histidine at site 18, but it was impossible to be sure from chemical evidence alone.

It was also known from chemical analysis that horse cytochrome *c* incorporates 12 glycines (the residues with only hydrogen as a side chain) and that these glycines were either invariant or else conserved in the great majority of species. It was known too that of the eight phenylalanines or tyrosines (with aromatic rings in their side chains) seven are either invariant in all species or replaceable only by one another. In the case of residue 36, phenylalanine or tyrosine is replaced in three species by isoleucine, whose side chain, although it is nonaromatic, is at least as large and hydrophobic as the side chains it replaces.

All these similarities and conservatisms were known before the X-ray analysis, but none could be explained in terms of structure. It was assumed that every residue had been placed where it was by natural selection and that it contained potentially important information about the working parts of the cytochrome molecule. Natural selection, however, does not act on an amino acid sequence but rather on the folded and operating molecule in its association with other biological molecules. Having a sequence without the folding instructions is like having a list of parts without a blueprint of the entire machine.

Cytochrome *c* and Evolution

Now that the blueprint for cytochrome *c* is revealed, let us look more closely at its representation on page 59. To keep the illustration simple no side chains have been included except for those that are bonded to the heme group. Moreover, along the main chain the illustration depicts only the alpha-carbon atoms from which side chains would, if they were shown, branch off. The amide groups (–CO–NH–) that connect alpha carbons are represented simply by straight lines. The picture is therefore a simplified folding diagram of the cytochrome molecule.

We see that the flat heme group, a symmetrical rosette of carbon and nitrogen atoms with an atom of iron at its center, sits in a crevice with only one edge exposed to the outside world. If the heme participates directly in shuttling electrons in and out of the molecule, the transfer probably takes place along this edge. Cysteines 14 and 17 and histidine 18 hold the heme in place from the right as depicted, and the other heme-binding group on the left is indeed methionine 80, as had been suspected.

It was known from earlier X-ray studies of proteins that sequences of amino acids frequently fold themselves into the helical configuration known as the alpha helix; in other cases the amino acids tend to assume a rippled or corrugated configuration called a beta sheet. Cytochrome *c* has no beta sheets and only two stretches of alpha helix, formed by residues 1 through 11 and 89 through 101. For the most part the protein chain is wrapped tightly around the heme group, leaving little room for the alpha and beta configurations that are prominent in other proteins.

Just as one can use cytochrome *c* to learn about evolution, one can also use evolution to learn about cytochrome *c*. As I have noted, the illustration on page 59 is color-coded to indicate the amount of variability in the kind of amino acid tolerated at each site. The structure is "hot" (red and orange) in the functionally important places in the molecule where differences among species are absent or rare, and it is "cool" (green, blue and violet) in regions that vary widely from one species to another and thus are presumably less important to a viable molecule of cytochrome *c*.

The heme crevice is hot, indicating that strong selection pressures tend to keep the environment of the heme group constant throughout evolution. The invariant residues 70 through 80 are also hot, and we now see that they are folded to make the left side of the molecule and the pocket in which the heme sits. The right side of the molecule is warm, consisting of sites where only one, two or three different amino acids are tolerated. The back of the molecule is its cool side; residues 58 and 60 and four more residues on the back of the alpha helix are each occupied by six or more different amino acids in various species. These are powerful clues to the important parts of the molecule, whether for electron transfer or for interaction with two large molecular complexes, the reductase and the oxidase.

How the Molecule Folds Itself

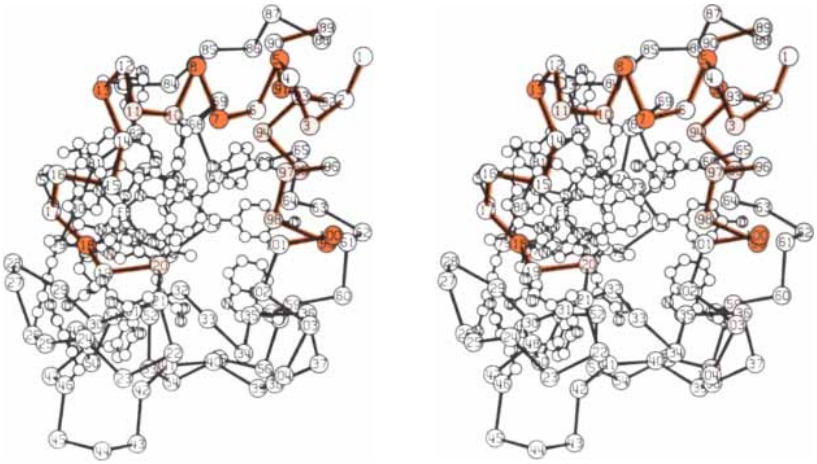
If we now turn to the illustration on page 62, which shows all the side chains of horse-heart cytochrome *c*, many of the evolutionary conservatisms become understandable. (As before, amide groups are still shown only as straight lines; their atomic positions are known but are not particularly relevant to this article.) In

this illustration the colors are selected to classify the various sites according to the character of the amino acid tolerated (hydrophilic, hydrophobic or ambivalent); the same color coding applies in the illustration on pages 60 and 61 showing the amino acid sequences in the cytochrome *c*'s of 38 different species.

Nonpolar, hydrophobic groups are found predominantly on the inside of the molecule, away from the external aqueous world, whereas charged groups, acidic or basic, are always on the outside. This arrangement is a good example of the "oil drop" model of a folded protein. According to this model, when an amino acid chain is synthesized inside a cell, it is helped to fold in the proper way by the natural tendency of hydrophobic, or "oily," side chains to retreat as far as possible from the aqueous environment and cluster in the center of the molecule. An even stronger statement can be made: If it is necessary for the successful operation of a protein molecule that certain portions of the polypeptide chain be folded into the interior, then natural selection will favor the retention of hydrophobic side chains at that point so that the proper folding is achieved. A charged, or hydrophilic, side chain can be pushed into the interior of a protein molecule, but a considerable price must be paid in terms of energy. Thus in most cases the presence of a charged group at a given site helps to ensure that the chain at that point will be on the outside of the folded molecule. (Charged groups inside a protein are known only in one or two cases where they play a role in the catalytic mechanism of the protein.)

We can now see the reason for the evolutionary conservatism of hydrophobic side chains, and one of the reasons for the conservatism of the hydrophilic residue lysine: they help to make the molecule fold properly. Radical changes of side chain that prevent proper folding are lethal. No folding, no cytochrome; no cytochrome, no respiration; no respiration, no life. It is seldom that cause and effect in evolution are quite so clear-cut.

There is still more to the lysine story. The lysines are not only on the outside; they are clustered in two positively charged regions of the molecular surface, separated by another zone of negative charge. This segregation of charge has not been found in any other protein structure and, as we shall see, probably occurs because cytochrome *c* interacts with two molecular complexes (the reductase and the oxidase) rather than with small substrate molecules as an enzyme does. The charge arrangements are believed to be part of the process by which large molecules recognize each other.

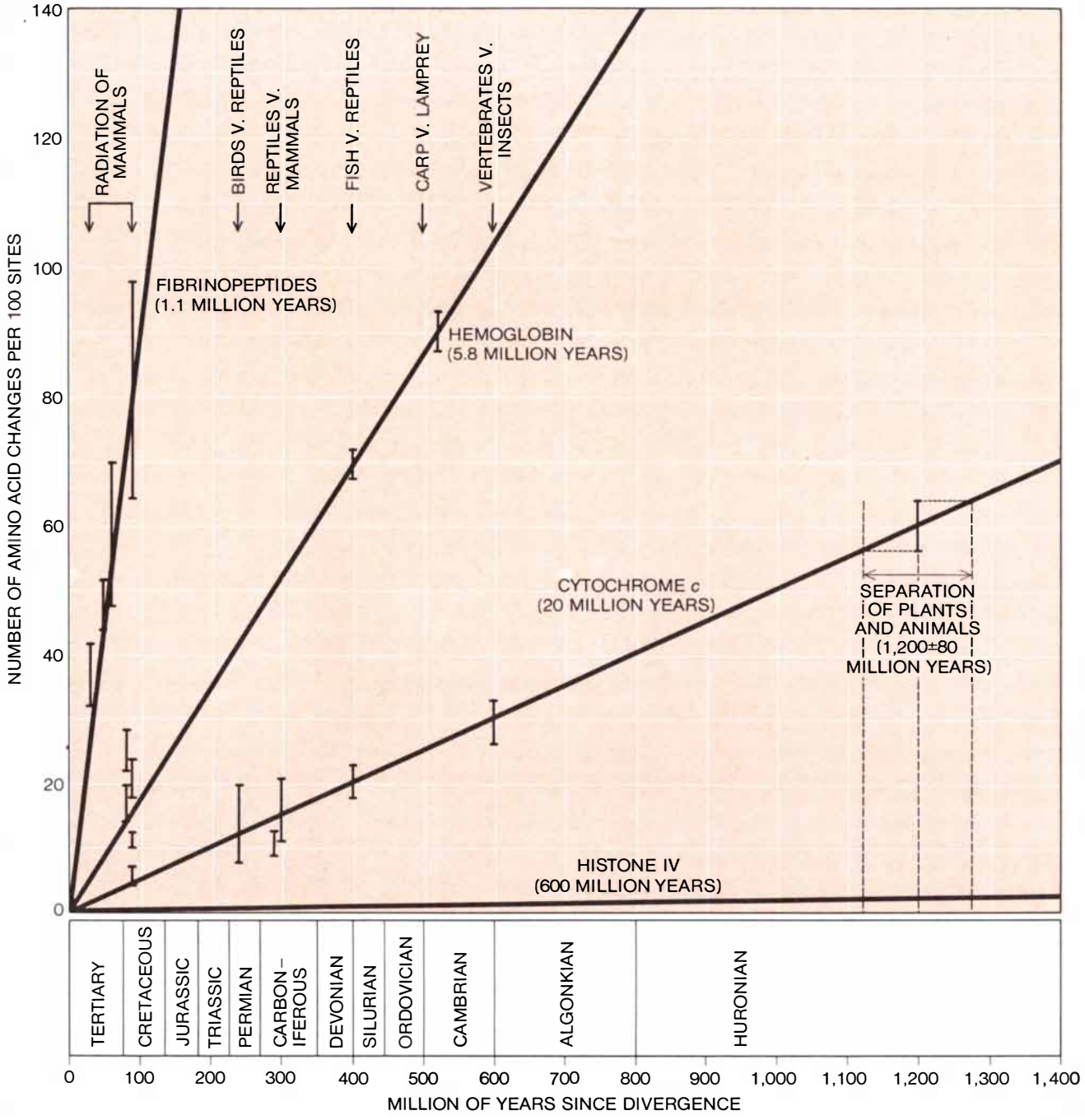


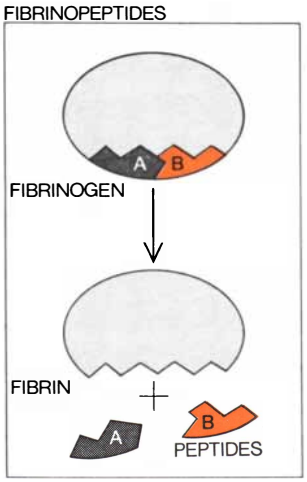
STEREOSCOPIC PAIR OF RIGHT SIDE of oxidized cytochrome *c* shows two sequences forming alpha helixes: the sequence from 1 through 11 and the sequence from 89 through 101. The two alpha helixes and the chain from 12 through 20 outline the right channel. Like the left channel it is lined with hydrophobic side groups, but it apparently contains a slot large enough to receive a hydrophobic side chain from another molecule. As in the stereoscopic drawing on the opposite page, alpha carbons with positively charged side chains around this channel are indicated in dark color; alpha carbons with strongly hydrophobic side chains are indicated in light color. The computer program for preparing the stereoscopic pictures was written by Carroll Johnson of the Oak Ridge National Laboratory.

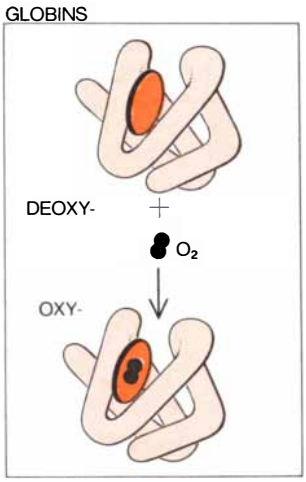
Most of the 19 lysines are found on the left and right sides of the molecule, viewed from the back on page 64. The left and right sides of the molecule can be examined separately in the two stereoscopic pairs on these two pages. On the left side eight lysines surround a loop of chain from sites 55 through 75 that is tightly packed with hydrophobic groups, including the invariant tyrosine 74, tryptophan 59 and tyrosine 67 farther inside. Although we do not yet know the electron-transfer mechanism, it has been suggested that the aromatic rings of the three invariant residues could provide an inward path for the electron when cytochrome *c* is reduced. Another eight lysines are found on the right side, on the periphery of what appears to be a true channel large enough to hold a hydrophobic side chain of another large molecule. This right-side channel is bounded by the two alpha helixes and by the continuation from the first alpha helix through residues 12 to 20. Within this channel are found two large aromatic side chains: phenylalanine 10 (which cannot change) and tyrosine 97 (which can also be phenylalanine but nothing else). In summary, on the right side is a channel lined with hydrophobic groups (including two aromatic rings) and surrounded by an outer circle of positive charges. As someone in our laboratory remarked on looking at the model, it resembles a docking ring for a spaceship.

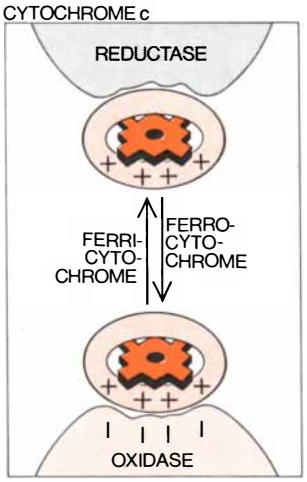
This remark may not be entirely frivolous. It is known from chemical work that the attraction between cytochrome *c* and the cytochrome oxidase complex is largely electrostatic, involving negatively charged groups on the oxidase and positively charged basic groups on cytochrome *c*. Either the left or the right cluster of lysines must be involved in this binding. Moreover, Kazuo Okunuki of the University of Osaka has shown that if just one positive charge, lysine 13, is blocked with a bulky aromatic chemical group, the reactivity of cytochrome with its oxidase is cut in half. Chemically blocking lysine 13 means physically blocking the upper part of the heme crevice. Lysine 13 is closer to the right cluster of positive charges than to the left; thus it would appear more likely that the heme crevice and the right channel together are the portions of the molecular surface that "see" the oxidase complex.

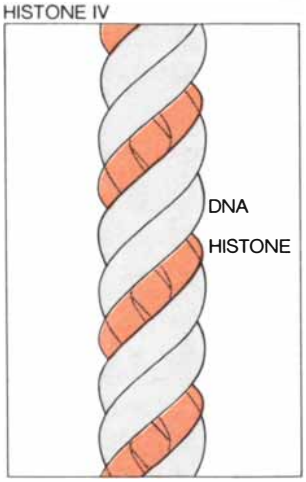
What, then, are the roles of the positive zone on the left side and the negative patch at the rear? The positive zone, with its three aromatic rings, may be the binding site to the reductase; we know virtually nothing about the chemical nature of this binding. The negative patch may be a "trash dump," an unimportant part of the molecule's surface where there are enough negative charges to prevent an excessively positive overall charge. The fact that the six most variable amino acid sites are in this part











of the molecule would support such an idea.

On the other hand, it is equally possible that this collection of negative charges has a function. Acidic amino acids are actually conserved throughout the various species, although in a subtle way that was overlooked in the earlier sequence comparisons. Selection pressures have kept this zone of the molecular surface negative, even though the individual residues that carry the negative charges differ from one species to another. Because several sections of the protein chain bend into and out of this acidic region, the conservation of negative charge is not immediately obvious if one looks only at the stretched-out sequence. This is a good illustration of the principle of molecular evolution that natural selection acts on the folded, functioning protein and not on its amino acid sequence alone.

If we look carefully at where the glycine residues are, we can appreciate why such a large number are evolutionarily invariant. The heme group is so large that 104 amino acids are barely enough to wrap around it. There are many places where a chain comes too close to the heme or to another chain for a side chain to fit in. It is just at these points that we find the glycines with their single hydrogen atom as a side chain.

The last type of conservatism, the conservatism of the aromatic side chains, is more difficult to explain. Tyrosines and phenylalanines tend to occur in nearby pairs in the folded cytochrome *c* molecule: residues 10 and 97 in the right channel, 46 and 48 below the heme crevice, 67 and 74 along with tryptophan 59 in the hydrophobic left channel. Only residue 36, which can be tyrosine, phenylalanine or isoleucine, seems to have merely a space-filling role on the back of the molecule; it is an "oily brick."

The three aromatic rings in the left channel may be involved in electron transfer during reduction. The two rings in the right channel could also be employed in electron transfer, or might only help to define the hydrophobic slot in the middle of that channel. Tyrosine 48 at the bottom of the molecule helps to hold the heme in place by making a hydrogen bond to one of the heme's propionic acid side chains. In cytochrome *c* from the tuna and the bonito, where residue 46 is tyrosine, electron-density maps have shown that this residue also holds the heme by a hydrogen bond to its other propionic acid group. These two tyrosines, along with cysteines 14 and 17, help to lock the heme in place in a way not seen in hemoglobin or myoglobin.

Phenylalanine 82 is the enigma. Never tyrosine or anything else, it extends its oily side chain out into the aqueous world on the left side of the heme crevice, where it has no visible role. A price must be paid in energy for its being there. Why should such a large hydrophobic group be on the outside of the molecule, and why should it be absolutely unchanging through the entire course of evolution? Viewing the oxidized molecule alone, it is impossible to say, but when at the end of this article we look briefly at the recently revealed structure of reduced cytochrome, we shall see the answer fall into place at once.

The structural reasons for the evolutionary conservatism in cytochrome *c* throughout the history of eukaryotic life can now largely be explained. Cytochrome *c* is unique among the structurally analyzed proteins in that it has segregated regions of charge on its surface. The roles assigned to these regions in the foregoing discussions have been speculative and may be quite wrong. What we can be sure of is that these regions do have roles in the operation of the molecule. Chance alone, or even common ancestry, could not maintain these positive and negative regions, along with paired and exposed aromatic groups, in all species through more than a billion years of molecular evolution. The conservative sequences are shouting to us, "Look!" Now we have to be clever enough to know what to look for.

Rates of Protein Evolution

With this background we are equipped to return to a question raised earlier: What determines the rates of evolution of different proteins? We begin by making a graph where the vertical axis represents the average difference in amino acid sequence between two species of organism on two sides of an evolutionary branch point, for example the branch point between fish and reptiles or between reptiles and mammals. The horizontal axis represents the time elapsed since the divergence of the two lines as determined by the geological record. If such a graph is plotted for cytochrome *c*, one finds that all the branch points fall close to a straight line, indicating a constant average rate of evolutionary change [*see illustration on opposite page*].

How can this be? How can cytochrome *c* change at so nearly a constant rate during the long period in which the external morphology of the organism was diversifying toward the present-day cotton plant, bread mold, fruit fly, rattlesnake and chimpanzee? This is an illustration of a fundamental advantage of proteins as tools in studying evolution. Natural selection ultimately operates on populations of whole living organisms, the only criterion of success being the ability of the population to survive, reproduce and leave behind a new generation. The farther down toward the molecular level one goes in examining living organisms, the more similar they become and the less important the morphological differences are that separate a clam from a horse. One kind of chemical machinery can serve many diverse organisms. Conversely, one external change in an organism that can be acted on by natural selection is usually the effect not of a single enzyme molecule but of an entire set of metabolic pathways.

The observed uniform rate of change in cytochrome *c* simply means that the biochemistry of the respiratory package, the mitochondrion, is so well adjusted, and the mitochondrion is so well insulated from natural selection, that the selection pressures become smoothed out at the molecular level over time spans of millions of years. A factory can convert

RATES OF EVOLUTION OF PROTEINS (*opposite page*) can be inferred by plotting average differences in amino acid sequences between species on two sides of an evolutionary branch point that can be dated, for example the branch point between fish and reptiles or between reptiles and mammals. The average differences (*vertical axis*) have been corrected to allow for the occurrence of more than one mutation at a given amino acid site. The length of the vertical data bars indicates the experimental scatter. Times since the divergence of two lines of organisms from a common branch point (*horizontal axis*) have been obtained from the geological record. The drawings below the graph show schematically the function (described in the text) of the molecules whose evolutionary rate of change is plotted. The rate of change is proportional to the steepness of the curve. It can be represented by a number called the unit evolutionary period, which is the time required for the amino acid sequence of a protein to change by 1 percent after two evolutionary lines have diverged. For fibrinopeptides this period is about 1.1 million years, whereas for histone IV it is 600 million years. The probable reasons for these differences are discussed in the text.

from making military tanks to making sports cars and keep the same machine tools and power source. Similarly, a primitive eukaryote cell line can lead to such diverse organisms as sunflowers and mammals and still retain a common metabolic chemistry, including the respiratory package that comprises cytochrome *c*. One of the advantages of proteins in studying the process of evolution is just this relative insulation from the immediate effects of external selection. Protein structure is farther removed from selection pressures and closer to the sources of genetic variation in DNA than gross anatomical features or inherited behavior patterns are.

The only other proteins for which enough sequence information is available to allow this kind of analysis are hemoglobin and the fibrinopeptides: the short amino acid chains left over when fibrinogen is converted to fibrin in the process of blood clotting. One hemoglobin chain consists of approximately 140 amino acids. Fibrinopeptides *A* and *B*, on the other hand, consist of only about 20 amino acids, which are cut out of fibrinogen and discarded during the clotting process. The hemoglobins and the fibrinopeptides also appear to be evolving individually at a uniform average rate, but their rates are quite different. Whereas 20 million years are required to produce a change of 1 percent in the amino acid sequence of two diverging lines of cytochrome *c*, the same amount of change takes a little less than six million years in hemoglobins and just over one million years in the fibrinopeptides, as indicated on page 68. The approximate time required for a 1 percent change in sequence to appear between diverging lines of the same protein is defined as the unit evolutionary period. That period has been roughly estimated for a number of proteins for which only two or three sequences from different species are known. Most simple enzymes evolve approximately as fast as hemoglobin and much more rapidly than cytochrome *c*. Does all of this mean that the genes for these proteins are mutating at different rates? Are we looking at differences in variation or in selection?

Since there is no evidence to suggest variable rates of mutation, one asks what case can be made for differences in selection pressure among the different proteins. The case appears to be quite convincing [*see inset figures in illustration on page 68*]. The fibrinopeptides are "spacers" that prevent fibrinogen from adopting the fibrin configuration before the clotting mechanism is triggered. As long as they can be cut out by an enzyme when the time comes for the blood to clot, they would seem to have few other requirements. Thus one would expect a fibrinogen molecule to tolerate many random changes in the fibrinopeptide spacers. If the unit evolutionary period measures not the rate of appearance of mutations but the rate of appearance of *harmless* mutations, then it is not surprising that a 1 percent change can occur in the sequence of fibrinopeptides in just over a million years.

A successful hemoglobin molecule has more constraints. Each hemoglobin molecule embodies four heme groups that not only bind oxygen but also cooperate in such a way that the oxygen is released more rapidly into the cell when the local acidity, created by the presence of carbon dioxide, builds up. The structural basis for this "breathing" mechanism has only recently been explained with the help of X-ray crystallography by M. F. Perutz and his co-workers at the Medical Research Council Laboratory of Molecular Biology in England. If a random mutation is five times as likely to be harmful in hemoglobin as in the fibrinopeptides, one can account for hemoglobin's having a unit evolutionary period that is five times as long.

The chances of randomly damaging cytochrome *c* are evidently three to four times greater than they are for hemoglobin. Why should this be, and why should the unit evolutionary period for cytochrome *c* be greater than the period for enzymes of comparable size? The X-ray structure has given us a clue to the answer. Cytochrome *c* is a small protein that interacts over a large portion of its surface with molecular complexes that are larger than itself. It is virtually a "substrate" for the reductase and oxidase complexes. A large fraction of its surface is subject to strong conservative selection pressures because of the requirement that it mate properly with other large molecules, each with its own genetic blueprint. This evidently explains why the patches of positive and negative charge are preserved so faithfully throughout the history of eukaryotic life. Hemoglobin and most enzymes, in contrast, interact principally with smaller molecules: with oxygen in the case of hemoglobin or with small substrate molecules at the active sites of enzymes. As long as these restricted regions of the molecule are preserved the rest of the molecular surface is relatively free to change. Mutations are weeded out less rigorously and sequences diverge faster.

A satisfying confirmation of these

LONG TRIPS IN LITTLE CARS NEEDN'T BE ENDURANCE TESTS.

One thing you don't expect in a little car is comfort. With Vega, you're in for a surprise.

We didn't work any miracles, mind you. No little car will ever give you the comfort of a limousine.

But a good little car should get you from here to there without your feeling every mile in between.

And Vega does.

For one thing, it's built to be little outside but big inside. That may sound contradictory, but it means we can provide as much room per passenger as many big cars.

Another thing. Vega has a couple of the most comfortable front buckets you're liable to find anywhere. Seats you sit in, not on.

There are other comforts, too. The engine, for instance. It's big enough to do its work without straining, even at prolonged highway speeds. If your engine feels easy going, you'll feel more at ease too.

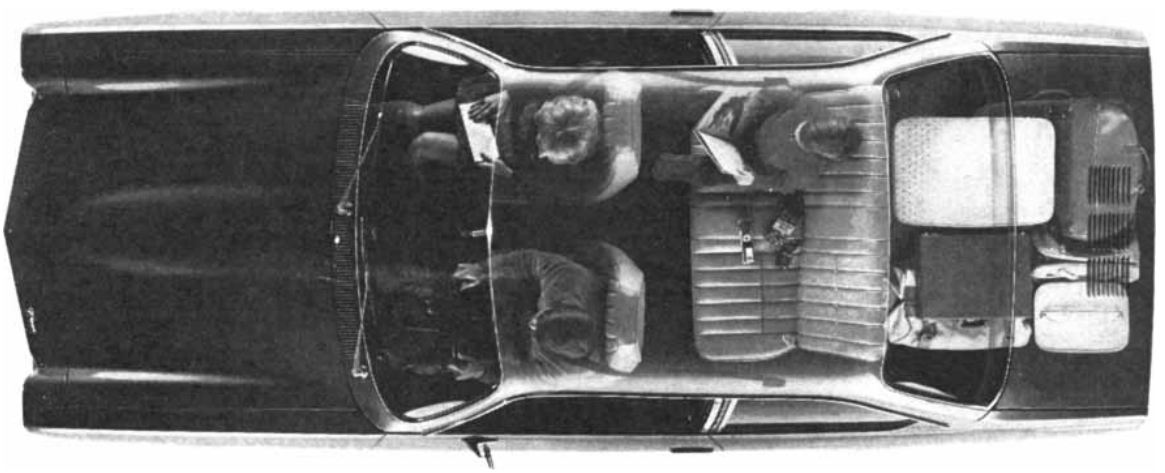
The Vega power ventilation system keeps air circulating inside the car as you travel. Or if you pull over just to admire the view.

Full Coil suspension front and rear softens the big bumps and makes the little ones all but disappear. Coupled with Vega's superb handling and stability, those long trips might even stand a few detours. All because we want your Vega to be the best car you ever owned.

GM
MARK OF EXCELLENCE

VEGA
CHEVROLET

Highway safety begins at home. Buckle up before you leave.



Chevrolet. Building a better way to see the U.S.A.

ideas comes from the amino acid sequences of histone IV, one of the basic proteins that binds to DNA in the chromosomes and that may play a role in expressing or suppressing genetic information. When molecules of histone IV from pea seedlings and calf thymus are compared, one finds that they differ in only two of their 102 amino acids. If we adopt an approximate date of 1.2 billion years ago from the cytochrome study for the divergence of plants and animals, we find that histone IV has a unit evolutionary period of 600 million years. Clearly the conservative selection pressure on histone IV must be intense. Since histone IV participates in the control processes that are at the heart of the genetic mechanism, its sensitivity to random changes is hardly surprising.

The date of 1.2 billion years ago for the divergence of plants and animals is based on the cytochrome-sequence comparisons, assuming that the observed linear rate of evolution of cytochrome *c* in more recent times can be extrapolated back to that remote epoch. Is this a fair extrapolation? It probably is for cytochrome *c* because the biochemistry of the mitochondrion evolved still earlier; the great similarity in respiratory reactions among all eukaryotes argues that there has not been much innovation in cytochrome systems since. The respiratory chain had probably "settled down" by 1.2 billion years ago. It is reassuring that the cytochrome figure of 1.2 billion years is in harmony with the relatively scarce fossil record of Precambrian life.

If one accepts the provocative suggestion that eukaryotes developed from a symbiotic association of several prokaryotes, one of which was a respiring bacterium that became the ancestor of present-day mitochondria, one is obliged to conclude that the respiratory machinery had stabilized in essentially its present form before or during this symbiosis. The same thing cannot be said for hemoglobin and its probable ancestor, myoglobin. They can provide no clue to the date when animals and plants diverged, since the globins were evolving to play several different roles during and after this period, as multicelled organisms arose. In no sense had the globins settled down 1.2 million years ago. Nevertheless, if the right proteins are selected, and if the data are not overextended, it should be possible to use the rates of protein evolution to assign times to events in the evolution of life that have left only faint traces in the geological record.

So far we have mentioned the electron-transfer mechanism of cytochrome *c* only in passing, virtually ignoring the structure of the reduced molecule. The mechanism is another story in itself, and one that cannot yet be written. One hopes that the clues supplied by X-ray analysis will suggest the best chemical experiments to try next, in order to learn the mechanism of the oxidation-reduction process. The reduced cytochrome structure has been obtained so recently that it would be premature to base many deductions on it.

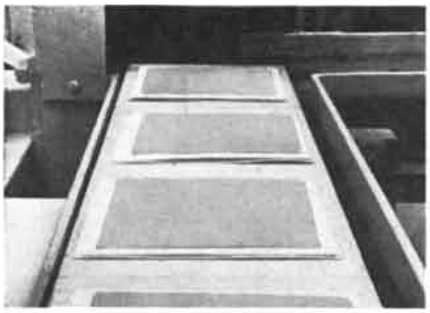
A Glimpse of Molecular Dynamics

One obvious structural feature, which undoubtedly has great physiological significance, is that in the reduced molecule the top of the heme crevice is closed. The chain from residues 80 to 83 swings to the right (as the molecule is depicted on page 62), the exposed phenylalanine 82 slips into the heme crevice to the left of the heme and nearly parallel to it, and the heme becomes less accessible to the outside world. The absolute preservation of this phenylalanine side chain throughout evolution, in an environment that is energetically unfavorable in the oxidized molecule, argues that closing of the heme crevice in the reduced molecule is important for its biological activity.

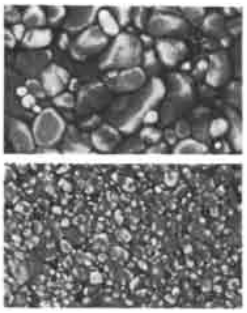
Several explanations might be offered. The aromatic ring of phenylalanine 82 may be part of the electron-transfer mechanism, or its removal from the heme crevice may be necessary to permit an electron-transferring group to enter beside the heme or just to approach the edge of the heme. At a minimum the refolding of the chain from residues 80 to 83 may be a "convulsive" motion that pushes the oxidase complex away from the protein after electron transfer is achieved by some other pathway.

This article has been speculative enough without making a choice between these or other alternatives. At this stage, as both oxidized and reduced cytochrome analyses are being extended to higher resolution, it is enough to say that we can see more refolding of the protein chain in passing between the two states than has been observed in any other protein. Phenylalanine 82 swings to an entirely new position and several other aromatic rings change orientation, including the three in the left channel. As the molecule is reduced, the right channel apparently is partly blocked by residues 20 and 21. We now have pictures of both strokes of a very ancient two-stroke molecular engine. We hope in time to be able to figure out how it operates.

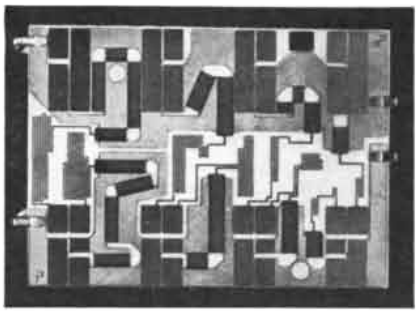
WESTERN ELECTRIC REPORTS



1500° C furnace was specially designed to fire these new substrates. The relatively low temperature results in smooth substrate surfaces for practically fault-free thin film bonding.



Electron micrographs show the great difference in grain size between new ceramic material (lower) and the previous material (upper).



Thin film integrated circuit shown here is part of a resistor network. It is one of many that benefit from the improved substrate. Metal leads on sides are bonded by thermocompression to tantalum nitride resistor film.

Smoothing the way for perfect thin film bonding.

Aluminum oxide, or alumina, is considered to have the best combination of properties for thin film circuit substrates. Until recently, however, the bonding of metal elements to gold-coated tantalum nitride resistor film on alumina was somewhat unpredictable.

Now, an advance at Western Electric has made it possible to get practically fault-free bonding of these materials.

This new perfection in bonding came through the development of finer grained alumina substrates.

The process has four basic steps: milling, casting, punching and firing.

During milling, alumina is combined with magnesium oxide, trichlorethylene, ethanol and a unique deflocculant. For 24 hours, this mixture is rotated in a ball mill. In a second 24-hour period, plasticizers and a binder are included.

The deflocculant plays a major role by dissipating the attraction forces that exist between the highly active alumina particles. This prevents thickening, which would ordinarily make an active alumina mixture unworkable.

The 48 hours of milling is followed by casting. When the material comes off the casting line, it is in the form of a flexible polymer/alumina tape, dry enough to be cut into easily handled sections.

After casting, a punch press cuts the material into the desired rectangles or other shapes. Holes can be punched at the same time.

Finally, because of the use of active alumina, the material is fired at an unusually low temperature which results in smooth substrate surfaces for reliable thin film bonding. The finished substrate is then ready for the various processes of thin film circuit production.

In developing this new process, engineers at Western Electric's Engineering Research Center worked together with engineers at the Allentown plant.

Conclusion: This new way to produce substrates is a truly significant contribution for thin film circuit production.

The ultimate gain from this smoother substrate is for communications itself. For through the achievement of nearly perfect bonding of metal leads to tantalum nitride, thin films can be produced with even greater reliability and economy.

Western Electric

We make things that bring people closer.

Some things are changing for the better.

Many people know us as an instrument manufacturer: we make more than 2,000 products for measurement, test and analysis. Others know us as a computer company: more than 10,000 own our programmable calculators and computers. We prefer to think that our business is to serve measurement, analysis and computation needs . . . in science, industry, medicine and education. This is the rationale behind every new instrument, computer or system that we tell you about in these ads. This month:

When the HP 9600 rolls through your door, your real-time and data acquisition tasks become a lot easier to perform. This new systems family's long suit is the efficient and economic handling of multitudes of analog and digital information, simultaneously.

A sensor-based system that makes real sense.

There's a growing demand in industry and research laboratories for sensor-based computer systems that handle great quantities of analog and digital information. Systems built from programmable instruments usually are too expensive; people pay for equipment features that they don't need. Yet the alternative has been a piecemeal approach – break down the customer's problem into several parts and use separate "mini-systems" to solve each part independently.

Now there's a third choice – Hewlett-Packard's new family of compact data acquisition and control systems for cost-effective automation in industry and research. A 9600 Series system monitors, collects, and processes information from sensor-based sources. It then can generate reports, control power supplies, alert operators, drive graphic displays and plotters, and produce control signals for closed loop operations. Although you can't be everywhere at once – supervising and trouble-shooting – our system can.

Two new subsystems within the 9600, one analog and one digital, now do the things a number of programmable instruments used to do. These instrument functions are contained on plug-in cards. Instead of adding individual instruments, you merely slip in an inexpensive printed cirucuit board.

The 9600 data acquisition systems are modular. Start with a minimum low-cost system to control a single test or experiment, and expand with your growing needs.

The full story on the 9600 System family is yours for the asking.

Nothing can outperform this new digital GC-even at twice the price.

Because the gas chromatograph (GC) is essentially a tool for qualitative and quantitative chemical analysis, its value ultimately depends on how well it does this job. Over the years, many new models have been introduced that perform more accurately than previous

This would be an unusual case – using a battery-powered counter to check out the frequency of a mountain rescue-team's radio equipment – but it illustrates that HP's portable instruments can go anywhere service is needed.

instruments – at a price. The truly amazing thing about the new HP 5700 GC is this: it produces more accurate and precise retention time (qualitative) and peak area (quantitative) data than any GC ever built. Yet it costs about half as much as top-of-the-line GCs of comparable quality.

A new bulletin on the 5700 fully documents this perhaps startling claim. Until you have a chance to study this data consider this: one of the first 5700s off the production line was used "as is" to make two series of replicate analytical runs, one series before and one after an overnight shutdown. The sample used in both series contained seven components, out to C_{17}.

The results speak for themselves. In terms of repeat accuracy, the mean retention time of each of the seven components differed less than 0.01 minute after the overnight shutdown; the normalized area % varied only within ±0.001%. In terms of precision, the standard deviations of the replicate retention time measurements fell within 0.0175, both before and after the overnight shutdown; the standard deviations of the area % data were all within 0.0038. No other GC, regardless of price, can do better.

For a fully documented proof of performance as well as a factual description of this new all-digital, computer-compatible automatic GC, write for Bulletin 5700.

Portable instruments go where the problem is.

Capital equipment such as mobile or remote communications systems and million dollar computers have at least two things in common. They are electronically complex, and they can't be taken into a service center when they need repair. Today's traveling field service engineer must have laboratory quality equipment that will go where he goes.

HP's portable instruments enable service engineers to diagnose and repair this equipment on the spot, reducing expensive downtime. Our portable scopes are small enough to fit under an airliner seat, and, at 24 pounds, are light enough to be carried up antenna masts and into other hard-to-reach places. An HP electronic counter can be held in one hand – it takes only seconds to snap on a function module that provides the specific measuring capability needed. Then there's our multi-function meter – a high performance, instant-reading voltmeter and ohmmeter rolled into one.

And the length of HP's portable measuring capability isn't limited by the distance to the nearest wall socket. Most of our portable instruments feature their own accessory battery pack. Many can run off ordinary car, plane or boat batteries as well as a standard power line. And all of them deliver HP precision in a rugged, portable package.

Ask for the full story on portable instruments that go where the problem is. Write Hewlett-Packard, 1503 Page Mill Road, Palo Alto, California 94304; Europe: 1217 Meyrin-Geneva, Switzerland.

00202

HEWLETT hp PACKARD

LANGUAGE AND THE BRAIN

Aphasias are speech disorders caused by brain damage. The relations between these disorders and specific kinds of brain damage suggest a model of how the language areas of the human brain are organized

by Norman Geschwind

Virtually everything we know of how the functions of language are organized in the human brain has been learned from abnormal conditions or under abnormal circumstances: brain damage, brain surgery, electrical stimulation of brains exposed during surgery and the effects of drugs on the brain. Of these the most fruitful has been the study of language disorders, followed by postmortem analysis of the brain, in patients who have suffered brain damage. From these studies has emerged a model of how the language areas of the brain are interconnected and what each area does.

A disturbance of language resulting from damage to the brain is called aphasia. Such disorders are not rare. Aphasia is a common aftereffect of the obstruction or rupture of blood vessels in the brain, which is the third leading cause of death in the U.S. Although loss of speech from damage to the brain had been described occasionally before the 19th century, the medical study of such cases was begun by a remarkable Frenchman, Paul Broca, who in 1861 published the first of a series of papers on language and the brain. Broca was the first to point out that damage to a specific portion of the brain results in disturbance of language output. The portion he identified, lying in the third frontal gyrus of the cerebral cortex, is now called Broca's area [*see illustration on page 78*].

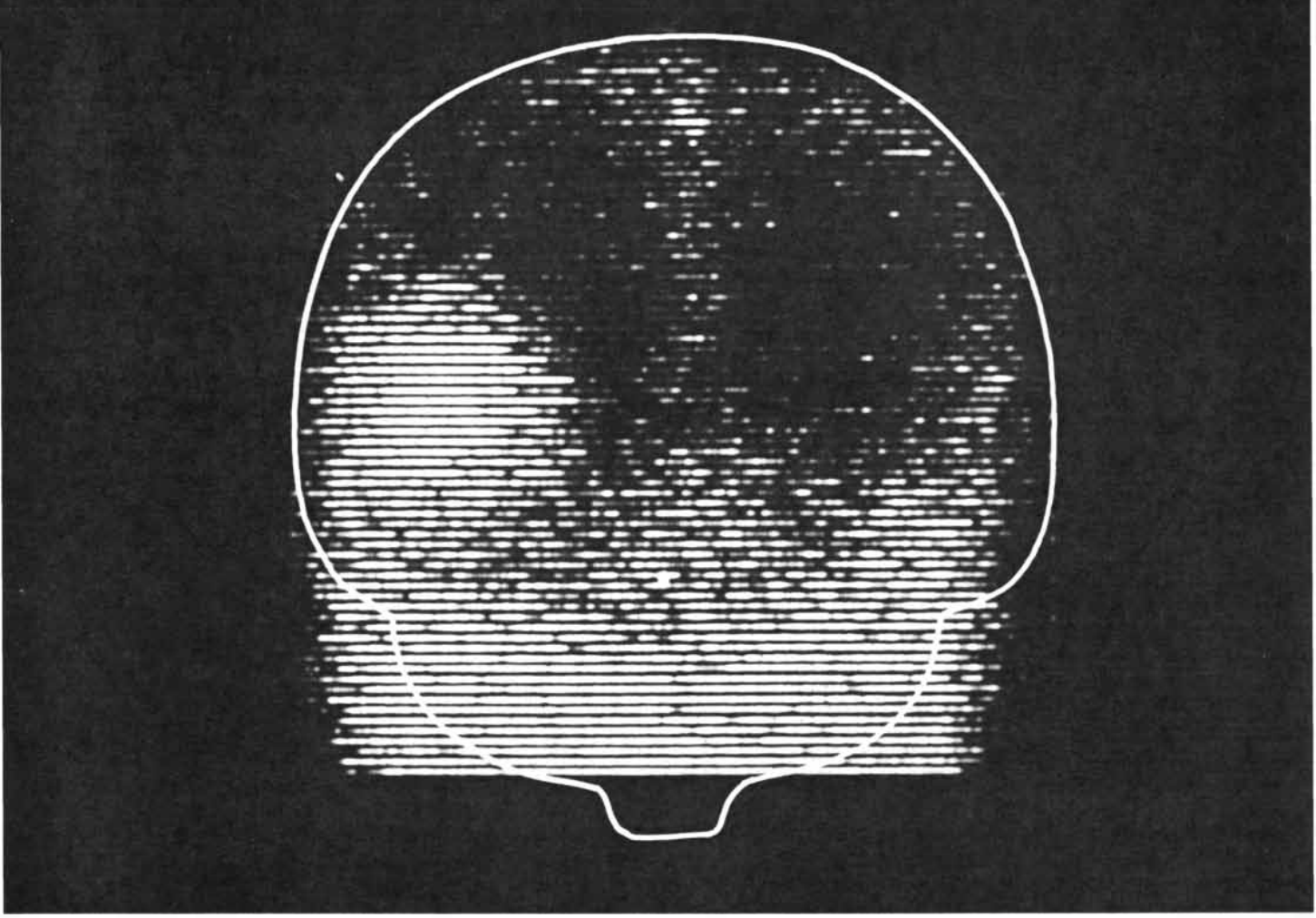
Broca's area lies immediately in front of the portion of the motor cortex that controls the muscles of the face, the jaw, the tongue, the palate and the larynx, in other words, the muscles involved in speech production. The region is often called the "motor face area." It might therefore seem that loss of speech from damage to Broca's area is the result of paralysis of these muscles. This explanation, however, is not the correct one. Direct damage to the area that controls these muscles often produces only mild weakness of the lower facial muscles on the side opposite the damage and no permanent weakness of the jaw, the tongue, the palate or the vocal cords. The reason is that most of these muscles can be controlled by either side of the brain. Damage to the motor face area on one side of the brain can be compensated by the control center on the opposite side. Broca named the lesion-produced language disorder "aphemia," but this term was soon replaced by "aphasia," which was suggested by Armand Trousseau.

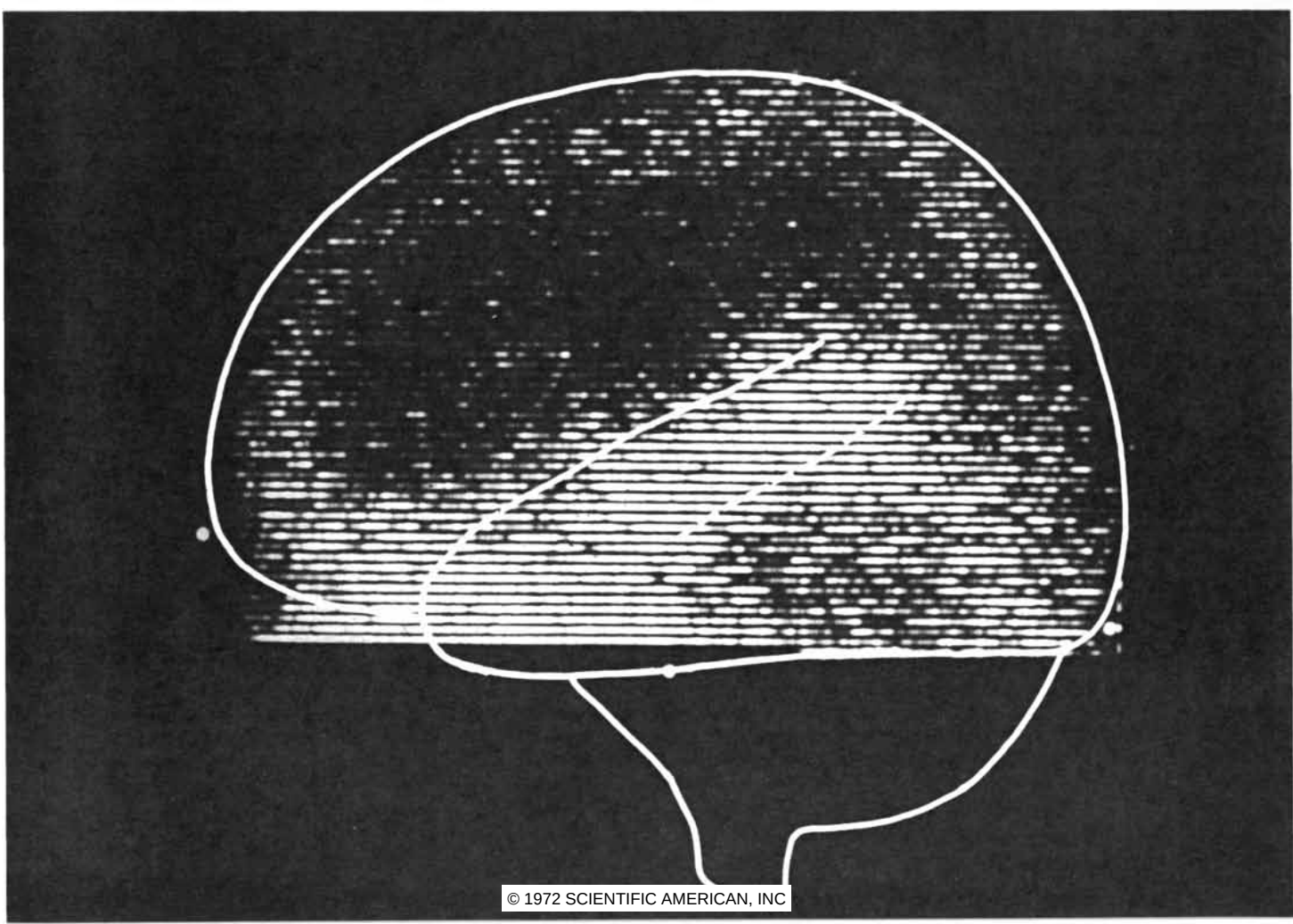
In 1865 Broca made a second major contribution to the study of language and the brain. He reported that damage to specific areas of the left half of the brain led to disorder of spoken language but that destruction of corresponding areas in the right side of the brain left language abilities intact. Broca based his conclusion on eight consecutive cases of aphasia, and in the century since his report his observation has been amply confirmed. Only rarely does damage to the right hemisphere of the brain lead to language disorder; out of 100 people with permanent language disorder caused by brain lesions approximately 97 will have damage on the left side. This unilateral control of certain functions is called cerebral dominance. As far as we know man is the only mammal in which learned behavior is controlled by one half of the brain. Fernando Nottebohm of Rockefeller University has found unilateral neural control of birdsong. It is an interesting fact that a person with aphasia of the Broca type who can utter at most only one or two slurred words may be able to sing a melody rapidly, correctly and even with elegance. This is another proof that aphasia is not the result of muscle paralysis.

In the decade following Broca's first report on brain lesions and language there was a profusion of papers on aphasias of the Broca type. In fact, there was a tendency to believe all aphasias were the result of damage to Broca's area. At this point another great pioneer of the brain appeared on the scene. Unlike Broca, who already had a reputation at the time of his first paper on aphasia, Carl Wernicke was an unknown with no previous publications; he was only 26 years old and a junior assistant in the neurological service in Breslau. In spite of his youth and obscurity his paper on aphasia, published in 1874, gained immediate attention. Wernicke described damage at a site in the left hemisphere outside Broca's area that results in a language disorder differing from Broca's aphasia.

In Broca's aphasia speech is slow and labored. Articulation is crude. Characteristically, small grammatical words and the endings of nouns and verbs are

LOCATION OF SOME LESIONS in the brain can be determined by injecting into the bloodstream a radioactive isotope of mercury, which is taken up by damaged brain tissue. The damaged region is identified by scanning the head for areas of high radioactivity. The top scan on the opposite page was made from the back of the head; the white area on the left shows that the damage is in the left hemisphere. The bottom scan is of the left side of the head and shows that the uptake of mercury was predominantly in the first temporal gyrus, indicating damage to Wernicke's speech area by occlusion of blood vessels. David Patten and Martin Albert of the Boston Veterans Administration Hospital supplied the scans.





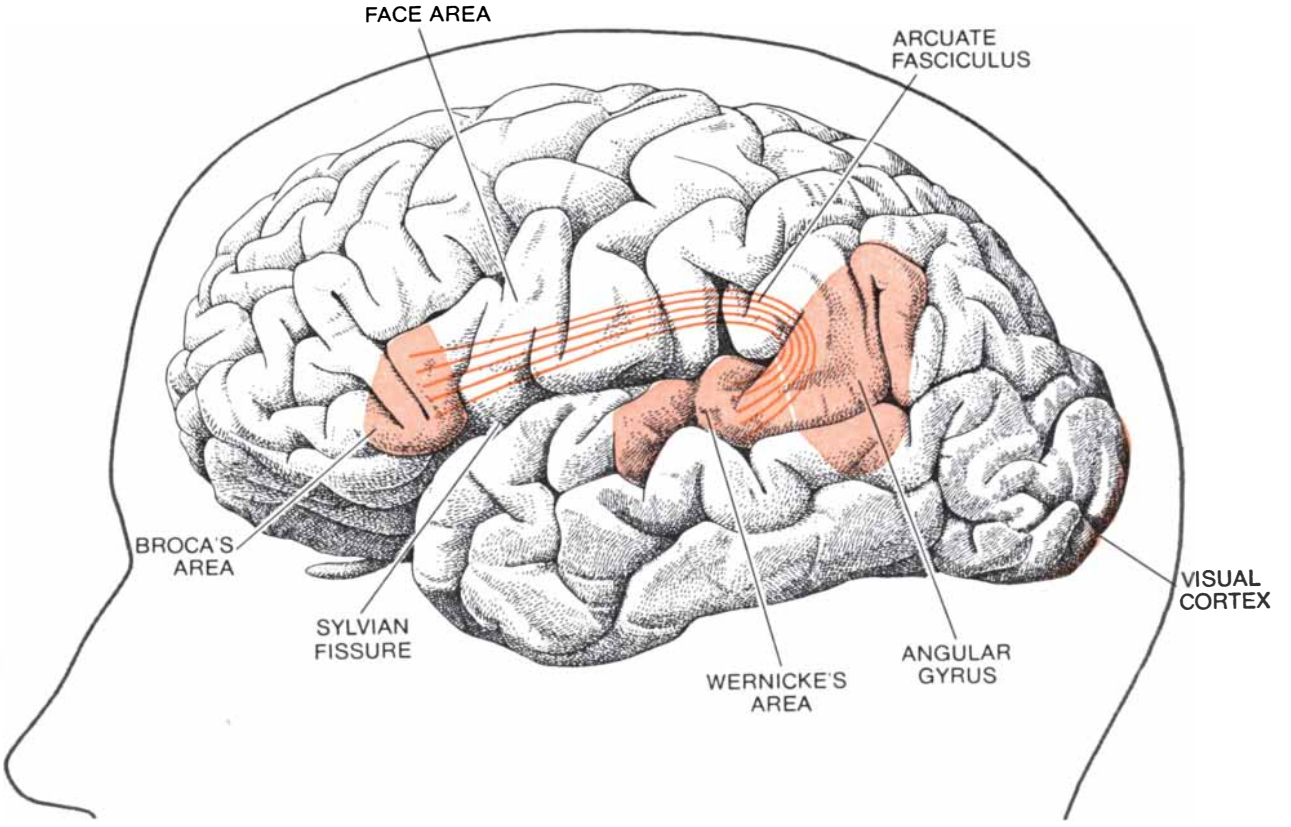
omitted, so that the speech has a telegraphic style. Asked to describe a trip he has taken, the patient may say "New York." When urged to produce a sentence, he may do no better than "Go... New York." This difficulty is not simply a desire to economize effort, as some have suggested. Even when the patient does his best to cooperate in repeating words, he has difficulty with certain grammatical words and phrases. "If he were here, I would go" is more difficult than "The general commands the army." The hardest phrase for such patients to repeat is "No ifs, ands or buts."

The aphasia described by Wernicke is quite different. The patient may speak very rapidly, preserving rhythm, grammar and articulation. The speech, if not listened to closely, may almost sound normal. For example, the patient may say: "Before I was in the one here, I was over in the other one. My sister had the department in the other one." It is abnormal in that it is remarkably devoid of content. The patient fails to use the correct word and substitutes for it by circumlocutory phrases ("what you use to cut with" for "knife") and empty words ("thing"). He also suffers from paraphasia, which is of two kinds. Verbal paraphasia is the substitution of one word or phrase for another, sometimes related in meaning ("knife" for "fork") and sometimes unrelated ("hammer" for "paper"). Literal or phonemic paraphasia is the substitution of incorrect sounds in otherwise correct words ("kench" for "wrench"). If there are several incorrect sounds in a word, it becomes a neologism, for example "pluver" or "flieber."

Wernicke also noted another difference between these aphasic patients and those with Broca's aphasia. A person with Broca's aphasia may have an essentially normal comprehension of language. Indeed, Broca had argued that no single lesion in the brain could cause a loss of comprehension. He was wrong. A lesion in Wernicke's area can produce a severe loss of understanding, even though hearing of nonverbal sounds and music may be fully normal.

Perhaps the most important contribution made by Wernicke was his model of how the language areas in the brain are connected. Wernicke modestly stated that his ideas were based on the teachings of Theodor Meynert, a Viennese neuroanatomist who had attempted to correlate the nervous system's structure with its function. Since Broca's area was adjacent to the cortical region of the brain that controlled the muscles of speech, it was reasonable to assume, Wernicke argued, that Broca's area incorporated the programs for complex coordination of these muscles. In addition Wernicke's area lay adjacent to the cortical region that received auditory stimuli [*see illustration below*]. Wernicke made the natural assumption that Broca's area and Wernicke's area must be connected. We now know that the two areas are indeed connected, by a bundle of nerve fibers known as the arcuate fasciculus. One can hypothesize that in the repetition of a heard word the auditory patterns are relayed from Wernicke's area to Broca's area.

Comprehension of written language



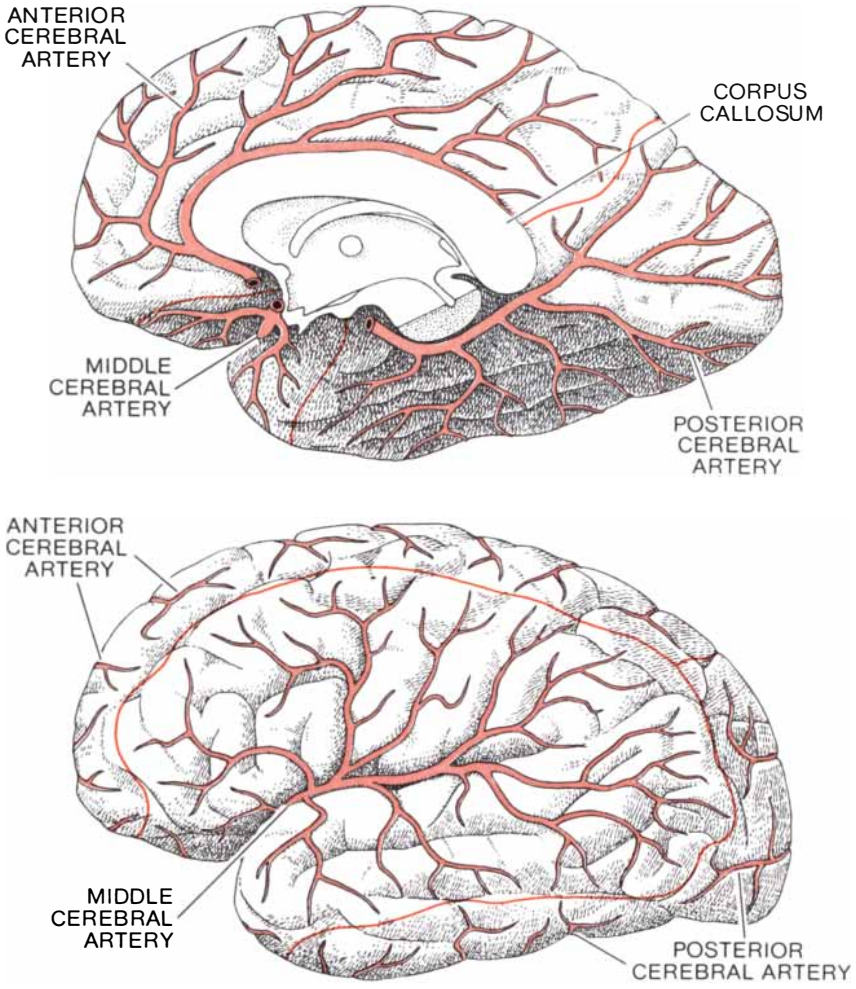
PRIMARY LANGUAGE AREAS of the human brain are thought to be located in the left hemisphere, because only rarely does damage to the right hemisphere cause language disorders. Broca's area, which is adjacent to the region of the motor cortex that controls the movement of the muscles of the lips, the jaw, the tongue, the soft palate and the vocal cords, apparently incorporates programs for the coordination of these muscles in speech. Damage to Broca's area results in slow and labored speech, but comprehension of language remains intact. Wernicke's area lies between Heschl's gyrus, which is the primary receiver of auditory stimuli, and the angular gyrus, which acts as a way station between the auditory and the visual regions. When Wernicke's area is damaged, speech is fluent but has little content and comprehension is usually lost. Wernicke and Broca areas are joined by a nerve bundle called the arcuate fasciculus. When it is damaged, speech is fluent but abnormal, and patient can comprehend words but cannot repeat them.

would require connections from the visual regions to the speech regions. This function is served by the angular gyrus, a cortical region just behind Wernicke's area. It acts in some way to convert a visual stimulus into the appropriate auditory form.

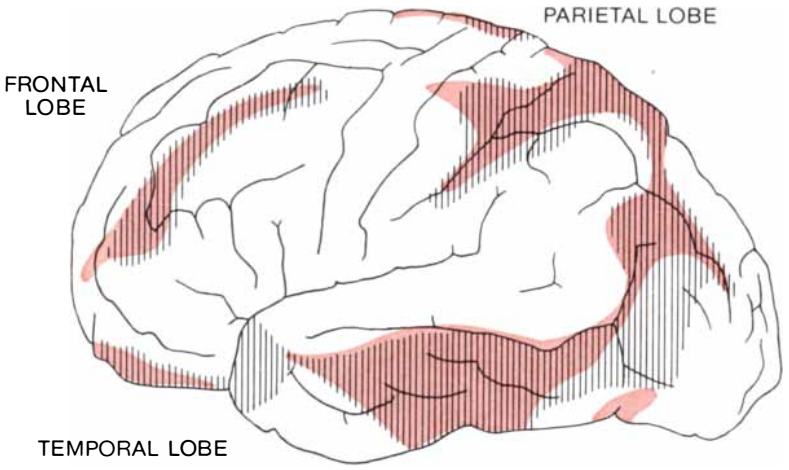
We can now deduce from the model what happens in the brain during the production of language. When a word is heard, the output from the primary auditory area of the cortex is received by Wernicke's area. If the word is to be spoken, the pattern is transmitted from Wernicke's area to Broca's area, where the articulatory form is aroused and passed on to the motor area that controls the movement of the muscles of speech. If the spoken word is to be spelled, the auditory pattern is passed to the angular gyrus, where it elicits the visual pattern. When a word is read, the output from the primary visual areas passes to the angular gyrus, which in turn arouses the corresponding auditory form of the word in Wernicke's area. It should be noted that in most people comprehension of a written word involves arousal of the auditory form in Wernicke's area. Wernicke argued that this was the result of the way most people learn written language. He thought, however, that in people who were born deaf, but had learned to read, Wernicke's area would not be in the circuit.

According to this model, if Wernicke's area is damaged, the person would have difficulty comprehending both spoken and written language. He should be unable to speak, repeat and write correctly. The fact that in such cases speech is fluent and well articulated suggests that Broca's area is intact but receiving inadequate information. If the damage were in Broca's area, the effect of the lesion would be to disrupt articulation. Speech would be slow and labored but comprehension should remain intact.

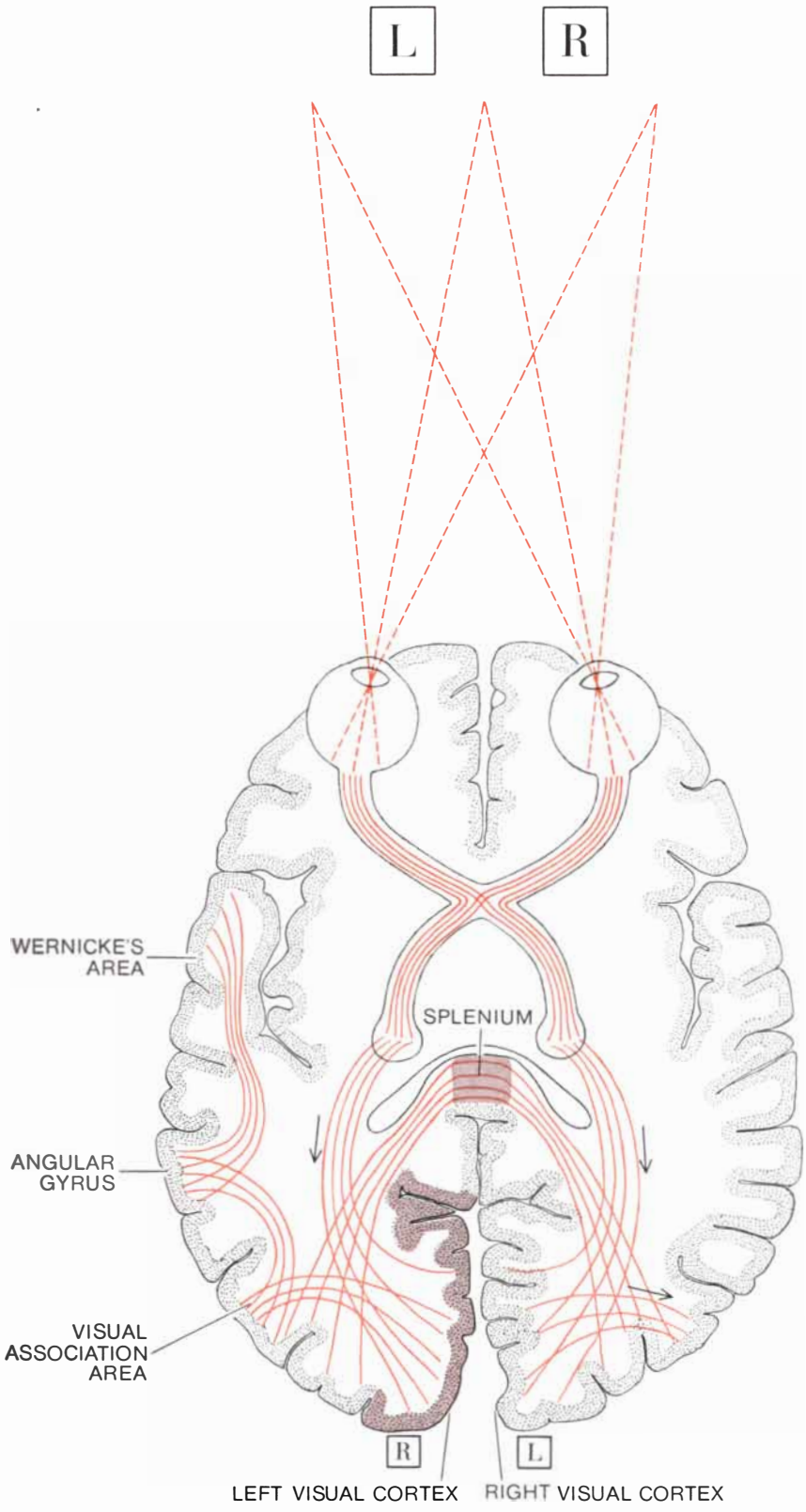
This model may appear to be rather simple, but it has shown itself to be remarkably fruitful. It is possible to use it to predict the sites of brain lesions on the basis of the type of language disorder. Moreover, it gave rise to some definite predictions that lesions in certain sites should produce types of aphasia not previously described. For example, if a lesion disconnected Wernicke's area from Broca's area while leaving the two areas intact, a special type of aphasia should be the result. Since Broca's area is preserved, speech should be fluent but abnormal. On the other hand, comprehension should be intact because Wernicke's area is still functioning. Rep-



CEREBRAL AREAS are nourished by several arteries, each supplying blood to a specific region. The speech and auditory region is nourished by the middle cerebral artery. The visual areas at the rear are supplied by the posterior cerebral artery. In patients who suffer from inadequate oxygen supply to the brain the damage is often not within the area of a single blood vessel but rather in the "border zones" (*colored lines*). These are the regions between the areas served by the major arteries where the blood supply is marginal.



ISOLATION OF SPEECH AREA by a large *C*-shaped lesion produced a remarkable syndrome in a woman who suffered from severe carbon monoxide poisoning. She could repeat words and learn new songs but could not comprehend the meaning of words. Postmortem examination of her brain revealed that in the regions surrounding the speech areas of the left hemisphere, either the cortex (*colored areas*) or the underlying white matter (*hatched areas*) was destroyed but that the cortical structures related to the production of language (Broca's area and Wernicke's area) and the connections between them were left intact.



CLASSIC CASE of a man who lost the ability to read even though he had normal visual acuity and could copy written words was described in 1892 by Joseph Jules Dejerine. Postmortem analysis of the man's brain showed that the left visual cortex and the splenium (*dark colored areas*) were destroyed as a result of an occlusion of the left posterior cerebral artery. The splenium is the section of the corpus callosum that transfers visual information between the two hemispheres. The man's left visual cortex was inoperative, making him blind in his right visual field. Words in his left visual field were properly received by the right visual cortex, but could not cross over to the language areas in the left hemisphere because of the damaged splenium. Thus words seen by the man remained as meaningless patterns.

etition of spoken language, however, should be grossly impaired. This syndrome has in fact been found. It is termed conduction aphasia.

The basic pattern of speech localization in the brain has been supported by the work of many investigators. A. R. Luria of the U.S.S.R. studied a large number of patients who suffered brain wounds during World War II [see "The Functional Organization of the Brain," by A. R. Luria; SCIENTIFIC AMERICAN, March, 1970]. When the wound site lay over Wernicke's or Broca's area, Luria found that the result was almost always severe and permanent aphasia. When the wounds were in other areas, aphasia was less frequent and less severe.

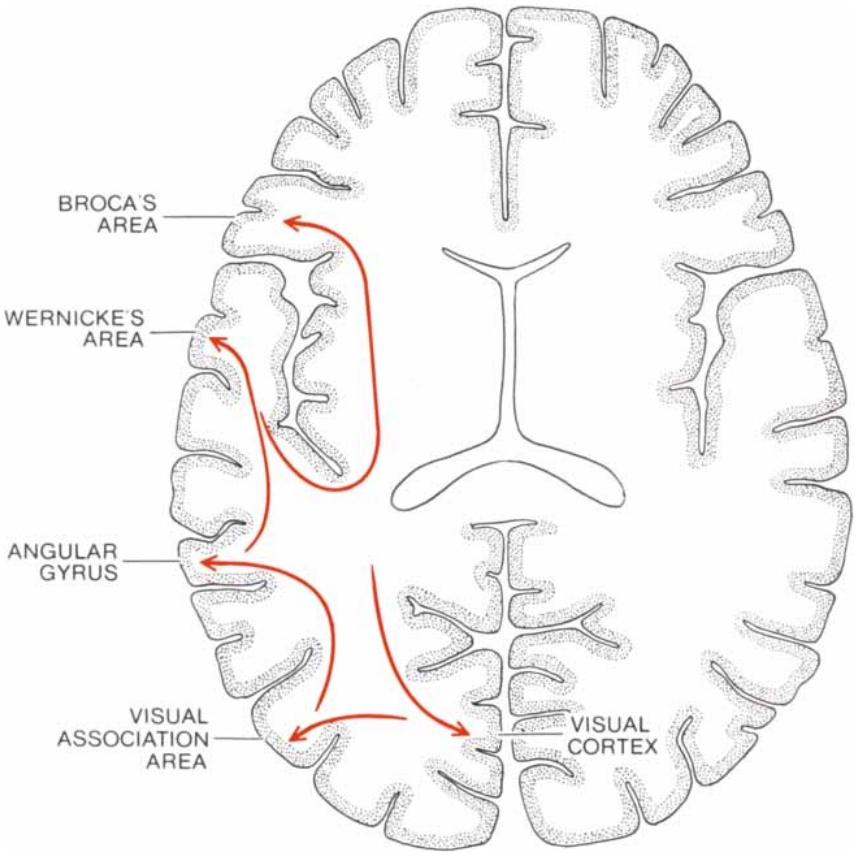
A remarkable case of aphasia has provided striking confirmation of Wernicke's model. The case, described by Fred Quadfasel, Jose Segarra and myself, involved a woman who had suffered from accidental carbon monoxide poisoning. During the nine years we studied her she was totally helpless and required complete nursing care. She never uttered speech spontaneously and showed no evidence of comprehending words. She could, however, repeat perfectly sentences that had just been said to her. In addition she would complete certain phrases. For example, if she heard "Roses are red," she would say "Roses are red, violets are blue, sugar is sweet and so are you." Even more surprising was her ability to learn songs. A song that had been written after her illness would be played to her and after a few repetitions she would begin to sing along with it. Eventually she would begin to sing as soon as the song started. If the song was stopped after a few bars, she would continue singing the song through to the end, making no errors in either words or melody.

On the basis of Wernicke's model we predicted that the lesions caused by the carbon monoxide poisoning lay outside the speech and auditory regions, and that both Broca's area and Wernicke's area were intact. Postmortem examination revealed a remarkable lesion that isolated the speech area from the rest of the cortex. The lesion fitted the prediction. Broca's area, Wernicke's area and the connection between them were intact. Also intact were the auditory pathways and the motor pathways to the speech organs. Around the speech area, however, either the cortex or the underlying white matter was destroyed [*see bottom illustration on preceding page*]. The woman could not comprehend speech because the words did not arouse

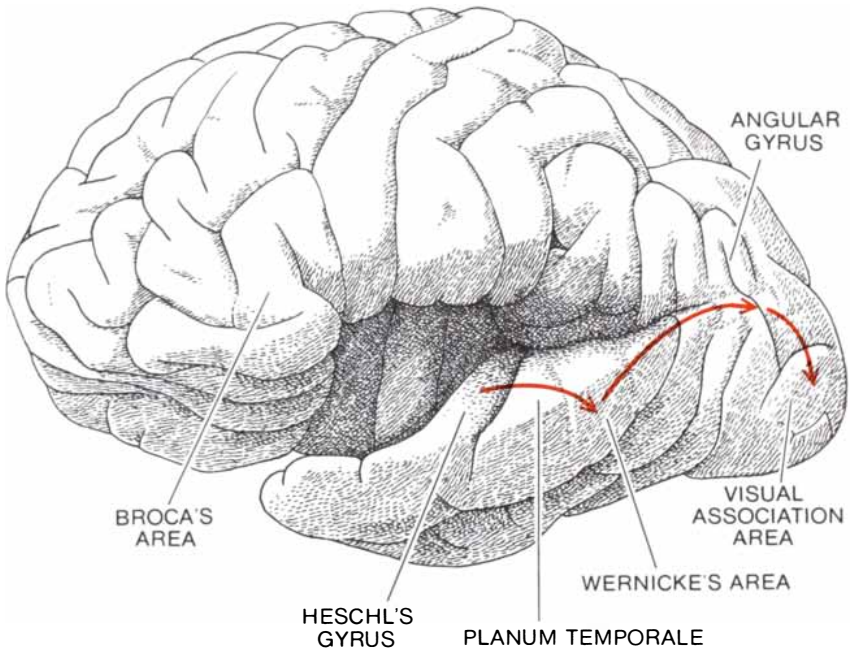
associations in other portions of the cortex. She could repeat speech correctly because the internal connections of the speech region were intact. Presumably well-learned word sequences stored in Broca's area could be triggered by the beginning phrases. This syndrome is called isolation of the speech area.

Two important extensions of the Wernicke model were advanced by a French neurologist, Joseph Jules Dejerine. In 1891 he described a disorder called alexia with agraphia: the loss of the ability to read and write. The patient could, however, speak and understand spoken language. Postmortem examination showed that there was a lesion in the angular gyrus of the left hemisphere, the area of the brain that acts as a way station between the visual and the auditory region. A lesion here would separate the visual and auditory language areas. Although words and letters would be seen correctly, they would be meaningless visual patterns, since the visual pattern must first be converted to the auditory form before the word can be comprehended. Conversely, the auditory pattern for a word must be transformed into the visual pattern before the word can be spelled. Patients suffering from alexia with agraphia cannot recognize words spelled aloud to them nor can they themselves spell aloud a spoken word.

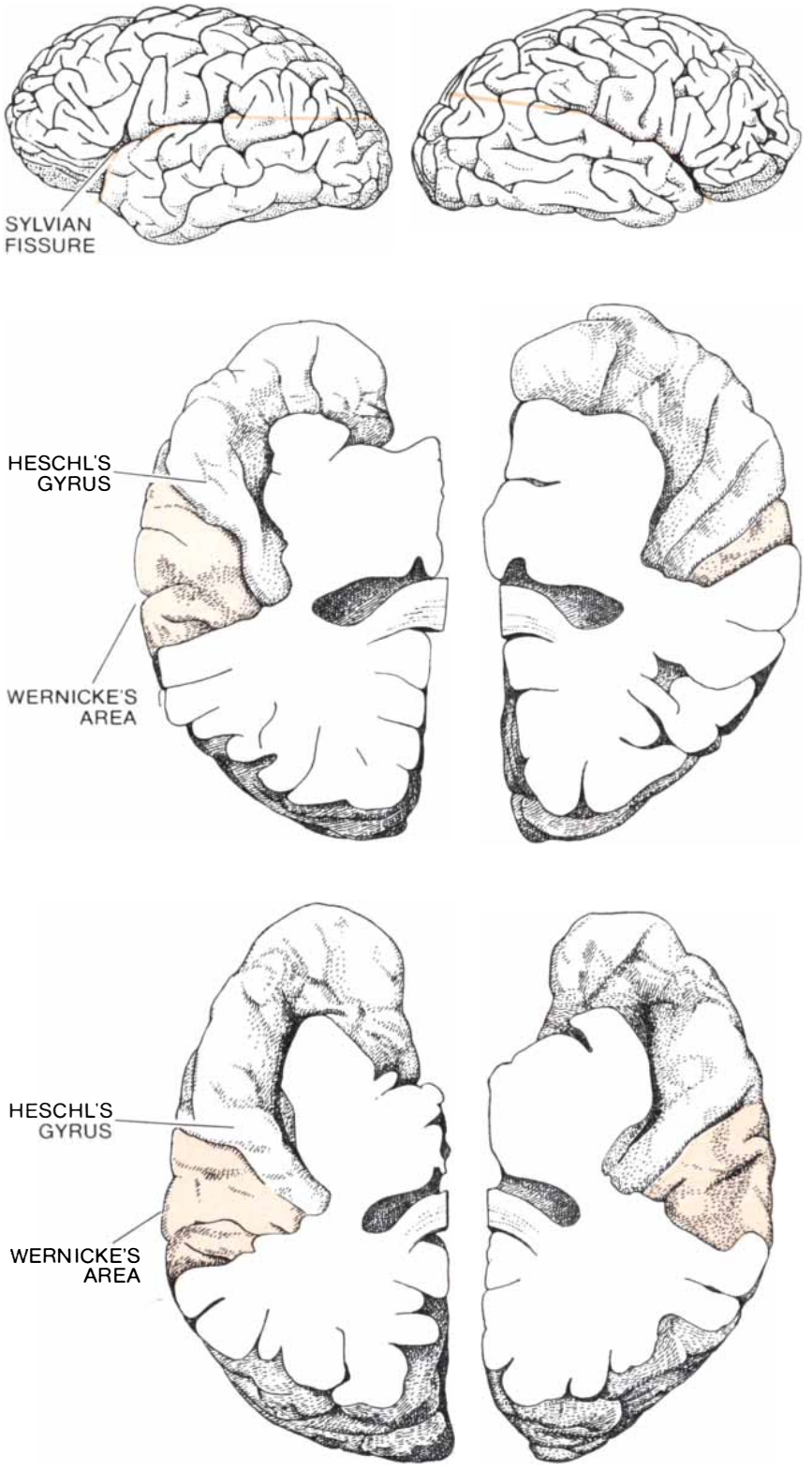
Dejerine's second contribution was showing the importance of information transfer between the hemispheres. His patient was an intelligent businessman who had awakened one morning to discover that he could no longer read. It was found that the man was blind in the right half of the visual field. Since the right half of the field is projected to the left cerebral hemisphere, it was obvious that the man suffered damage to the visual pathways on the left side of the brain [*see illustration on opposite page*]. He could speak and comprehend spoken language and could write, but he could not read even though he had normal visual acuity. In fact, although he could not comprehend written words, he could copy them correctly. Postmortem examination of the man's brain by Dejerine revealed two lesions that were the result of the occlusion of the left posterior cerebral artery. The visual cortex of the left hemisphere was totally destroyed. Also destroyed was a portion of the corpus callosum: the mass of nerve fibers that interconnect the two cerebral hemispheres. That portion was the splenium, which carries the visual information between the hemispheres. The destruction of the splenium prevented stimuli from the visual cortex of the right hemisphere



SAYING THE NAME of a seen object, according to Wernicke's model, involves the transfer of the visual pattern to the angular gyrus, which contains the "rules" for arousing the auditory form of the pattern in Wernicke's area. From here the auditory form is transmitted by way of the arcuate fasciculus to Broca's area. There the articulatory form is aroused, is passed on to the face area of the motor cortex and the word then is spoken.



UNDERSTANDING the spoken name of an object involves the transfer of the auditory stimuli from Heschl's gyrus (the primary auditory cortex) to Wernicke's area and then to the angular gyrus, which arouses the comparable visual pattern in the visual association cortex. Here the Sylvian fissure has been spread apart to show the pathway more clearly.



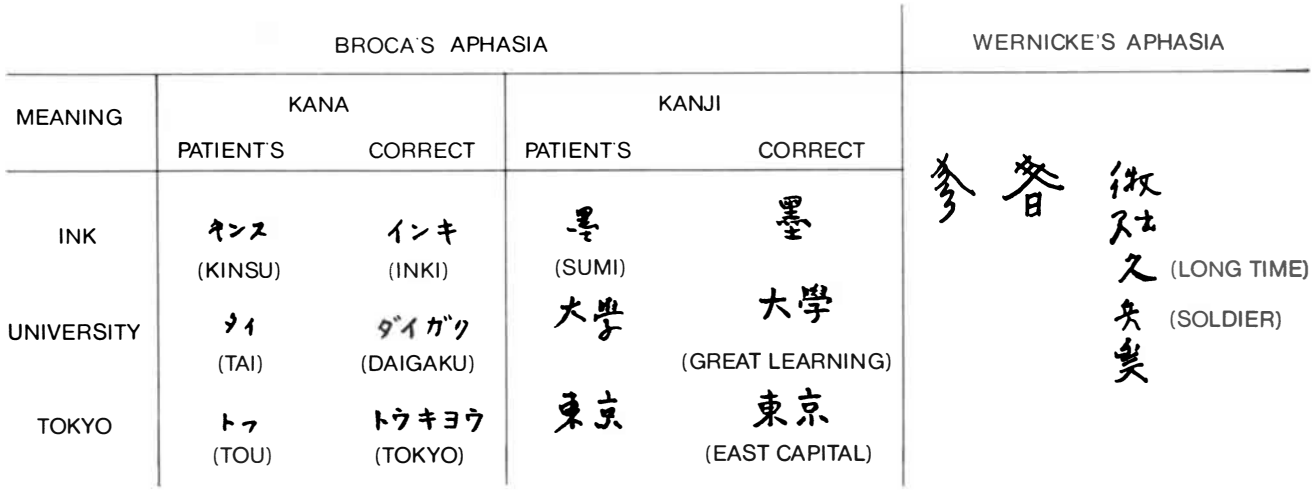
ANATOMICAL DIFFERENCES between the two hemispheres of the human brain are found on the upper surface of the temporal lobe, which cannot be seen in an intact brain because it lies within the Sylvian fissure. Typically the Sylvian fissure in the left hemisphere appears to be pushed down compared with the Sylvian fissure on the right side (*top illustration*). In order to expose the surface of the temporal lobe a knife is moved along the fissure (*broken line*) and then through the brain, cutting away the top portion (*solid line*). The region studied was the planum temporale (*colored areas*), an extension of Wernicke's area. The middle illustration shows a brain with a larger left planum; the bottom illustration shows left and right planums of about the same size. In a study of 100 normal human brains planum temporale was larger on the left side in 65 percent of the cases, equal on both sides in 24 percent of the cases and larger on the right side in 11 percent.

from reaching the angular gyrus of the left hemisphere. According to Wernicke's model, it is the left angular gyrus that converts the visual pattern of a word into the auditory pattern; without such conversion a seen word cannot be comprehended. Other workers have since shown that when a person is blind in the right half of the visual field but is still capable of reading, the portion of the corpus callosum that transfers visual information between the hemispheres is not damaged.

In 1937 the first case in which surgical section of the corpus callosum stopped the transfer of information between the hemispheres was reported by John Trescher and Frank Ford. The patient had the rear portion of his corpus callosum severed during an operation to remove a brain tumor. According to Wernicke's model, this should have resulted in the loss of reading ability in the left half of the visual field. Trescher and Ford found that the patient could read normally when words appeared in his right visual field but could not read at all in his left visual field.

Hugo Liepmann, who was one of Wernicke's assistants in Breslau, made an extensive study of syndromes of the corpus callosum, and descriptions of these disorders were a standard part of German neurology before World War I. Much of this work was neglected, and only recently has its full importance been appreciated. Liepmann's analysis of corpus callosum syndromes was based on Wernicke's model. In cases such as those described by Liepmann the front four-fifths of the corpus callosum is destroyed by occlusion of the cerebral artery that nourishes it. Since the splenium is preserved the patient can read in either visual field. Such a lesion, however, gives rise to three characteristic disorders. The patient writes correctly with his right hand but incorrectly with the left. He carries out commands with his right arm but not with the left; although the left hemisphere can understand the command, it cannot transmit the message to the right hemisphere. Finally, the patient cannot name objects held in his left hand because the somesthetic sensations cannot reach the verbal centers in the left hemisphere.

The problem of cerebral dominance in humans has intrigued investigators since Broca first discovered it. Many early neurologists claimed that there were anatomical differences between the hemispheres, but in the past few decades there has been a tendency to assume that the left and right hemispheres are



JAPANESE APHASICS display some characteristics rarely found in Western patients because of the unique writing system used in Japan. There are two separate forms of such writing. One is Kana, which is syllabic. The other is Kanji, which is ideographic. Kana words are articulated syllable by syllable and are not easily identified at a glance, whereas each Kanji character simultaneously represents both a sound and a meaning. A patient with Broca's aphasia, studied by Tsuneo Imura and his colleagues at the Nihon University College of Medicine, was able to write a dictated word correctly in Kanji but not in Kana (*top left*). When the patient was asked to write the word "ink," even though there is no Kanji character for the word, his first effort was the Kanji character "sumi," which means india ink. When required to write in Kana, the symbols he produced were correct but the word was wrong. Another patient who had Wernicke's aphasia wrote Kanji quickly and without hesitation. He was completely unaware that he was producing meaningless ideograms, as are patients who exhibit paraphasias in speech. Only two of characters had meaning (*top right*).

symmetrical. It has been thought that cerebral dominance is based on undetected subtle physiological differences not reflected in gross structure. Walter Levitsky and I decided to look again into the possibility that the human brain is anatomically asymmetrical. We studied 100 normal human brains, and we were surprised to find that striking asymmetries were readily visible. The area we studied was the upper surface of the temporal lobe, which is not seen in the intact brain because it lies within the depths of the Sylvian fissure. The asymmetrical area we found and measured was the planum temporale, an extension of Wernicke's area [*see illustration on opposite page*]. This region was larger on the left side of the brain in 65 percent of the cases, equal in 24 percent and larger on the right side in 11 percent. In absolute terms the left planum was nine millimeters longer on the average than the right planum. In relative terms the left planum was one-third longer than the right. Statistically all the differences were highly significant. Juhn A. Wada of the University of British Columbia subsequently reported a study that confirmed our results. In addition Wada studied a series of brains from infants who had died soon after birth and found that the planum asymmetry was present. It seems likely that the asymmetries of the brain are genetically determined.

It is sometimes asserted that the anatomical approach neglects the plasticity of the nervous system and makes the likelihood of therapy for language disorders rather hopeless. This is not the case. Even the earliest investigators of aphasia were aware that some patients developed symptoms that were much milder than expected. Other patients recovered completely from a lesion that normally would have produced permanent aphasia. There is recovery or partial recovery of language functions in some cases, as Luria's large-scale study of the war wounded has shown. Of all the patients with wounds in the primary speech area of the left hemisphere, 97.2 percent were aphasic when Luria first examined them. A follow-up examination found that 93.3 percent were still aphasic, although in most cases they were aphasic to a lesser degree.

How does one account for the apparent recovery of language function in some cases? Some partial answers are available. Children have been known to make a much better recovery than adults with the same type of lesion. This suggests that at least in childhood the right hemisphere has some capacity to take over speech functions. Some cases of adult recovery are patients who had suffered brain damage in childhood. A number of patients who have undergone surgical removal of portions of the speech area for the control of epileptic seizures often show milder language disorders than had been expected. This probably is owing to the fact that the patients had suffered from left temporal epilepsy involving the left side of the brain from childhood and had been using the right hemisphere for language functions to a considerable degree.

Left-handed people also show on the average milder disorders than expected when the speech regions are damaged, even though for most left-handers the left hemisphere is dominant for speech just as it is for right-handers. It is an interesting fact that right-handers with a strong family history of left-handedness show better speech recovery than people without left-handed inheritance.

Effective and safe methods for studying cerebral dominance and localization of language function in the intact, normal human brain have begun to appear. Doreen Kimura of the University of Western Ontario has adapted the technique of dichotic listening to investigate the auditory asymmetries of the brain. More recently several investigators have found increased electrical activity over the speech areas of the left hemisphere during the production or perception of speech. Refinement of these techniques could lead to a better understanding of how the normal human brain is organized for language. A deeper understanding of the neural mechanisms of speech should lead in turn to more precise methods of dealing with disorders of man's most characteristic attribute, language.

SUPERCONDUCTORS FOR POWER TRANSMISSION

The rising demand for electricity conflicts with opposition to overhead transmission lines. Underground lines that are refrigerated to be superconducting may provide a solution

by Donald P. Snowden

Throughout this century the use of electrical energy in the U.S. has doubled approximately every 10 years. Although ecological concerns and delays in construction may tend to depress this rate of growth somewhat in the future, it still seems likely that consumption will double and then double again in the next two decades. The demand for electric power is growing twice as fast as the gross national product and more than three times as fast as the population. In the present decade the nation's installed electric generating capacity can be expected to grow from about 300,000 megawatts, the 1970 figure, to about 600,000 megawatts in 1980. In the same period the annual per capita consumption will climb from between 6,000 and 7,000 kilowatt-hours to more than 10,000. In order to meet this growth rate the utility industry will have to install, in the single year 1980, 25,000 megawatts of generating capacity, with transmission and distribution facilities to match, at an estimated total cost of $13 billion. The comparable investment in 1970 was only $7.4 billion.

Finding acceptable sites for the new generating stations and routes for the new transmission lines is already proving difficult and is bound to become more so in the years immediately ahead. The public's desire to have new generating stations located well outside populous areas comes into direct conflict with the desire not to disfigure the countryside with mile after mile of overhead transmission lines. Even if all power plants were located within city limits, extensive networks would still be needed for sharing the load between systems. The obvious solution is to place transmission systems underground. With present technology, however, the cost is prohibitive.

Perhaps the best hope for lowering the cost of underground systems is to develop a technology based on superconductors: conductors that will operate with extremely low (but not necessarily zero) resistive losses when they are refrigerated to a temperature range between 4.2 and 20 degrees Celsius above absolute zero. As we shall see, superconducting systems will be in competition with other high-capacity systems in which the conductors are also refrigerated but not to the point of being superconducting.

In the utility industry a distinction is made between transmission systems, which transport electric power in large blocks from generating stations to major load centers or from point to point for load-sharing, and distribution systems, which carry energy to individual customers. As a rough rule lines operating at 138,000 volts or more are considered to be part of the transmission-system network. Distribution lines usually operate at lower voltage, and within many cities they are now being installed underground routinely. The technology required for these lower-voltage, lower-capacity distribution systems, however, is much less complex than the technology required for transmission systems.

As generating stations have been made steadily larger to take advantage of the economies that come with increasing size, transmission lines have been designed to carry steadily higher voltages. The higher the voltage, the higher the capacity. The highest voltage in general use today in overhead transmission lines is 500,000 volts; some 4,000 miles of transmission line operate at that voltage. Recently a few hundred miles of 765,000-volt line have been placed in service. Systems capable of carrying up to 1.5 million volts are being studied. Although overhead transmission lines will probably be tolerated in remote areas for many years to come, their continued use in and around major load centers, at least for new installations, appears unlikely.

The actual cost of a transmission-line system depends primarily on local conditions. Costs are commonly expressed as the cost per mile per million-volt ampere, representing the product of the line's voltage and current capacity. Typically a 345,000-volt overhead line costs $95 per million-volt ampere per mile. Optimum loading for such a line is about 580 million-volt amperes, which yields a cost of $55,000 per mile. An underground line of conventional design with comparable capacity would cost $1,450 per million-volt ampere per mile, or, with the same loading, $840,000 per mile.

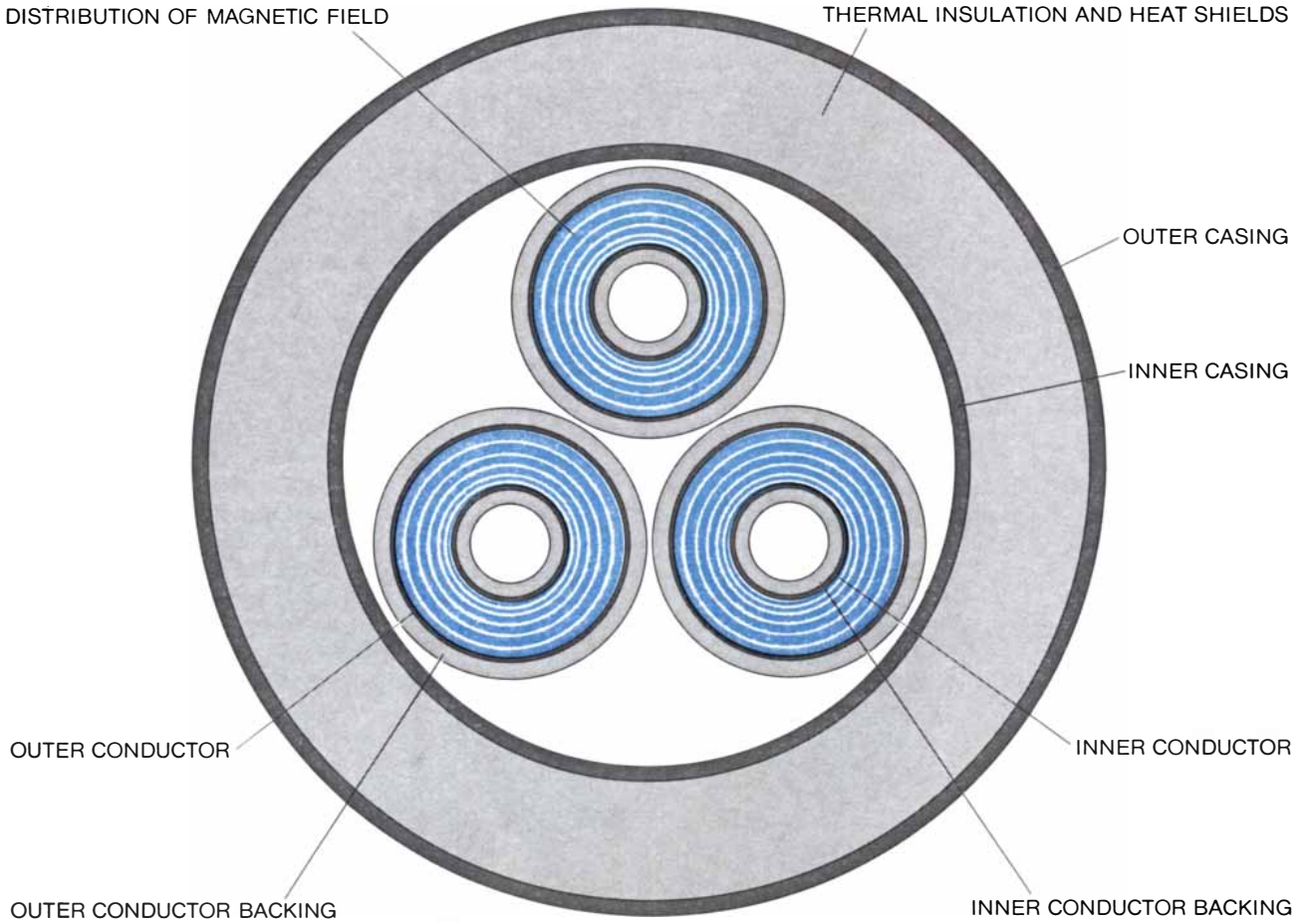
These costs are exclusive of right-of-way. An overhead line of 345,000 volts normally requires a right-of-way 100 feet wide, which amounts to 12 acres per mile of route. If land were priced at $4,500 per acre, the right-of-way cost would just equal the cost of the overhead line itself. In suburban areas land cost can be several times $4,500 per acre. Although land cost would then dominate the total cost of a new transmission line, an overhead line could still be justified economically. If, however, the cost of land were as much as $70,000 per acre, which it might be in densely populated areas, the right-of-way alone would cost as much as an under-

ground line ($70,000 \times 12 = $840,000). In that case the utility would be better off to choose an underground system for which right-of-way costs are negligible. In urban areas a utility can usually obtain permission to bury cable alongside highways or under city streets. It is estimated that increasing land costs combined with public pressure against overhead installations will create an annual market of $1 billion for underground transmission lines by 1980. Therefore the utility industry has considerable incentive to find ways to reduce the cost of putting its power lines underground.

In evaluating new technologies of power transmission utilities are also looking closely at the advantages offered by direct-current systems. Historically alternating current has been used almost exclusively because of the ease with which alternating-current line voltages can be raised or lowered by means of transformers. Except in special cases alternating current will continue to be what is delivered to the customer. For moving large blocks of power from point to point, however, direct current has three important advantages over alternating current.

The first advantage is that a direct-current line requires only two conductors, whereas an alternating-current line requires three, one for each phase in the typical three-phase system. The second advantage results from the elimination of problems inherent in lines carrying alternating current. In overhead lines the magnetic field that is created, destroyed and re-created around the conductor as the current alternates from plus to minus 60 times per second induces an out-of-phase voltage drop along the conductor that reduces the amount of power that can be delivered to the load. In a cable, where the three conductors of an alternating-current system are each surrounded by a grounded shield, the maintenance of the alternating electric field between each conductor and ground requires an out-of-phase current to flow in the conductor. This similarly reduces the amount of current available to deliver power to the load. These effects can be "tuned out" by installing special equipment that adds a capacitance to match the inductance in overhead systems and an inductance to match the capacitance in underground systems, but the equipment is costly. Partly because of this problem the longest underground circuits now in service above 230,000 volts are no longer than about 20 miles.

The third advantage of direct current over alternating current comes into play when one wishes to interconnect systems operating at different frequencies or load levels. In an interconnected alternating-current system all generators must rotate synchronously and in phase with one another. A disruption in one generator can be easily propagated throughout the entire system if adequate provisions for disconnection are not made. The problem can be forestalled by using a direct-current link to tie two alternating-current systems together. An



SUPERCONDUCTING TRANSMISSION CABLE is shown in cross section. Current would flow in the three conductors at center, each consisting of an inner conductor and an outer one, which is a return conductor. The magnetic field associated with the flow of current would be confined in the space between each inner and outer conductor. In a system with helium as a coolant to maintain the superconductivity of the conductors the helium would also occupy the space between the inner and outer conductor. Thermal insulation would help to maintain low temperature. A 138-kilovolt system such as this one would be about 55 centimeters in diameter.

example is a direct-current system only 35 feet long being installed by the New Brunswick Electric Power Commission in Canada to serve as a link between two alternating-current systems. If the link had been made with alternating current, the stability of the two systems would have been compromised.

These three advantages are offset, however, by the high cost of converting alternating current to direct current for transmission and back again to alternating current for the customer. The conversion cost is on the order of $35 million per system, or the equivalent of 40 miles of underground line. As a result direct-current transmission can normally be justified only for lines that are several hundred miles long. The largest direct-current system has recently been installed between hydroelectric generators on the Columbia River in the Pacific Northwest and in southern California. The 850-mile line carries 1,350,000 kilowatts at ±375,000 volts.

The high cost of conventional underground cable is attributable to the difficulty of adequately insulating three high-voltage conductors and placing them side by side in a conduit only a few feet in diameter [*see illustration below*]. Each conductor consists of a bundle of several hundred strands of copper wire. Because alternating current tends to flow on the surface of a conductor (the "skin effect") the strands are usually transposed and twisted so that the current is forced to travel throughout the volume of the conductor. The skin effect is further reduced by dividing the multisectioned conductor into sectors of a circle that are lightly insulated from one another. The entire conductor is wrapped with paper insulating tape. Finally, the wrapped cable is shielded with a combination of metal and polymer tapes and the inner layers of paper are impregnated with oil. Three such cables are pulled into a steel pipe, which is filled with oil at a pressure of between 200 and 225 pounds per square inch. Meticulous care is required during both manufacture and installation to exclude moisture, which would seriously degrade the insulating properties of the system. The first cable of this type capable of carrying 345,000 volts of alternating current was installed in New York City eight years ago.

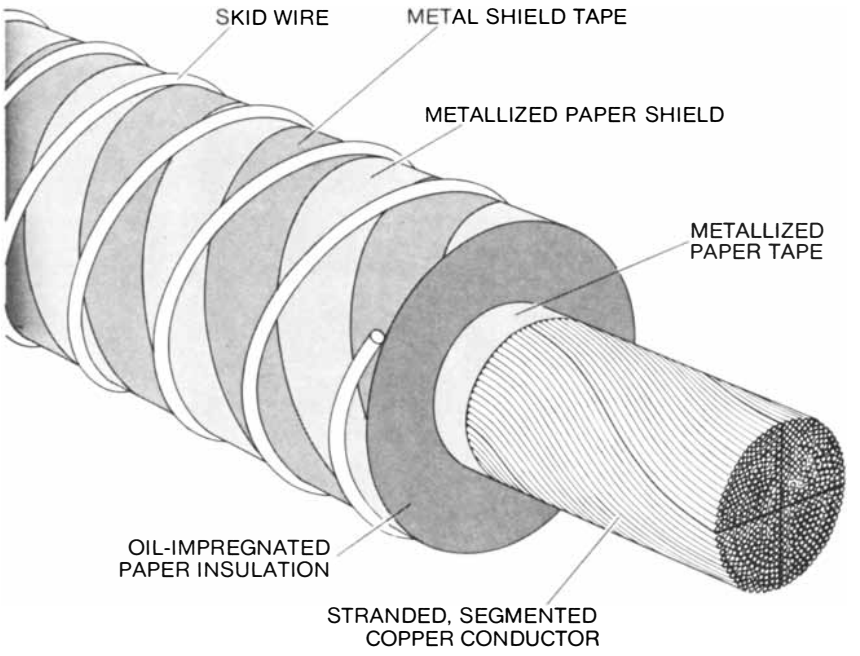
The principal limitation on the capacity of an underground cable is the buildup of heat, caused by the motion of charged particles (ions) or polarized molecules (dipoles) in the electric field around the conductors. For a 345,000-volt line operating at rated capacity this dielectric loss, as it is called, accounts for more than a third of the total line loss. The temperature at which an underground cable operates and therefore the total loss that can be tolerated are functions of the thermal properties of the soil in which the line is buried. The trench containing the line is usually filled with soil that has a high thermal conductivity in order to carry away as much heat as possible, thereby maximizing the capacity of the line.

Because of the buildup of heat a 345,000-volt underground line will normally have less than half the capacity of an overhead line with conductors of equal size. Additional capacity can be obtained, however, by circulating and cooling the insulating oil in the pipe containing the cables or by circulating a coolant adjacent to the pipe. For example, the Commonwealth Edison Company in Chicago has recently added an oil-circulating and refrigerating system to a 138,000-volt cable that increased its capacity from 200 million-volt amperes to 300.

A number of developments promise to increase the capacity and reduce the cost of underground transmission systems. One direction is to use insulating materials with properties superior to those of oil-impregnated paper. One promising material is extruded polyethylene. Because of its low dielectric constant it should make possible the fabrication of a cable that can carry a given amount of power 30 percent farther than the maximum distance now possible with a cable that has oil-and-paper insulation. In addition the dielectric loss in polyethylene is only 10 percent of the loss in oil-impregnated paper insulation. So far, however, satisfactory techniques have not been developed for extruding void-free polyethylene in the size needed for a 345,000-volt cable.

Of even greater promise are cables in which either compressed nitrogen or sulfur hexafluoride is used as the insulating medium. Both gases have a dielectric constant of unity, equal to that of a vacuum. Moreover, they combine a low dielectric loss with good heat-transfer properties. In spite of these favorable characteristics high-capacity cables insulated with either of these gases will require forced cooling to overcome the low thermal conductivity of the surrounding soil. A short line insulated with sulfur hexafluoride has recently been installed by the Consolidated Edison Company of New York.

Since cooling appears to be necessary for any high-capacity underground line,

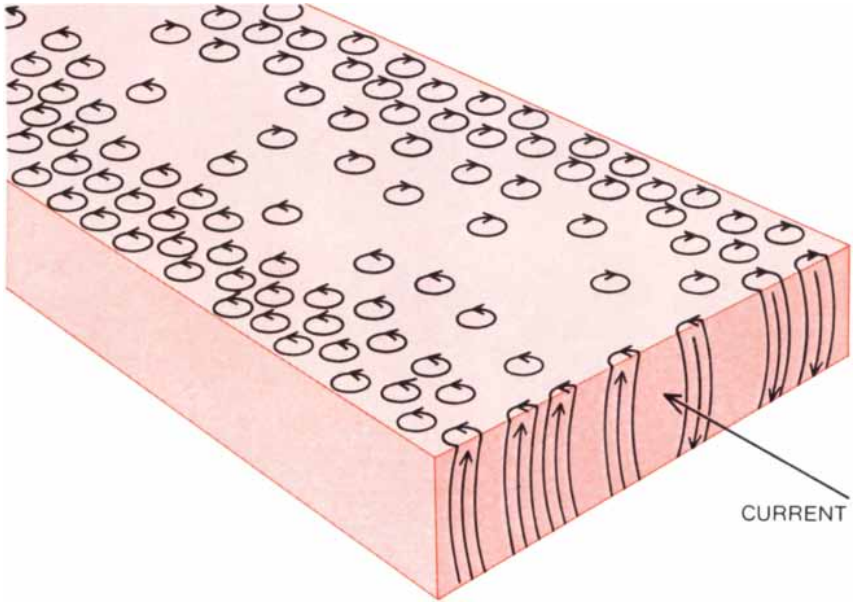


SINGLE CABLE of the type now used in high-voltage but nonsuperconducting underground transmission of power consists of a bundle of strands of copper wire, twisted and segmented to counteract the tendency of current to flow on the surface of a wire. Three such conductors are elaborately insulated and pulled into a steel conduit, which is filled with oil that is under pressure. The pressurized oil further improves the insulation of the system.

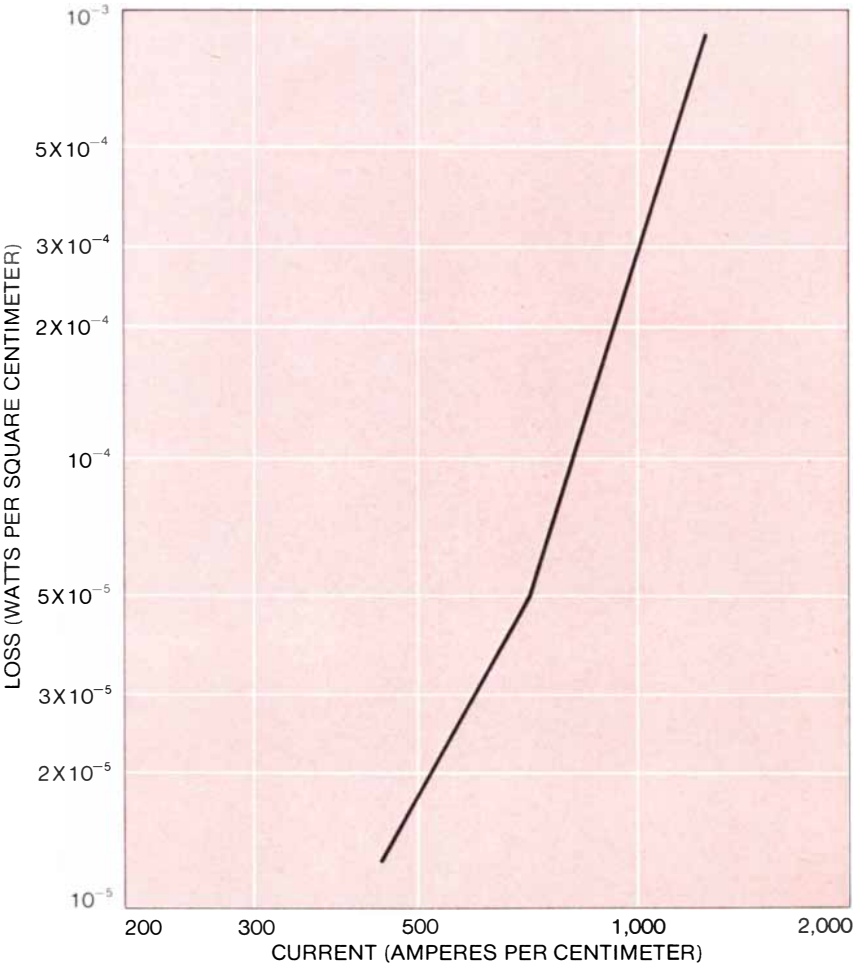
it should be recognized that there are advantages to refrigerating the line well below ordinary soil temperatures, thereby lowering the resistivity of the conventional metals used as conductors. For example, at the boiling point of liquid nitrogen, 77 degrees Kelvin (degrees Celsius above absolute zero), the resistivity of aluminum or copper is reduced by a factor of 10, which just about pays for the cost of refrigeration. The improvement at 20 degrees K., the boiling point of liquid hydrogen, is greater: the resistance loss is reduced by a factor of 500 and the cooling cost increases only by a factor of 30 to 100.

It must be recognized, however, that the goal of new types of transmission systems is not a reduction in operating losses but rather a reduction in capital costs. In other words, it is clearly preferable to generate an extra kilowatt-hour of energy for half a cent than to spend a cent on refrigerating a line to keep from losing that amount of energy. Only for high-capacity overhead lines operating at full capacity (that is, with maximum losses) do the capitalized losses become comparable to the capital costs. With this exception the losses capitalized by any reasonable system are a small fraction of the cost of any present or contemplated transmission line—overhead or underground.

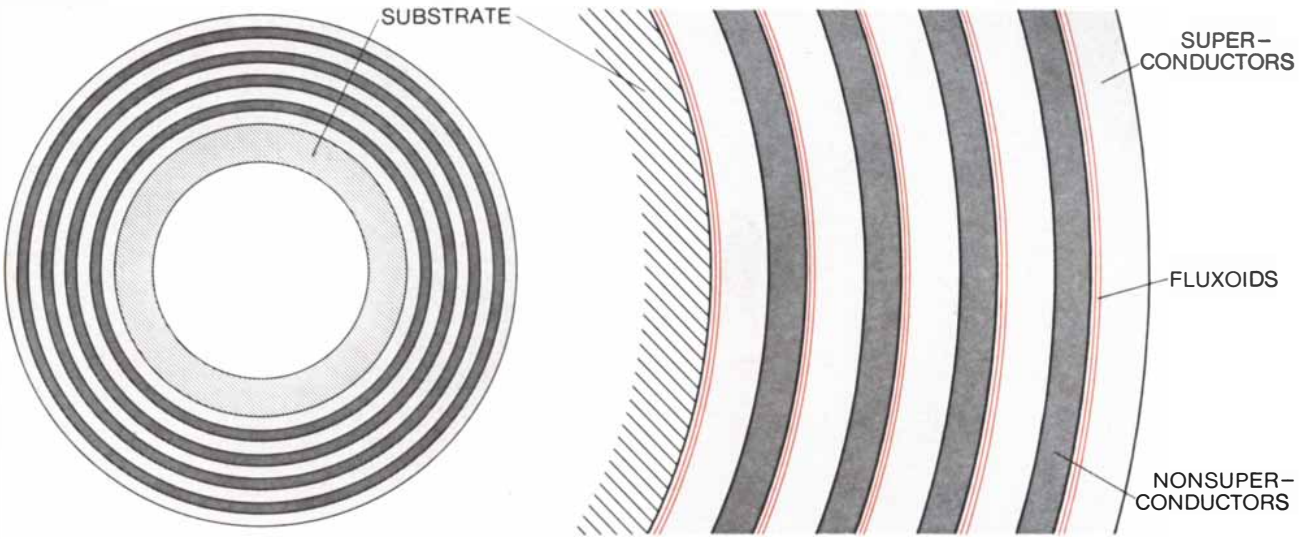
Cryogenic transmission systems, even in the present state of the art, have one major economic advantage over other underground systems: the capacity of the facility can be increased at a later time without tearing up the line or building a new one. Since the capital cost of the refrigeration system can amount to as much as half of the total cost of the line, one can install a line of reduced capacity initially, and when additional capacity is needed, it can be provided by adding more refrigeration. Moreover, one can allow the line temperature to rise somewhat during periods of large load, accepting increased losses for a short period. It is necessary only that the refrigeration capacity be equal to the capacity required by the average load. Finally, if the system is cooled with an inexpensive and inert refrigerant such as liquid nitrogen, one can meet short-term emergency overloads by allowing some of the refrigerant to boil off to the atmosphere, carrying the heat of vaporization with it. When the emergency is past, the nitrogen can be replaced. Several industrial firms, supported in part by the Edison Electric Institute, are pressing the development of cryogenically cooled ca-



DISTRIBUTION OF FLUXOIDS in a strip of Type II superconductor carrying a current is depicted. Fluxoids are magnetic fields in the form of quantized vortexes. They penetrate the interior of the conductor in such a way that current can flow in most of the conducting material, rather than on the surface as in a Type I superconductor. The distribution shown is typical for the case in which the current has been initially increased from zero.



POWER LOSS in a Type II superconductor made of niobium increases approximately as the third power of the current at lower currents and as fourth power at higher ones. Chart shows loss of alternating current at 60 cycles per second at temperature of liquid helium.

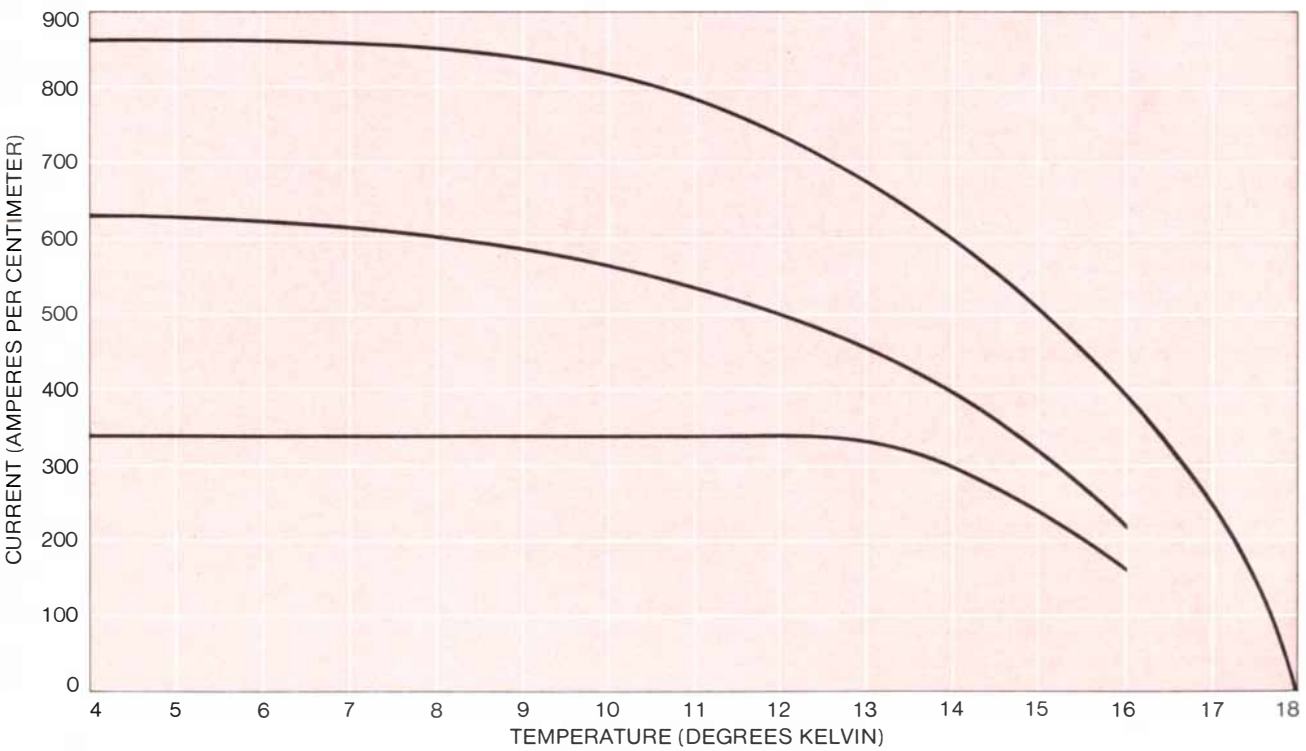


FLUXOID-PINNING reduces power loss by increasing the interaction of superconducting metal and fluxoids. Pinning is achieved in a cylindrical structure (*left*) consisting of thin concentric layers of a superconducting compound of niobium and tin alternating with layers of a nonsuperconducting material. The interface between the superconductor and the nonsuperconductor serves to pin the fluxoids, as depicted schematically in the enlarged view at right. Reduction of power loss allows higher operating temperature.

bles with the expectation that a liquid-nitrogen cable can be in operation by the mid-1970's.

For meeting the transmission needs of the somewhat more distant future the superconducting cable offers the greatest promise. It will also require the greatest development and engineering effort. The primary attractiveness of a superconducting transmission line is not so much its potential ability to carry current with a low loss as it is its ability to carry a large amount of current without very high voltages. Until now increasing the voltage has been the only way to increase a transmission line's capacity.

I shall deal primarily with alternating-current systems, which is not to say that direct-current superconducting lines may not someday be used. The first superconducting systems will be tailored for metropolitan areas where the distances involved are not great. Direct-



TEMPERATURE DEPENDENCE of current in a layered niobium-tin cylinder is charted. Critical current for this superconductor is shown by the top curve; next two curves respectively show the current for a loss of 10^{-5} and 10^{-6} watt per square centimeter.

current systems, as we have seen, become economic only when the lines are long.

In 1911 the Dutch physicist Heike Kamerlingh Onnes discovered that certain metals, such as aluminum, tin, lead and mercury, lose all resistance to the flow of electric current when they are cooled almost to absolute zero. Unfortunately these metals, now known as Type I superconductors, could carry very little current without losing their superconductivity. Beginning in the 1950's the situation was dramatically changed with the discovery of a new class of metals, known as Type II superconductors, that remain superconducting while carrying substantial amounts of current [see "The Magnetic Structure of Superconductors," by Uwe Essmann and Hermann Traüble; SCIENTIFIC AMERICAN, March, 1971].

Type I materials are unable to carry much current because such a conductor in the superconducting state does not allow a magnetic field to penetrate its interior. A fundamental law of electromagnetic theory states that a current cannot flow unless it is accompanied by a spatially varying magnetic field. Type II materials allow magnetic fields to penetrate the interior of the conductor in the form of quantized vortexes, or fluxoids, thereby enabling current to flow in the bulk of material [*see top illustration on page 87*].

Fluxoids participate in two types of interaction of special importance. First, they interact with the current and are driven in a direction mutually perpendicular to the direction of current flow and to the direction of the magnetic field. Therefore fluxoids created by the self-magnetic field of a current are driven from the surface to the interior of the conductor. Second, the motions of fluxoids are impeded by imperfections present in the superconducting metal. These two fluxoid interactions allow the transfer of energy from the moving current to the stationary metal and thus lead to a substantial dissipation of power.

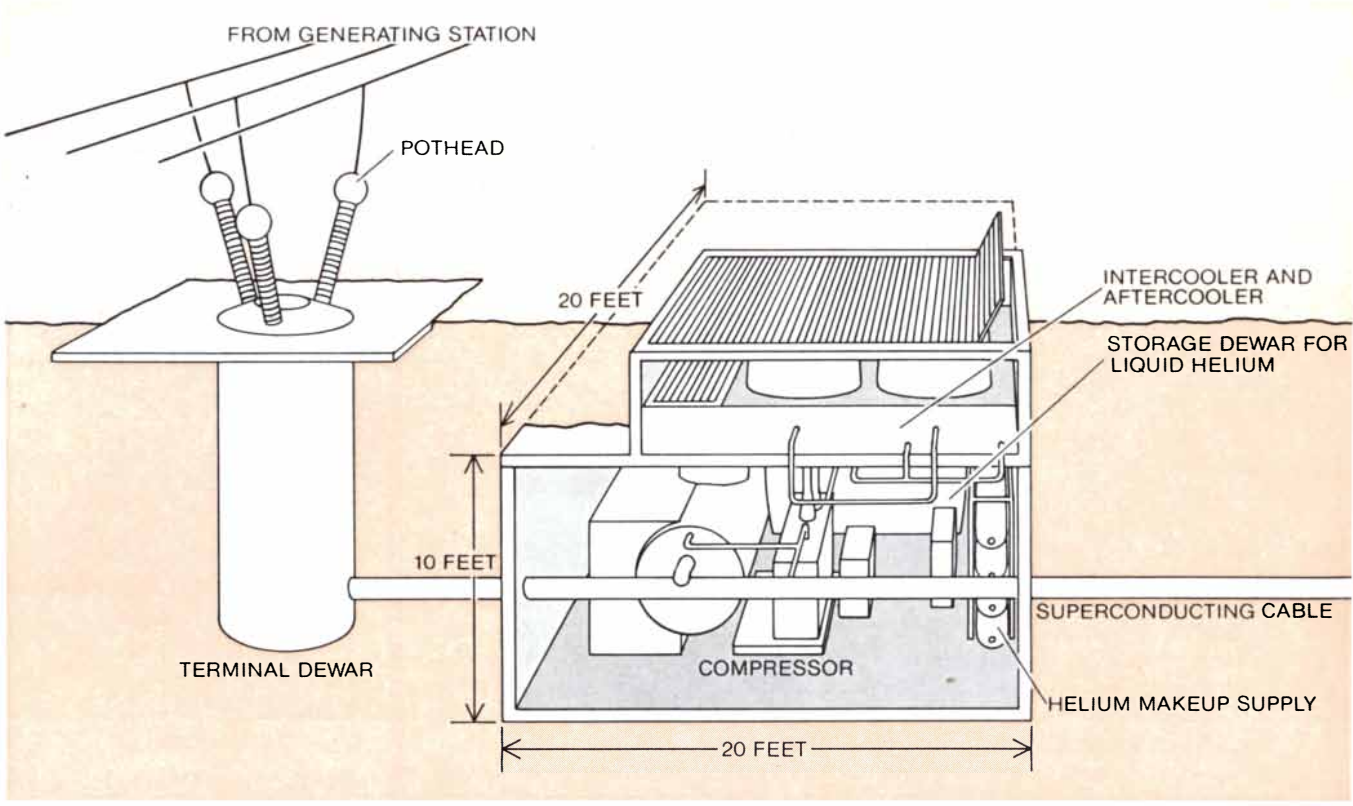
This loss is acceptably small in a Type II superconductor carrying a direct current because a fixed distribution of fluxoids can support a direct current and the only interaction with the superconductor is that produced by a small slippage of fluxoids past the fixed distribution. With alternating currents, however, the fluxoids are continuously in motion because of the changing current. The motion causes a power dissipation that increases linearly with frequency; worse yet, the power loss often increases as the third or fourth power of the magnitude of current flowing.

The losses are so large that it was originally thought Type II superconductors would not be useful as carriers of alternating current. Recent research, however, has shown ways in which they can be used. In a Type II material fluxoids form only above a certain magnetic field called the lower critical field. Below this critical value the material acts as a Type I superconductor, so that current is carried only in a thin layer at the surface. In this layer there are no losses whether the current is direct or alternating. In most materials the magnitude of the current flow is too small to be useful in a power-transmission system.

One Type II superconductor, however, has a lower critical field whose value is uniquely high. The material is niobium, in which fluxoid formation first occurs when the magnetic field reaches 1,400 oersteds at 4.1 degrees K. (the boiling point of liquid helium). This value enables a cylindrical conductor of niobium to carry an average of 790 amperes of alternating current for each centimeter of its circumference without exceeding the lower critical field. (The average flow of an alternating current, technically called the root-mean-square current, amounts to about 70 percent of the peak flow.) With a current flow of 790 amperes per centimeter it is possible to design a superconducting transmission system.

In order to obtain the benefits of the cylindrical geometry in a three-phase system it is necessary to surround each of the three conductors with a coaxial return conductor that is also superconducting [*see illustration on page 85*]. The magnetic field associated with each conductor is then confined between the inner and the outer conductor; therefore no interaction can take place between the magnetic field of one line and the current in another, nor can the magnetic field interact with the walls of the refrigerated jacket or conduit. Triaxial and quadriaxial geometries have also been proposed. Niobium superconducting transmission lines are being studied by the Linde Division of the Union Carbide Corporation in the U.S. and by the Central Electricity Research Laboratories and British Insulated Callender's Cables Ltd. in Britain.

The dielectric for insulating a superconducting power line will probably be either a vacuum or liquid helium. Most solid dielectrics would give rise to prohibitive losses, although polyethylene or



INSTALLED SYSTEM for an underground superconducting line carrying alternating current through a heavily populated area is portrayed according to a design by the Union Carbide Corporation. The power enters the system at left by way of an overhead line that carries it from the generating station some miles away. The line then passes through a refrigerating station that maintains in

polypropylene may prove useful in systems of moderate voltage. Helium will probably emerge as the final choice because it can also serve as a refrigerant.

A critical problem for a cryogenic transmission line is providing absolutely reliable refrigeration equipment. Until now few, if any, refrigerators have had to operate unattended for years on end. Total heat loads at 4.2 degrees K. will be of the order of 150 to 300 watts per mile, or only somewhat more than the load on a home freezer. But whereas a home freezer can extract 300 watts of heat with an expenditure of only about 357 watts of electricity, a cryogenic refrigerator may have to expend as much as 100,000 watts (100 kilowatts) per mile of line. The reason is that low-temperature refrigerators are inherently very inefficient.

Although an expenditure of 100 kilowatts per mile may seem a high figure, it will constitute only a small fraction of the total power transmitted. The spacing between refrigerators can be anywhere from five to 50 miles, depending at least in part on whether liquid, gaseous or supercritical helium is the coolant. One way to achieve reliability is to provide standby compressors, which could also be used during cooldown. Another proposal is to make the refrigerators oversize so that adjacent units could pick up the load in case one unit failed.

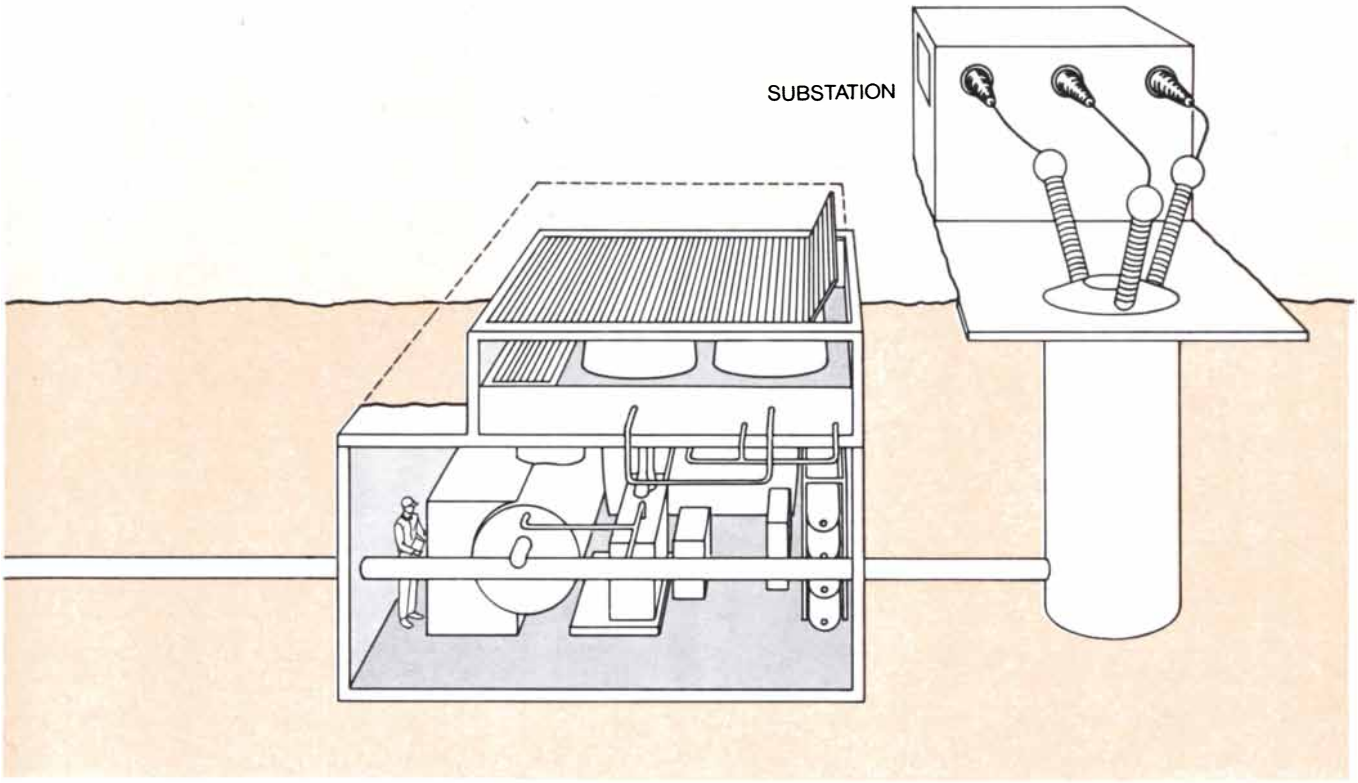
In experiments with niobium conductors done by Linde and others one striking and apparently general result has emerged. Even though the niobium is held below the lower critical field, the losses are not zero, as one would expect, but increase approximately as the third power of the current [*see bottom illustration on page 87*]. The size of the losses in niobium is strongly dependent on the quality of the conductor's surface. Apparently any roughness comparable in scale to the size of a fluxoid allows the local magnetic field to be much greater than the average value. Fluxoids then form in the surface imperfection and give rise to a loss.

Although the alternating-current losses in niobium are not zero, they are small enough to be acceptable in a transmission system. A typical design envisions coaxial conductors with an inner conductor 6.7 centimeters in diameter capable of carrying 4,710 million-volt amperes at 230,000 volts. (This is more than eight times the capacity of the typical 345,000-volt line mentioned above.) The conductor loss per mile for such a line held at 4.2 degrees K. would amount to 79 watts, and the heat leaking into the insulated cable would add another 225 watts.

The Linde group has estimated that a superconducting 138,000-volt line 10 miles long with a capacity of 1,690 million-volt amperes would cost about $15 million, or about $880 per million-volt ampere per mile, including capitalized operating costs. The cost per mile is only about 60 percent of the cost of a conventional underground transmission line today. The Linde group believes the superconducting line could be in operation by the 1980's.

Since the expectation of zero loss in an alternating-current niobium line evidently cannot be achieved, one can consider other superconducting materials. If one could use a material that remained superconducting at a higher temperature, one could trade possibly higher transmission losses for a saving in the cost of refrigeration. Operation at higher temperatures should also increase the thermal stability of the line.

At Gulf General Atomic we have been studying the feasibility of using a compound of niobium and tin (Nb_3Sn), which remains superconducting at about twice the temperature of niobium. The lower critical field of Nb_3Sn, however, is very low, which means that it will be



the system the low temperature necessary for superconduction. The station is shown in a cutaway view. Additional refrigerating stations would be necessary about every five miles; a second station that is some five miles from the first one is depicted at right. Such a superconducting system is envisioned as being from five to 20 miles long. At the end it would deliver power to a substation.

useful in transmission lines only if it can be used in the Type II, or fluxoid, regime. This in turn will be possible only if the alternating-current losses due to the motion of fluxoids are acceptably low.

It has been demonstrated that the alternating-current loss can be reduced by increasing the interaction of the superconducting metal and the fluxoids. This interaction, called fluxoid-pinning, operates through various inhomogeneities in the metal that provide energetically favorable locations for the fluxoids. As the density of pinning centers is increased the current loss decreases.

We have been experimenting with cylindrical structures consisting of thin concentric layers of Nb_3Sn alternating with layers of a nonsuperconductor, fabricated by vapor deposition in a vacuum [*see top illustration on page 88*]. The interface between the superconductor and the nonsuperconductor serves to pin the fluxoids. The currents carried by our samples decrease only a small amount over a temperature range from 4.2 degrees K. to 15 degrees K., verifying the expectation that Nb_3Sn would be useful at significantly higher temperatures than niobium alone [*see bottom illustration on page 88*]. The losses and current capacity of Nb_3Sn composite structures at between 10 and 12 degrees K. are comparable to those of niobium at 4.2 degrees K., a difference that should yield a saving in the capital cost of the refrigeration system of more than a factor of three.

In spite of the arguments presented above, the question of the design of a high-capacity underground transmission line for the 1980's and 1990's is still undecided. Clearly the line will be force-cooled, but the situation is too uncertain to predict now what temperatures will prove to be the most economic. It has been shown that the line could be superconducting. The results from the physics laboratory have been favorable, and preliminary engineering designs show no insurmountable obstacles, but many details must still be worked out. Economic estimates have also been favorable. Much the same could be said, however, about the other advanced-concept transmission systems. Indeed, the system in the forefront in future years may be adopted only because it received the most development effort.

The particularly farsighted workers, or perhaps the more adventurous dreamers, look well beyond the need for transmitting a few thousand megawatts into a densely populated area to the day when the entire power-generation capacity of the country will consist of large nuclear generating stations located on the East and West coasts (possibly offshore) or alternatively on the continental divide, with large-capacity underground transmission lines feeding power to consumers throughout the nation. Because of their high current capacity, superconductors are probably the only candidates for the conductors in such lines.

One early detailed proposal by Richard L. Garwin and Juri Matisoo of the Thomas J. Watson Research Center of the International Business Machines Corporation envisions a 600-mile superconducting direct-current line with a capacity of 100,000 megawatts. Such a line, including the converters, was estimated to cost $13 per million-volt ampere per mile, considerably below even the cost of overhead transmission today. Such a transmission line, although it is still far in the future, would be a natural evolution of the smaller superconducting lines now being studied. It could lead to a complete restructuring of the nation's power-generation and transmission capabilities. Such a large-scale upgrading of capacity is not possible for the other advanced-concept transmission systems now under consideration.

Environmental Control in the Beehive

Honeybees efficiently regulate temperature and humidity, eliminate polluted air, remove foreign objects, wastes and dead bodies and control parasites and pathogens that attack them and their food

by Roger A. Morse

A colony of honeybees consists of as many as 50,000 individuals, which live together under conditions of much greater crowding than human beings ever do. Yet the bee colony can maintain its internal environment with remarkable constancy. It can regulate temperature and humidity, eliminate polluted air, remove foreign objects, wastes and dead bodies and, not least, control most of the parasites and pathogens that attack not only the bees but also their food supplies in the hive. How do the bees do it?

A bee colony consists of the queen, a large complement of worker bees and, during part of the year, as many as 3,000 drones. The function of the drones is to mate with the queen; only about six to eight of them do so, and the rest are superfluous. In the fall the drones are driven out of the colony.

The three castes live together in a natural cavity or a man-made hive. In nature a hollow tree seems to be preferred. Bees usually build their combs in a dark place; indeed, light appears to inhibit comb construction. The result is that bees seldom build combs in exposed areas where the colony could not survive the winter.

Honey is accumulated by bees as a reserve against the time when nectar is not available from flowers. Honey is the bees' chief food; a normal colony will maintain a reserve of from 15 to 100 pounds of it. In the northern U.S. a colony requires at least 60 pounds of honey to get through the winter.

Bees obtain their protein from pollen. The pollen reserve of a colony will normally range between one pound and 15 pounds. Honey and pollen are the only foods the bees eat, and both are stored in the cells of the comb. Since honey is largely sugar and pollen is rich in protein and fats, the reserve could be an important source of food for organisms other than bees. Protecting the stored food against larger animals and parasites is therefore a matter of prime importance to the colony.

The home usually selected by bees in nature is strong enough to protect the nest against attack by other animals. If the nest fails to give sufficient protection, a second line of defense is the sting. Even in the winter bees too cold to fly can protrude their stingers; a mass of cold clustered bees with protruded stingers reminds one of a porcupine. Any animal that touches such a cluster is likely to retreat in haste. Predators are therefore more of an occasional hazard than a constant threat.

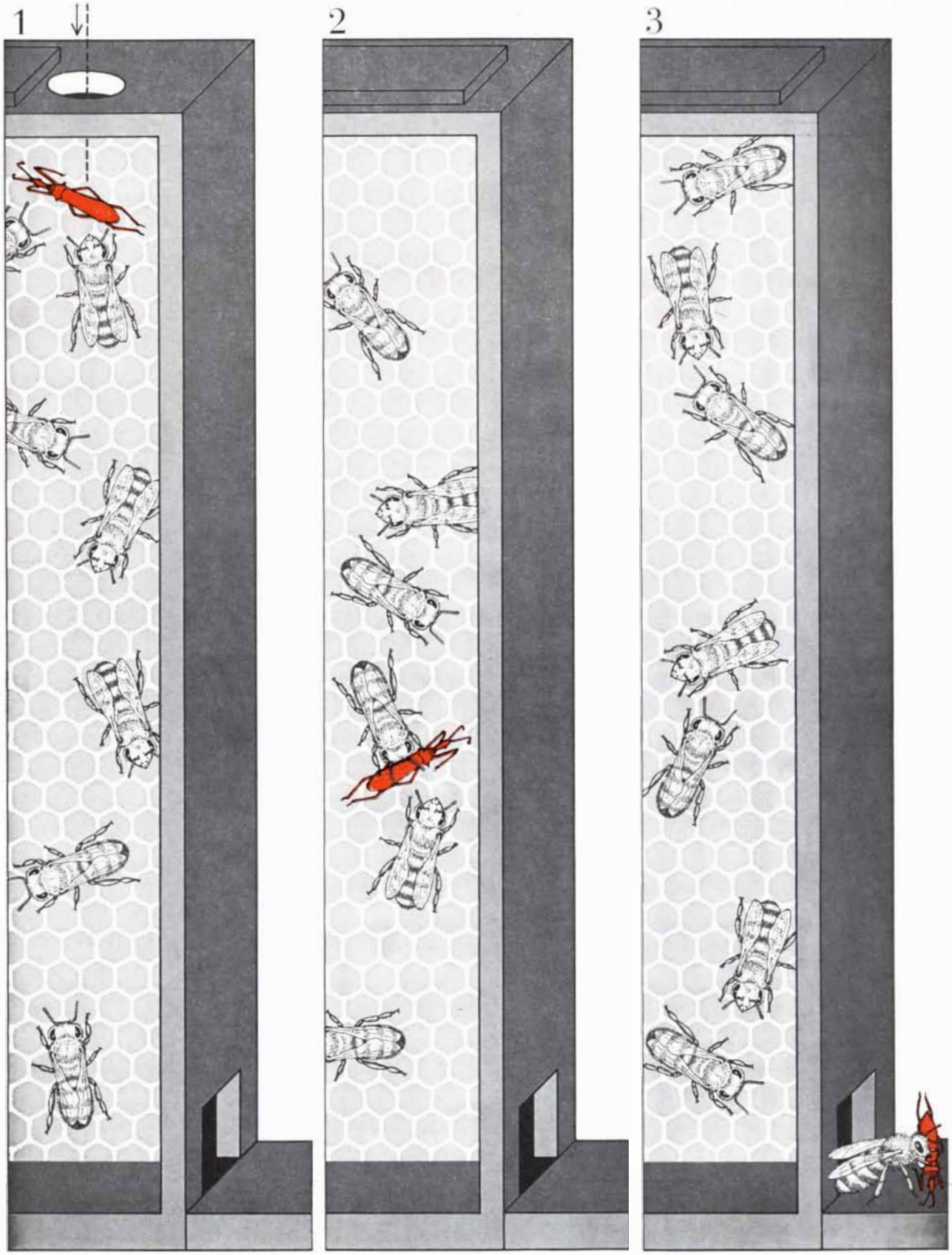
Protecting the reserve against parasites such as bacteria calls for other mechanisms. The problem is exacerbated by the fact that the bees maintain a temperature of about 92 degrees Fahrenheit in the nest during the period of about 10 months a year when they are rearing brood. Humidity is also maintained at a high level, although within wider limits. Such conditions would be favorable for the growth of yeasts and bacteria if the bee colony did not have protective mechanisms.

The investigation of these mechanisms began early in this century, coincident with the establishment of pure food and drug laws. At the time there was much interest in foods as possible carriers of typhoid fever and other infectious diseases. Walter G. Sackett of the Colorado Agricultural College investigated honey as a possible carrier of intestinal diseases in 1919. He introduced 10 different microorganisms into honey, including the organisms that cause typhoid fever and bacterial dysentery. He also introduced the same bacteria into honey that had been diluted with water.

Sackett found that in the undiluted honey the microorganisms were dead within two days and that in honey diluted to 60 percent of its original sugar content they were dead within a day. (Undiluted honey is from 82 to 84 percent solids, mostly sugar.) In solutions of honey with less than 50 percent solids the organisms died more slowly, but even in a solution of 10 percent honey they were killed within a few days.

Sackett found the failure of the microorganisms to thrive in honey "rather surprising," particularly in the case of the diluted honey. He could only suggest that the phenomenon was related to the physical state of the sugar particles in the honey. Later investigation showed that one antibacterial mechanism has to do with the fact that honey has a low water content and a high osmotic pressure. In such an environment osmotic pressure, resulting from the imbalance between the amount of water in a bacterial cell and the amount in the honey, causes water to leak out of bacterial cells. The result is bactericidal or bacteriostatic: some bacteria are killed and others, although they remain viable, cannot grow. Osmotic pressure therefore accounts for the failure of bacteria to flourish in undiluted honey.

INTERIOR OF HIVE of a colony of bees active in Florida appears in the photograph on the opposite page. The absence of foreign objects is conspicuous; bees promptly remove such an object. The cells visible in the photograph are for storing pollen and honey. Pollen is the source of protein and fat in the bees' food as honey is the source of sugar. Protective devices for pollen cells include putting a layer of honey in the top fifth of the cell and capping the cell with wax.



REMOVAL OF FOREIGN OBJECT from a hive by worker bees is portrayed. Soon after the object is introduced (*1*) a worker bee grasps it and moves it some distance, usually toward the entrance. If the first bee abandons the object, another bee will pick it up (*2*). Eventually the object reaches the entrance (*3*), where a bee flies away with it, carrying it at least 50 feet from the hive before dropping it. If the object is too large for a bee to carry in flight, the bee will try to drag it several feet over the ground.

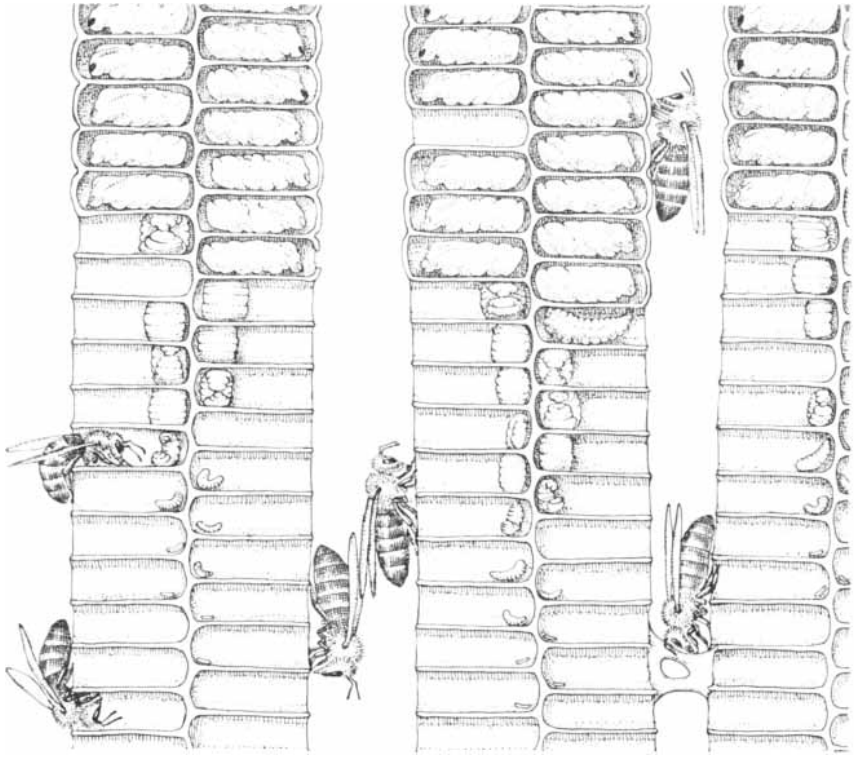
It took longer to find the mechanism that accounts for the stronger effect of partly diluted honey. In 1937 German experimenters called attention to the bactericidal effect of honey and termed it the "inhibine" effect. The laboratory investigating the phenomenon reported that whatever caused the effect was destroyed by heat and light. This finding was confirmed in other laboratories; indeed, the "inhibine number," a measure of bactericidal activity, has been used in some countries since the 1930's to indicate the quality of honey.

In 1962 Jonathan W. White of the U.S. Department of Agriculture discovered that honey contains the enzyme glucose oxidase, which is added to the honey by worker bees. Glucose oxidase attacks glucose, which is one of the two common sugars in honey. (The other is levulose, a form of fructose.) In the process hydrogen peroxide, which is of course a powerful bactericidal agent, is released. White's discovery showed the real basis of the inhibine effect.

Moreover, White found that glucose oxidase is largely inactive in ordinary honey with a moisture content of 19 percent or less. He showed that when honey is diluted, the generation of hydrogen peroxide rises sharply. Here, then, is the reason for the increased bactericidal effect of diluted honey. (It is an interesting fact that the honey fed to larvae in the colony is diluted by the nurse bees that do the feeding.)

Glucose oxidase is not commonly found in insects or in association with them. There is no reference to it in standard textbooks on insect physiology. It would be worthwhile to know how widely the enzyme is used by insects that collect or store food, including ants and solitary and semisocial bees.

Antibiotic activity is found in royal jelly, the rich creamlike substance worker bees secrete from glands in their head and feed to larvae. (Larvae that will become queen bees get more royal jelly than larvae destined to become workers; the difference is crucial, since queens and workers arise from the same eggs.) In 1959 Murray Blum and Arthur F. Novak at Louisiana State University and Stephen Taber III of the Department of Agriculture found that royal jelly contained a fatty acid with antibiotic properties. Although the antibiotic is only 20 to 25 percent as effective as penicillin and chlorotetracycline against the microorganisms exposed to it, its activity is still high enough to give the royal jelly considerable protection against bacteria.



BEE SPACE can be likened to corridors that bees maintain around and between combs. The space is seldom less than 1/4 inch or more than 3/8 inch wide. Larger spaces are filled with comb, and smaller ones are plugged with a varnish-like material called propolis (*lower right*), which apparently serves to fill cracks that might harbor microorganisms.

Another antibacterial agent that works for the bee colony is the natural acidity of honey. For many microorganisms an acid environment is inhospitable. Certain yeasts, however, will tolerate such an environment, and yeasts are found in honey. Because of the high osmotic pressure of honey, they are yeasts of the type called osmophilic, meaning that they live or thrive in a medium that has a high osmotic pressure.

Most yeasts, including the bread yeasts and the yeasts used to make beer and wine, belong to the genus *Saccharomyces;* they cannot grow or multiply in solutions that contain more than about 30 percent sugar. The osmophilic yeasts, belonging to the genus *Zygosaccharomyces,* will not grow in dilute sugar solutions. They flourish only in products such as maple syrup, which has a sugar concentration of about 66 percent, and in honey—provided that it contains more than the normal amount of water, which is about 19 percent.

Grading rules of the Department of Agriculture stipulate that Grade A honey must contain less than 18.6 percent water. Nectar that bees collect in the field usually contains between 10 and 50 percent sugar; the remaining content is mostly water. The bees "ripen" the nectar, reducing its moisture content both by warming the hive and by passing large volumes of air over the droplets of honey that have been put in comb cells. Further processing is visible among the house bees that receive nectar from the field bees; the house bees manipulate the nectar with their long tongue, swallowing it, regurgitating it, forming a droplet on the tip of the tongue and taking it into the honey stomach again. This repeated treatment of the honey drives off water and adds enzymes.

Fortunately for both bees and beekeepers normal honey containing less than 19 percent water does not ferment even though it contains yeasts. The high osmotic pressure and lack of water keep the yeasts from growing, although they remain alive. Fermentation in stored honey is occasionally a problem for bees and more often vexes the beekeeper after he has removed honey from the hive. Honey is hygroscopic: it absorbs moisture from the air. When stored honey crystallizes, as it usually does because it is a supersaturated sugar solution, a certain amount of water is incorporated into the crystal nuclei. The amount is proportionately less than is found in the honey overall, however, so that the liquid por-

tion of partly crystallized honey has a higher water content than the honey had before crystallization started. As soon as the moisture content of the uncrystallized portion rises above about 19 percent, the osmophilic yeasts begin to grow (unless they have been killed by heat treatment) and fermentation becomes evident.

In colonies with a large population of bees the high interior temperature helps to remove the excess moisture. In colonies that cannot protect all their stores, either because the stores are too large or because the bee population is too small, one can see bubbles of carbon dioxide forming as a result of yeast growth in the honey. The fermenting honey will leak from the cells and run down the surface of the comb, where it is likely to attract more moisture. At this point the glucose oxidase system offers a degree of protection, but mold may nonetheless appear on the dilute honey. Alcohol produced by the yeasts is attacked by *Acetobacter*, the organism that converts alcohol into acetic acid, or vinegar. The odor of the vinegar may attract flies, particularly fruit flies (*Drosophila*). Under these rare circumstances food stores within a bee colony may be destroyed because of the failure of the sanitation and protection system.

The honey sold in grocery stores is usually pasteurized to prevent fermentation. The treatment requires heating the honey to 140 degrees F. for 30 minutes or 160 degrees for one minute or some intermediate combination. Such heating also destroys the glucose oxidase; some people feel it further harms the honey, although the claim is questionable. Overheating can certainly damage the flavor of honey.

How do bees protect their stores of pollen? Again the mechanisms vary. One of them is the glucose oxidase system, which comes into play because bees add nectar or honey to pollen as they collect it.

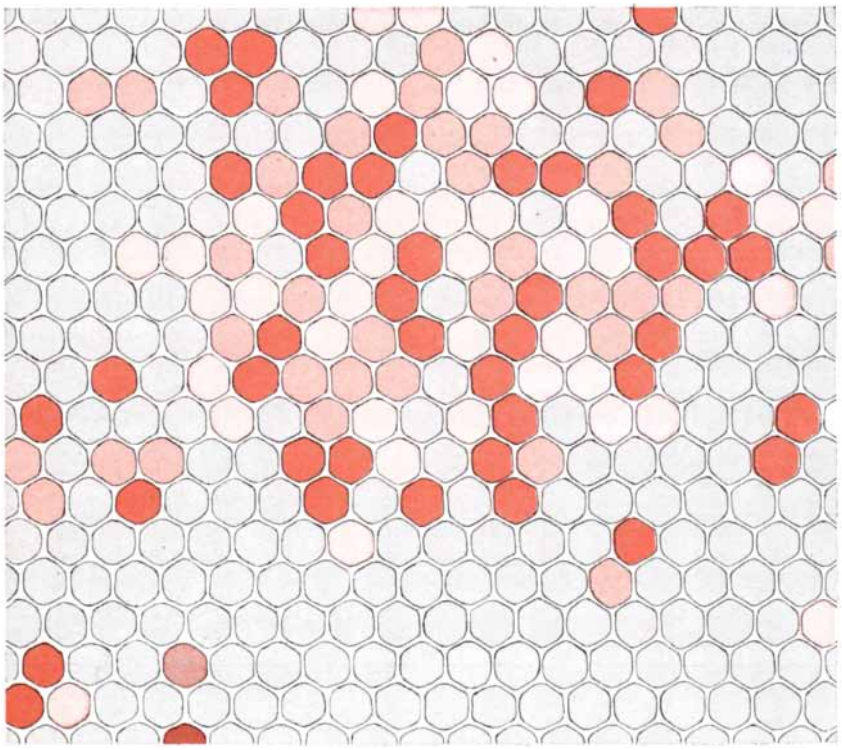
In 1966 Janine Pain and Jacques Maugenet of the French government's bee-research station at Bures-sur-Yvette near Paris found that the bacteria producing lactic acid (*Lactobacillus*) serve to protect pollen stored in the hive. The lactic acid makes the pollen mixture a natural ensilage, protecting the pollen against destruction by other microbial agents. Other bacteria (*Pseudomonas*) and ordinary yeasts (*Saccharomyces*) were also found in stored pollen. The effect of *Pseudomonas* is probably to remove oxygen, making the pollen a better growth medium for *Lactobacillus*. Yeast apparently serves to break down the pollen, making its use by the bees easier.

Although the honey cells in the comb are filled to the brim, the cells containing pollen are never filled to more than about 75 or 80 percent of their depth. When the pollen is being consumed rapidly, as it is in the early spring for brood-rearing, there is little danger of spoilage because the supply turns over quickly. If pollen is stored for any length of time, as it is in winter for the early spring activity, the pollen cells need further protection and are covered with honey and then with wax caps. That is why the pollen cells are only partly filled. The addition of honey and a cap helps to protect the pollen cell against contamination and bacterial spoilage.

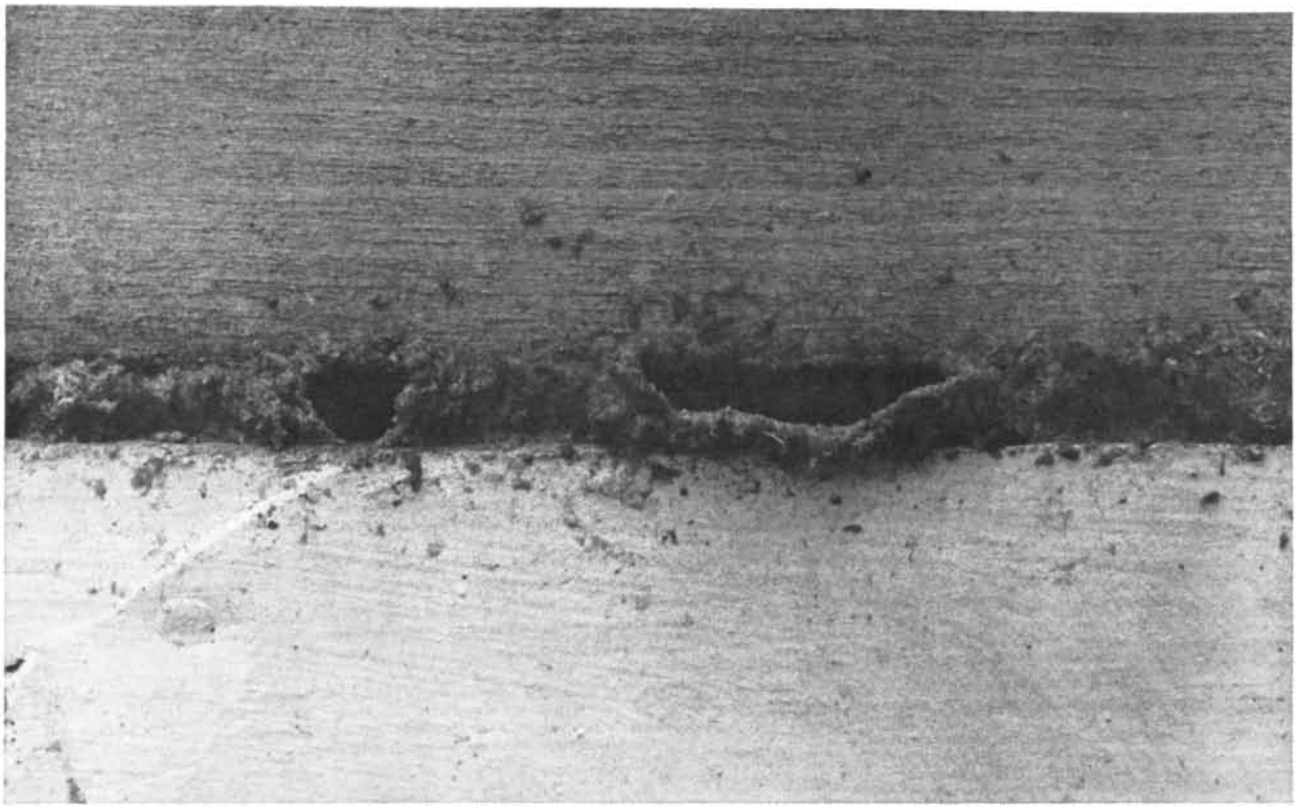
Another protective measure appears to lie in the fact that bees store honey and pollen by color: a given cell will have honey or pollen of a given color. (Since it is dark in the hive, one must assume that pollens and honeys of different colors also have different odors, which provide the basis for the bees' discrimination.) It seems probable that this separation of food from different sources might serve to protect the colony in the event that food from any one source had poor storage qualities.

It may be that human beekeepers, by failing to take sufficient advantage of the bees' ability to distinguish colors, have contributed to a malfunction of the bee colony's sanitation system. In an apiary the colonies may be only a foot or two apart, which is a situation that would not be found in nature. A returning bee may therefore enter the wrong hive. Such a bee is usually accepted, particularly if it is carrying a load of nectar or pollen. If the stray bee is performing a housecleaning task, however, having perhaps removed a dead larva or adult from its own hive, it may by entering the wrong hive spread the disease that killed the insect. It is believed American foulbrood, a bacterial disease of honeybee larvae, has been spread in this way. Since part of the sanitation system is that bees recognize their own hive, beekeepers are advised to paint the hives different colors and to provide landmarks such as small trees that would help bees return to the right hive.

If a bit of grass or straw is placed in a hive just under the cover, an observer watching the entrance will usually see a bee carrying the object out of the hive



STORAGE BY COLOR is apparently another protective device employed by honeybees. A given cell is filled with pollen or honey of a given color; stored pollen is shown here. Presumably the procedure serves to protect the colony by separating food from different sources in case the food from a particular source has poor storage qualities. Since the interior of a hive is usually dark, odor is probably the basis for the storage procedure.



PROPOLIS FILLS SPACE between two boards in a man-made hive. Propolis is obtained from gums and resins that bees gather from trees such as the pine and the poplar. After the material has been exposed to air in a hive it dries to form a hard surface.

within about five minutes. Moreover, if the object is small enough so that the bee can fly with it, the bee carrying it will take it 50 feet or more from the hive before dropping it. If the object is too heavy to be carried in flight, the bee will try to carry it at least several feet from the entrance by dragging it over the ground.

If one watches this housecleaning operation in an observation hive, which has glass sides, an interesting fact of bee behavior becomes apparent. The first bee to detect and pick up a foreign object is usually not the one that flies out of the entrance with it. My students at Cornell University, who do this exercise as part of a laboratory course, have observed that as many as five bees (the average is three) will participate in the removal of a single piece of debris, even though the distance from the place where the object is introduced to the entrance of the hive is seldom more than 12 inches. Usually only one bee at a time grasps the object. Occasionally a worker may actually carry it away from the entrance, since not all the workers are aware of where the entrance is.

Three points about the colony's housecleaning system emerge in a debris-removing action. The first is that most of the bees in a hive are quick to take action against a foreign object; any bee that encounters the object will carry it at least some distance away from where the bee is working. Second is the priority that worker bees give emergency tasks, such as removing a foreign object. Even if a bee is already occupied, it will switch at almost any time to an emergency task. Collectively the phenomenon is important to the hive in times of danger, when a sudden need may arise for guards, attackers, ventilators and so on. Third, foreign objects are not merely removed from the hive; they are carried some distance from it, so that if they represent a source of danger or infection, they are rendered unlikely to cause trouble.

Bees have still another sanitational mechanism involving the gums and resins they gather from trees such as the pine and the poplar. This material, which beekeepers call propolis, serves to protect the hive in several ways. One effect arises from the fact that the resins contain terpenes, which have both a bactericidal and a bacteriostatic action. In addition propolis is employed by the bees to put a hard coating over surfaces and objects that might create health hazards.

Plant resins, being gummy and sticky, are not easy for bees to collect. It is also a major operation for the house bees to remove propolis from the field bees that bring it into the hive. A field bee uses its mandibles and legs to work propolis into the pollen baskets on its hind legs. In the hive the field bee holds on to the comb or a part of the hive while several house bees pull long threads of propolis from the pollen baskets. After exposure to the air in the hive the propolis dries and forms a hard surface.

Large insects or small animals that enter a hive (either by mistake or in search of food) are usually killed by the bees. Such objects are too large for the bees to remove. Instead they encase them in propolis. It is not uncommon for beekeepers to find these entombed animals in a hive. The layer of propolis, usually at least a sixteenth of an inch thick, serves to remove the odor and, in effect, the dead animal.

If one puts a rough object, such as a piece of unsanded wood, into a hive, the bees will cover it with propolis and thus make its surface smooth. They also use propolis to close cracks in the hive wall. Indeed, they plug with propolis any space that is too small to serve as what beekeepers call "bee space," which can be likened to corridors that the bees maintain around and between combs to provide working room and walking

room. Bee space is from a quarter to three-eighths of an inch wide; anything larger is used to build comb, and anything smaller is closed with propolis.

One can only conjecture that the function of propolis as a space-filler is to eliminate cracks and crevices that might harbor bacteria, mold and small insects. Since the natural home of bees is usually a hollow tree, one can see how propolis might serve to protect the nest and make it more habitable. It has also been suggested that propolis protects the inside of the hive against water. Bees normally choose a dry place for a nest, however, and have been known to abandon nests that became wet.

The life-span of a worker bee during the summer is about six weeks. The death rate in a colony may therefore be more than 1,000 bees per day during the busy season, but if one checks the ground near the entrance to a colony, one will usually find no more than one or two bodies of newly dead bees per day. What happens in many cases is that worker bees die in the field. They do not retire or reduce their activity in old age; they simply keep working until they die. Laboratory examination has revealed that old bees may suffer from a variety of diseases, many of them infectious; the fact that they die away from the hive in such large numbers appears to be still another defense mechanism for the hive.

If a bee does die in the hive, the body is treated as a foreign object. Workers carry it to the entrance, and the worker that flies away with it carries it a considerable distance from the hive before dropping it. The tendency of the bees is not so much to keep the ground near the entrance to the hive clean (although if the entrance is near the ground, they do include the area around the entrance in their cleaning activity) as it is to conduct their affairs in such a way that the ground near the hive never gets dirty.

Ventilation figures importantly in the sanitation program of a colony. Bees ventilate the hive to cool it and also when they are collecting nectar in large amounts and need to evaporate water from exposed droplets of nectar in order to make it into honey of the right moisture content. The same ventilation system may be used to rid the hive of smoke or some other contamination in the air, provided that the contamination is solely internal.

Bees can be easily seen ventilating a hive on a warm day. The entrance will be crowded with bees, gripping the bottom board with their feet to remain stationary and fanning with their wings. In a standard hive air will be forced in one side of the entrance and out the other side. Within the hive additional fanners will be found moving the air around. It is possible to blow a little smoke into a hive through a hole made for the purpose and to observe it being forced out at the entrance or elsewhere by a ventilating response from the bees.

A major contribution to the cleanliness of the colony comes from the fact that bees do not void feces inside the hive. They do so only in flight and when they are some distance from the hive. As a result the colony sometimes confronts a problem in winter when a succession of cold days prevents the bees from flying out. If bees are confined too long, one bee may void its feces in the hive. In such a case other bees shortly do the same, and within a few hours the social order of the hive breaks down. In these circumstances the colony perishes within a few days. The need for an occasional warm day or part of a day—at least about half an hour—is critical for the survival of a honeybee colony in winter.

An unsolved question is what is done with the queen's feces. A queen takes flight only to mate when she is six to 12 days old; on rare occasions a queen will leave a hive to accompany a swarm to a new nest. The assumption is that worker bees remove the queen's feces. To my knowledge, however, no one has observed a queen voiding feces or seen a worker removing a queen's fecal matter.

I would not want to leave the impression that bees have no major disease problems. Honeybees suffer from a variety of bacterial, fungal, viral and protozoan diseases and also from certain pests. Contaminated water is occasionally a source of disease. Bees gather water from the nearest source, and if it is stagnant, it can cause illness in individual bees and can also contaminate the hive; bees often deposit droplets of water in the hive to assist in the cooling achieved by the ventilation system.

Bee diseases have been carefully studied, and methods of controlling them have been developed. Some methods work better than others. Men have complicated the problem of bee diseases by crowding colonies into apiaries. Nonetheless, it is usually possible to maintain the health of bees in apiaries by appropriate measures. The methods of the bee contribute significantly to the health of the colony. The honeybee is an example of an animal that has evolved good methods of protecting itself, its nest and its stored food from attack and damage by predators, parasites and microorganisms.

Four New Books

The Social Animal

ELLIOT ARONSON,
The University of Texas at Austin

"Aronson is one of the few people in our field who could write a book of this type . . . a succinct, enthusiastic, and quite accurate representation of social psychology."
—Ellen Berscheid, University of Minnesota

—A supplement for general psychology courses; an introductory text for social psychology courses; background reading for sociology and social science courses; a monograph for the researcher; enjoyable and enlightening reading for everyone.

To be published Summer 1972.
Approximately 150 pages, cloth $6.00 (tentative), paper $2.95 (tentative)

Physiological Psychology

Readings from **SCIENTIFIC AMERICAN**
With Introductions by
RICHARD F. THOMPSON,
University of California, Irvine

"Psychobiology represents the common ground shared by psychology and biology. As such, it is incorporated into both biology and psychology courses and programs. This reader is designed to meet both needs." —From the Preface

Forty-six readings for courses in physiological psychology or psychobiology, reproduced with full text, illustrations, and color.

To be published Spring 1972.
Cloth $12.00, paper $5.95

Perception: Mechanisms and Models

Readings from **SCIENTIFIC AMERICAN**
With Introductions by
RICHARD HELD and WHITMAN RICHARDS,
both of the Massachusetts Institute of Technology

"We have used this collection of reprints as the sole assigned reading material for several years in a course on Perception. . . . Student response has borne out our enthusiasm for these readings."
—From the Preface

Here are 39 readings, with comprehensive introductions, intended for courses in the psychology of perception.

To be published Spring 1972.

Altered States of Awareness

Readings from **SCIENTIFIC AMERICAN**
With Introductions by TIMOTHY J. TEYLER,
University of California, Irvine

These 14 readings provide excellent background in a significant area of psychobiology. The introductions review this material in the context of the latest information and discoveries. Useful supplementary reading for courses in introductory psychology, psychobiology, perception, cognition, physiological psychology, and other courses in behavioral and biological sciences.

To be published Spring 1972.

From your bookseller, or from Dept. M

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
58 Kings Road, Reading, England RG1 3AA

MATHEMATICAL GAMES

A topological problem with a fresh twist, and eight other new recreational puzzles

by Martin Gardner

The following nine problems, unrelated to one another, will be answered next month. Most of the problems were supplied by readers of this department and are published here for the first time.

1. Gustavus J. Simmons, in charge of research and development at Rolamite Inc., Albuquerque, N.M., sent this curious topological problem. Work at Rolamite involves complex banded rolling systems. One of the Rolamite engineers, Virgil Erbert, was confronted in the course of his work with the problem shown in the illustration below. End *A* of a flexible band was fastened to an object that was too large to pass through the slot at end *B*. It was essential that the band be formed into the looped configuration shown in the illustration without detaching end *A* from the object to which it was fastened. Can it be done?

It looks impossible, but the answer is yes. The reader is invited to draw a rough facsimile of the band on a sheet of paper, cut it out and tape end *A* to a tabletop. The puzzle, which is not difficult, is to manipulate the strip into the looped configuration.

2. Six players—call them *A*, *B*, *C*, *D*, *E* and *F*—sit around a circular table divided into six equal parts. At the center of the table is a disk mounted on a central pin around which it can rotate [*see top illustration on opposite page*]. The disk is marked with arrows and digits as indicated.

The wheel is spun five times. After each spin each player scores the number of points within his segment of the table. (If the wheel stops with its arrows exactly between adjacent players, the spin is not counted.) The players keep a running total of points, and the one with the largest total after the fifth spin is the winner. If there are ties for the highest score, no one wins and the game is played again.

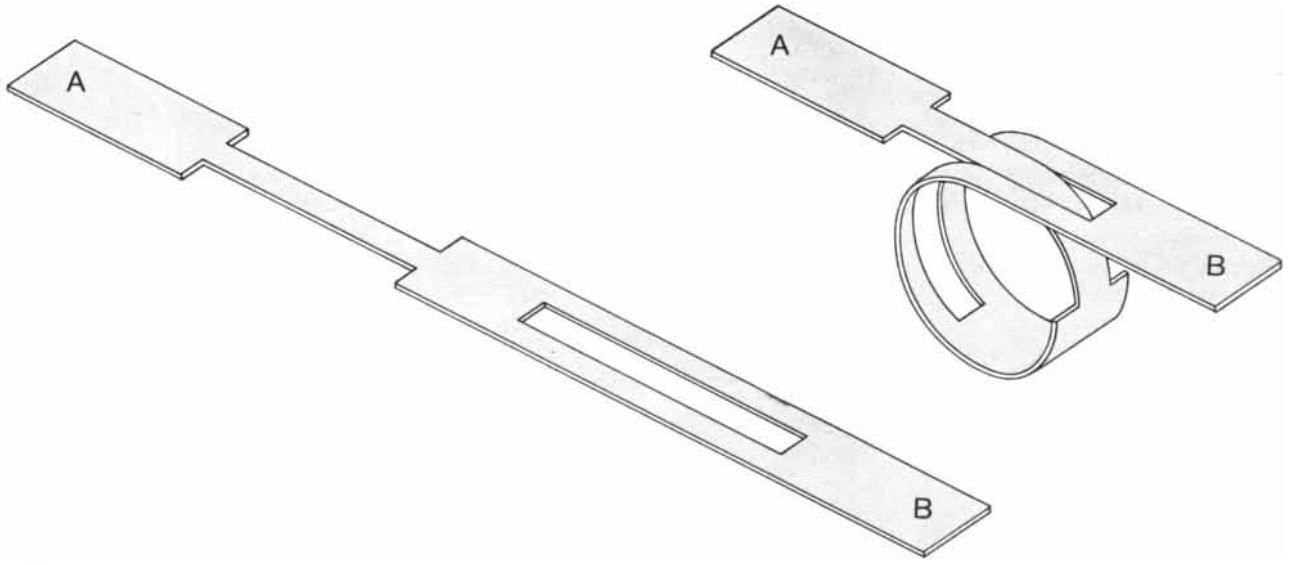
The outcome of the first spin is shown in the illustration. *C* is ahead with five points. After the second spin *D* is ahead. After the fifth spin *A* is the winner. What was each player's final score? The information seems to be insufficient, yet the question can be answered accurately by deductive reasoning. This unusual logic problem is adapted from a puzzle in one of D. St. P. Barnard's popular "Brain-twister" columns in the British *Observer*.

3. A frieze is a pattern that endlessly repeats itself along an infinite strip. Such patterns can exhibit different kinds of basic symmetry, but here we shall be concerned only with what is called "glide symmetry." A glide consists of a slide (more technically a "translation") combined with mirror reflection and a half-turn. For example, repeatedly gliding the letter *R* to the right along a strip generates the following frieze:

RЯRЯRЯRЯRЯRЯR...

H. S. M. Coxeter, a geometer at the University of Toronto, recently investigated in depth a remarkable class of frieze patterns that can be constructed very simply by using nonnegative integers, if the lack of symmetry in the shapes of the numerals is ignored [*see bottom illustration on opposite page*]. Think of the numerals as representing spots of colors, all 1's the same color, all 2's another color and so on. In this instance any rectangular portion of the frieze that is nine columns wide, such as the one shown in color, can be regarded as the unit pattern. By gliding it left or right—that is, sliding and simultaneously reflecting and inverting—the infinite frieze pattern is generated.

To produce this type of frieze pattern, begin with infinite borders of 0's and 1's at top and bottom, and a path of numbers from top to bottom such as the zigzag path of eight 1's shown on the left



A looped-band topological puzzle

between the borders of 0's. The numbers in such a path (which may be straight, or crooked as it is here), as well as the length of the path, can be varied to produce different patterns. A simple formation rule, common to all such patterns, is now applied to obtain all the other integers. The surprising glide symmetry that results is a nontrivial consequence of this rule.

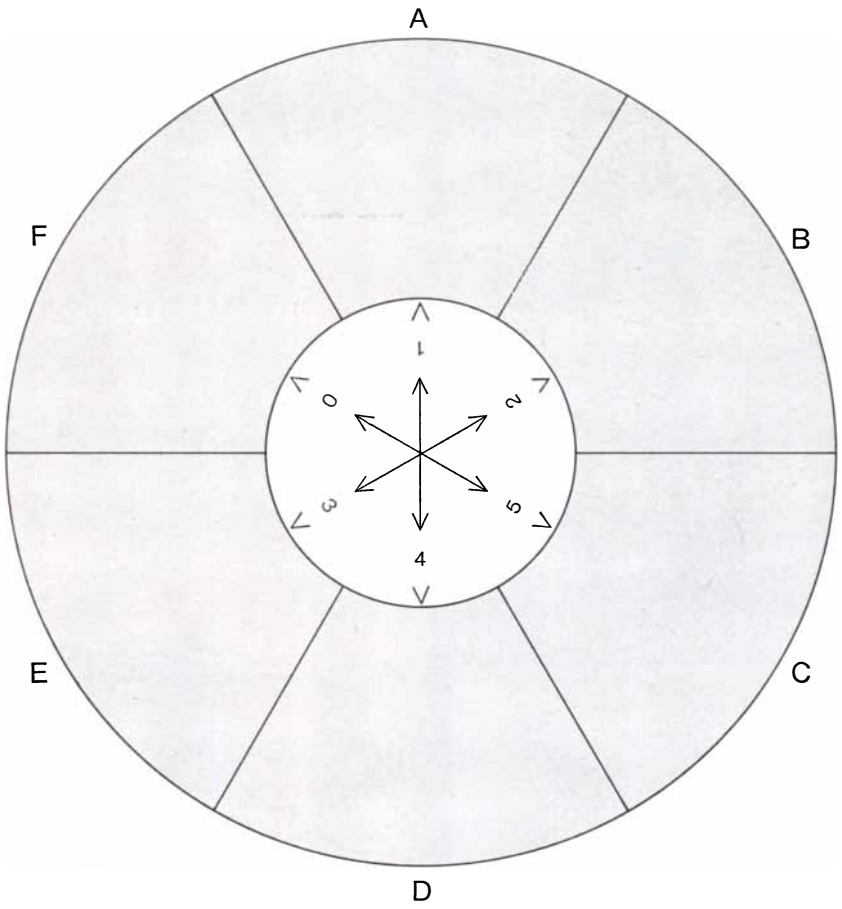
Our puzzle, suggested by Coxeter, is to guess the simple rule. Hint: It can be written as a linear equation with three terms involving nothing more than multiplication and addition. When Coxeter first showed the pattern given here to the mathematician Paul Erdös, Erdös guessed the rule in 20 seconds.

A discussion of the properties of such friezes, their fascinating historical background and their applications to determinants, continued fractions and geometry can be found in Coxeter's "Frieze Patterns" in *Acta Arithmetica,* Volume 18 (1971), pages 297–310. On friezes in general and their seven basic kinds of symmetry see Coxeter's modern classic, *Introduction to Geometry* (Wiley, 1961), pages 47–49.

4. On a picnic not long ago Walter van B. Roberts of Princeton, N.J., was handed a freshly opened can of beer. "I started to put it down," he writes, "but the ground was not level and I thought it would be well to drink some of the beer first in order to lower the center of gravity. Since the can is cylindrical, obviously the center of gravity is at the center of a full can and will go down as the beer level is decreased. When the can is empty, however, the center of gravity is back at the center. There must therefore be a point at which the center of gravity is lowest."

Knowing the weight of an empty can and its weight when filled, how can one determine what level of beer in an upright can will move the center of gravity to its lowest possible point? When Roberts and his friends worked on this problem, they found themselves involved with calculus: expressing the height of the center of gravity as a function of the height of the beer, differentiating, equating to zero and solving for the minimum value of the height of the center of gravity. Later Roberts thought of an easy way to solve the problem without calculus. Indeed, the solution is simple enough to get in one's head.

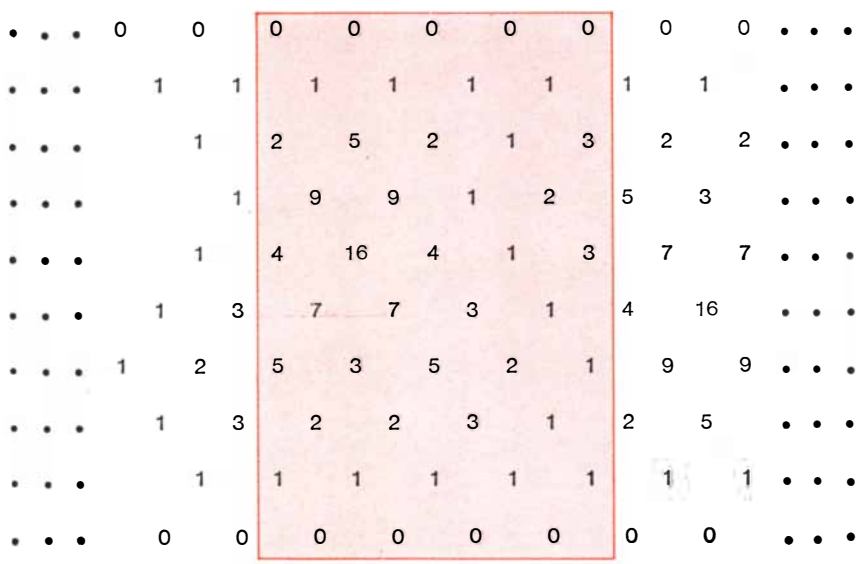
To devise a precise problem assume that the empty can weighs 1½ ounces. It is a perfect cylinder and any asymmetry introduced by punching holes in the top is disregarded. The can holds 12 ounces of beer, therefore its total weight, when filled, is 13½ ounces. The can is eight inches high. Without using calculus determine the level of the beer at which the center of gravity is at its lowest point.



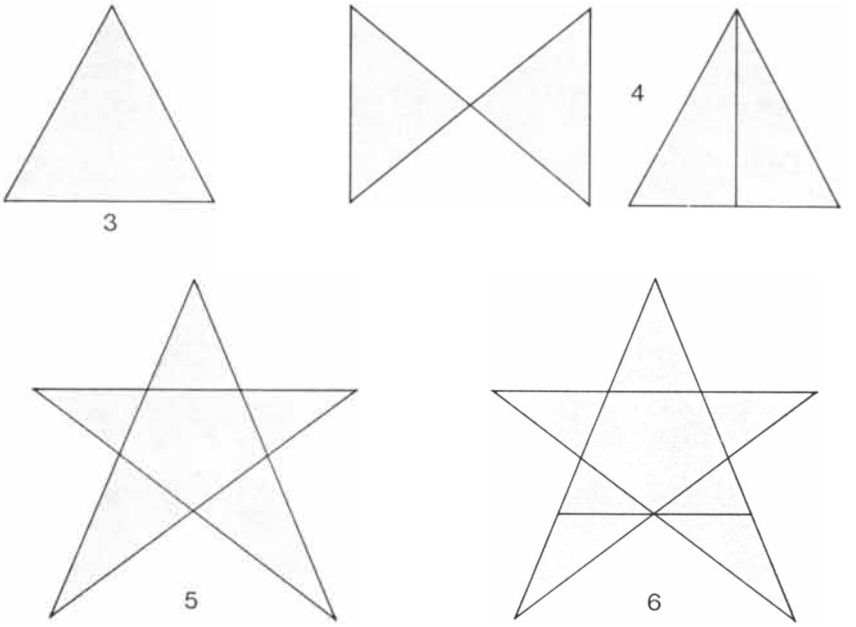
D. St. P. Barnard's game problem

5. Three coins are on a table: a quarter, a half-dollar and a silver dollar. Smith takes one coin and Jones takes the other two. All three coins are tossed simultaneously.

It is agreed that any coin falling tails counts zero. Any coin falling heads counts its value in cents. The tosser who



A frieze pattern with glide symmetry



Maximum number of nonoverlapping triangles for three, four, five and six lines

gets the larger score wins all three coins. If all three come up tails, no one wins and the toss is repeated.

What coin should Smith choose so that the game is fair, that is, so that the expected monetary win for each player is zero?

David L. Silverman, author of the excellent book of game puzzles called *Your Move* (McGraw-Hill, 1971), is responsible for this new and unpublished problem. It has an amazing answer. Even more astonishing is a generalization, formally proved by Benjamin L. Schwartz of Mathematica Inc. of Bethesda, Md., of which this problem is a special case.

6. Kobon Fujimura, a Japanese puzzle expert, recently invented a problem in combinatorial geometry. It is simple to state, but no general solution has yet been found. What is the largest number of nonoverlapping triangles that can be produced by n straight line segments?

It is not hard to discover by trial and error that for $n = 3, 4, 5$ and 6 the maximum number of triangles is respectively one, two, five and seven [*see illustration above*]. For seven lines the problem is no longer easy. The reader is asked to search for the maximum number of nonoverlapping triangles that can be produced by seven, eight and nine lines.

The problem of finding a formula for the maximum number of triangles as a function of the number of lines appears to be extremely difficult.

7. One of the satisfactions of recreational mathematics comes from finding better solutions for problems thought to have been already solved in the best possible way. Consider the following digital problem that appears as Number 81 in Henry Ernest Dudeney's *Amusements in Mathematics*. (There is a Dover reprint of this 1917 book.) Nine digits (0 is excluded) are arranged in two groups. On the left a three-digit number is to be multiplied by a two-digit number. On the right both numbers have two digits each:

$$\begin{array}{rr} 158 & 79 \\ 23 & 46 \end{array}$$

In each case the product is the same: 3,634. How, Dudeney asked, can the same nine digits be arranged in the same pattern to produce as large a product as possible, and a product that is identical in both cases? Dudeney's answer, which he said "is not to be found without the exercise of some judgment and patience," was

$$\begin{array}{rr} 174 & 96 \\ 32 & 58 \\ \hline 5{,}568 & 5{,}568 \end{array}$$

Victor Meally of Dublin County in Ireland later greatly improved on Dudeney's answer with

$$\begin{array}{rr} 584 & 96 \\ 12 & 73 \\ \hline 7{,}008 & 7{,}008 \end{array}$$

This remained the record until last

year, when a Japanese friend of Fujimura's found an even better solution. It is believed, although it has not yet been proved, to give the highest possible product. Can the reader find it without the aid of a computer?

8. A well-known problem with checkers is begun by placing eight checkers in a row. A move consists in picking up a checker, carrying it right or left over exactly two checkers, then placing it on a checker to make a king. (Carrying a checker over a king counts as moving it over two checkers.) In four moves form four kings. The problem is not difficult, and it is easy to show that for any even number of checkers, n, when n is at least 8, a row of $n/2$ kings can always be produced in $n/2$ moves.

Numerous variants on this old problem have been proposed by Dudeney and other puzzle inventors. The following variation on the theme, which I believe is new, was suggested and solved by W. Lloyd Milligan of Columbia, S.C.

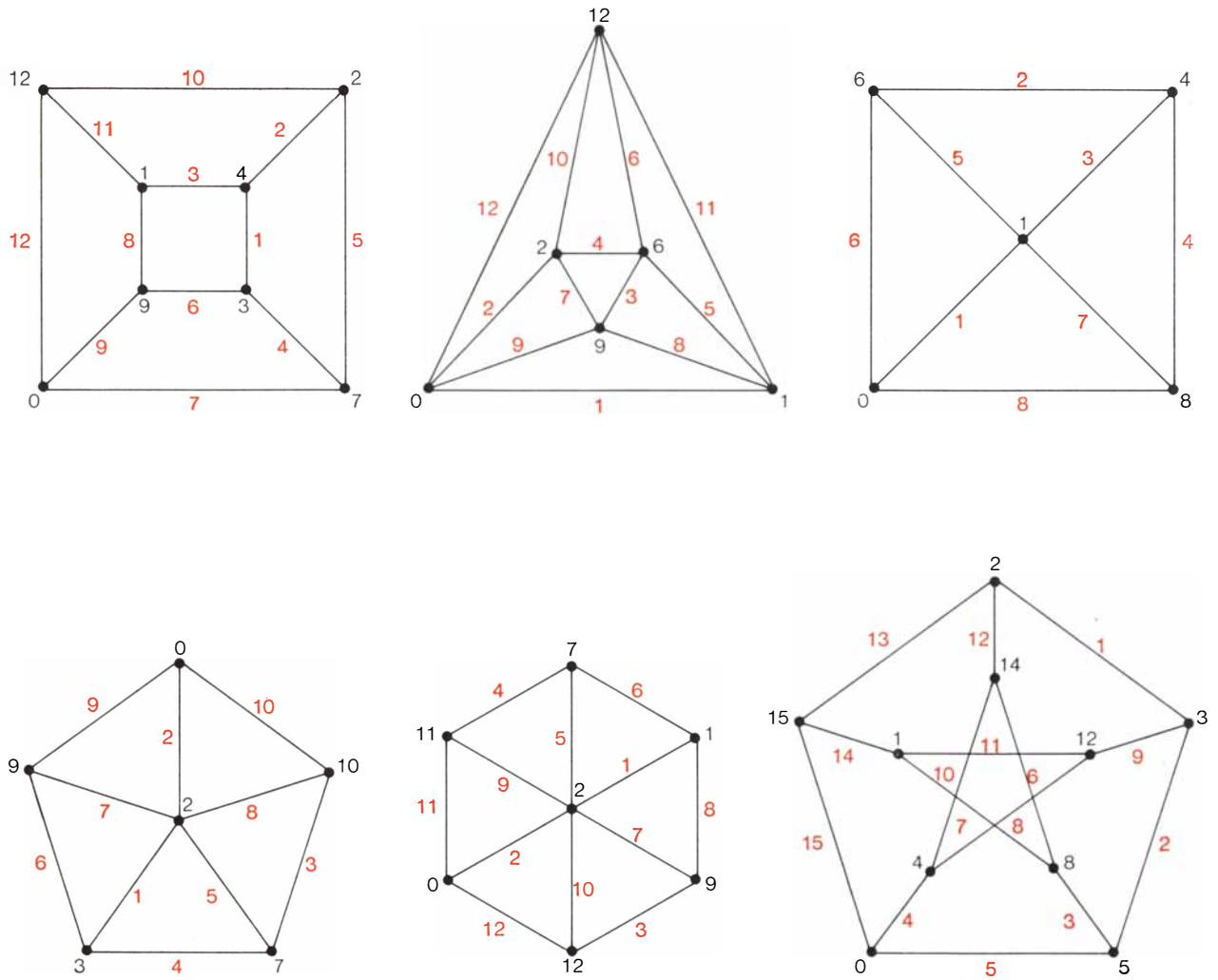
An even number of checkers, n, are placed in a row. First move a checker over one checker to make a king, then move a checker over two checkers, then a checker over three checkers, and so on, each time increasing by one the number of checkers to be passed over. The objective is to form $n/2$ kings in $n/2$ moves.

Can the reader prove that the problem cannot be solved unless n is a multiple of 4, and give a simple algorithm (procedure) for obtaining a solution in all cases where n is a multiple of 4? A solution is easily found by trial and error when n is 4 or 8, but for $n = 16$ or higher it is not so easy without a systematic method.

9. Single-panel gag cartoons, like Irish bulls, are often based on outrageous logical or physical impossibilities. Lewis Carroll liked to tell about a man who had such big feet that he had to put his pants on over his head. Almost the same kind of impossibility is the basis of a famous *New Yorker* cartoon by Charles Addams of a woman skier going down a slope. Behind her you see her parallel ski tracks approaching a tree, going around the tree with a track on each side and then becoming parallel again.

Suppose you came on a pair of such ski tracks on a snowy slope, going around a tree exactly as in Addams' cartoon. Assume that they are, in truth, tracks made by skis. Can you think of at least six explanations that are physically possible?

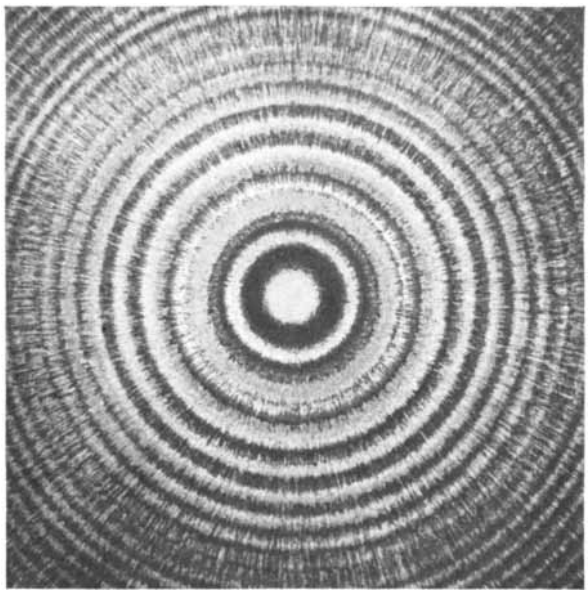
Solutions to the six graphs that readers were asked last month to number "gracefully" are shown in the illustrations below. None of these numberings is unique. If any readers succeed in gracefully numbering the skeletons of the dodecahedron and icosahedron, I shall report it later.



Solutions to last month's graceful-graph problems

"LASER LIGHT"

A SCIENTIFIC AMERICAN FILM



Almost everyone has heard of lasers, but relatively few people have seen them in action. The Editors of SCIENTIFIC AMERICAN now present "LASER LIGHT," a 16-millimeter sound film about lasers: what they are, how they work, the marvelously pure and curiously scintillating light they produce, how they are being used and how they may be used in the near future. The film is in color and lasts 37½ minutes. It is now available for sale or rent.

A few highlights of the film are:

- **Computer-generated animation explaining stimulated emission and resonant optical cavities.**
- **Ripple-tank and oscilloscope demonstrations explaining the wave principles underlying laser action and holography.**
- **Holograms, their three-dimensionality dramatically evoked by the moving camera.**
- **A 600-foot, 8.8-kilowatt laser in action.**
- **Tunable lasers.**
- **A television picture transmitted by laser beam.**
- **The laser chalkline for the San Francisco Bay tunnel.**
- **Laser interferometry.**
- **Gas, solid and organic-liquid lasers.**
- **An experiment on the use of holography in a computer memory.**
- **Original musical score.**

"LASER LIGHT" is recommended for general audiences with an interest in science and technology, and for use in conjunction with the teaching of physics and optics. The film is accompanied by a selection of five SCIENTIFIC AMERICAN articles on lasers and holography, written by leading authorities in these fields.

The sale price per print is $375, the rental price $37.50 for a booking of three days. Write Motion Picture Department SA, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, New York 10017. Telephone (212) 585-5319. Cable SCIAMER. In Canada write: International Tele-Film Enterprises, 221 Victoria Street, Toronto 205. Telephone (416) 362-2321. In Great Britain write: The British Film Institute, 81 Dean Street, London WIV 6AA. Telephone 01-437-4355.

THE AMATEUR SCIENTIST

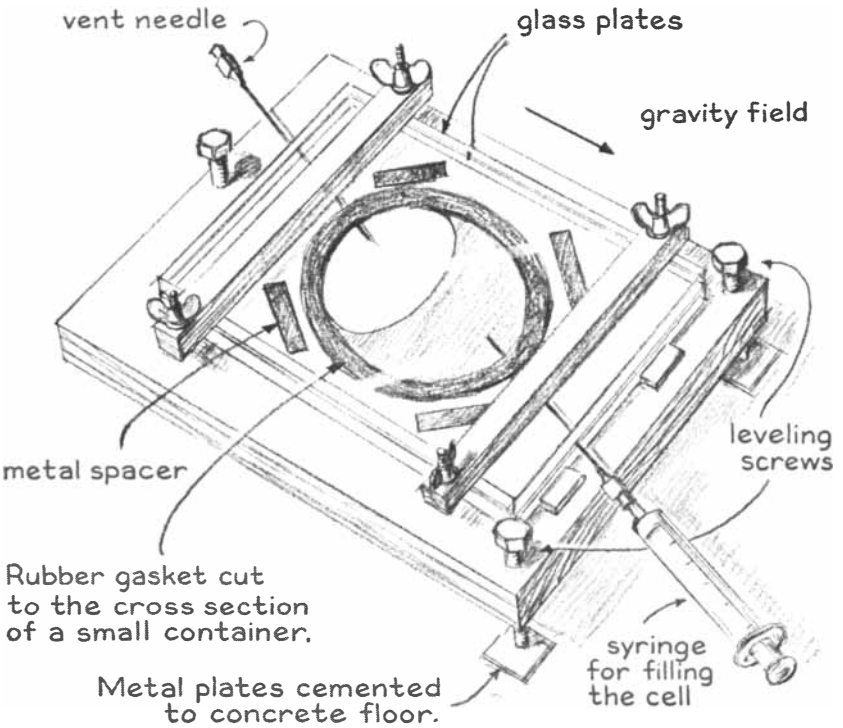
What happens to the fluid in the tanks of a spacecraft falling freely in space?

Conducted by C. L. Stong

An automotive engineer designing a fuel tank is not likely to give much thought to the effects of surface tension. He knows from experience that gravity pulls gasoline to the bottom of the tank, where it flows into the fuel line. He ignores the meniscus pulled upward by surface tension around the wall of the tank where the fluid wets the metal. In tanks of conventional size the total upward force of surface tension amounts to less than a quarter of an ounce. It is easily overcome by the weight of the gasoline.

In contrast, surface tension is a major concern of engineers who design space vehicles for operation in gravitational fields that range from 1 *g* (the force of gravity at the earth's surface) to the weightless state. As the force of gravity approaches zero, fluid in the tanks of space vehicles tends to creep over all accessible surfaces, creating a bubble of vapor that can drift randomly through the fluid. The surface configuration of the interface between the fluid and the vapor must be controlled by forces other than gravity if the tank is to be vented and fuel is to be reliably fed to the engine.

Various schemes have been devised for investigating the behavior of propellants in tanks subjected to reduced gravity. For example, fluid in model tanks made of transparent materials has been photographed in aircraft flown on zero-gravity trajectories. Models have also been dropped from towers and photographed in free fall by high-speed flash lamps.



William Olsen's reduced-gravity simulator

An interesting device for investigating how surface tension alters the shape of fluid in a closed container is the reduced-gravity simulator developed at the Lewis Research Center of the National Aeronautics and Space Administration. This apparatus enables the investigator to observe in two dimensions the successive changes in the shape of a fluid as gravity is reduced. A simplified version of the simulator that amateurs can make was suggested by William Olsen of 209 Inwood Boulevard, Avon Lake, Ohio 44012, who developed the simulator at the Lewis Center. Armed with Olsen's description and several NASA research papers on zero gravity, I made a simulator and conducted various experiments with it.

The simulator consists essentially of a pair of closely spaced glass plates separated by rubber gaskets that conform to the boundary of a selected cross section of the tank [*see illustration on this page*]. The space enclosed by the gaskets is partly filled with ethyl alcohol, which serves as a two-dimensional model of the fuel. The angle the sandwich makes with respect to the horizontal can be varied from 90 degrees to 0 degrees. By changes in the angle the two-dimensional model can in effect be subjected to gravitational acceleration that ranges from 1 *g* to almost zero. The influence of gravity on the shape of the interface diminishes as the horizontal position is approached, and the influence of surface tension increases. When the assembly lies in the horizontal plane, the interface assumes the shape it would take if the container were in free fall.

Essentially the simulator is an application of an apparatus devised about 1900 by H. S. Hele-Shaw of University College in Liverpool and subsequently developed by A. D. Moore, now professor emeritus of electrical engineering at the University of Michigan, for simulating potential fields [see "The Amateur Scientist"; SCIENTIFIC AMERICAN, July, 1967]. Both mathematical and experimental demonstrations have established that the behavior of ethyl alcohol in the simulator closely approximates the be-

havior of propellants such as liquid hydrogen in full-scale, three-dimensional tanks that are subjected to comparable gravitational fields. The simulator also enables the experimenter to investigate techniques for controlling the position the liquid and vapor occupy in the tank during free fall so that propellant can be fed to the engine when desired and vapor can be vented continuously with minimum loss of liquid.

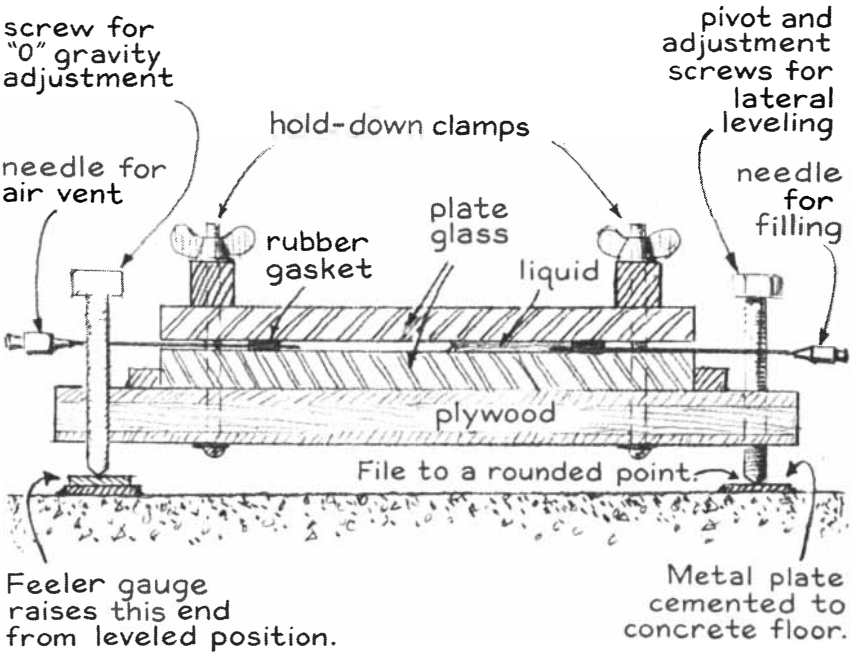
The assembly includes two pieces of plate glass at least 3/8 inch thick that can be as large as eight inches wide and 10 inches long. The glass should be as flat as possible and free of scratches. One plate is supported by a plywood base about 12 inches wide, 14 inches long and 3/4 inch thick [*see top illustration on this page*]. The glass can be held in position by strips of wood glued to the base.

The spacing between the sheets of glass is fixed by four strips of sheet steel approximately 1/2 inch wide, two inches long and .05 inch thick. The strips can be made of steel shim stock that is available from dealers in automobile supplies. When you cut the metal, take care to preserve its flatness. File all burrs from the edges. The four strips must be of identical thickness, the object being to create a space of uniform thickness between parallel glass plates. The four spacers are placed at the corners of the lower plate.

The boundary of the model tank to be investigated is cut from soft rubber sheeting, preferably neoprene, that is a few thousandths of an inch thicker than the spacers. A spherical tank is represented in two dimensions by a circular gasket, a cylindrical tank by a rectangle and so on. The rubber model is centered on the lower plate and covered by the upper plate. Before assembly the glass and rubber must be thoroughly cleaned with detergent and water followed by a rinse with alcohol.

The sandwich is clamped together. A pair of adjustment screws are added to the base at one end to serve as pivots on which the apparatus tilts. A third adjustment screw at the center of the opposite end controls the angle of tilt. In operation the apparatus rests on three pads (small rectangles of metal) cemented to a solid, vibrationless, horizontal foundation such as a concrete floor.

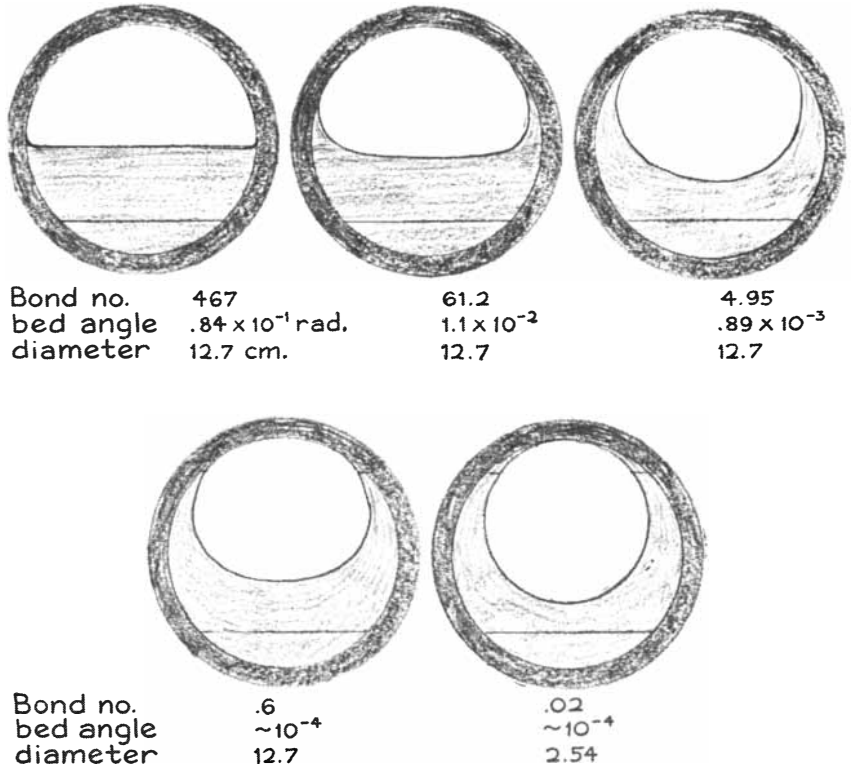
To conduct an experiment place the apparatus on its pads. Insert a fine hypodermic needle through the rubber model at the top, which is the end closest to the adjustment screw that controls the tilt. This needle serves as a vent. A similar needle at the bottom admits fluid to the model.



Elements of the simulator as seen from the side

With a hypodermic syringe completely fill the model with alcohol. I use 200-proof ethyl alcohol. Rubbing alcohol should work almost as well, but I have not tried it. After the model is filled inject a small bubble of air, approximately four millimeters in diameter, by sucking alcohol into the syringe. Turn the pivot (lateral-adjustment) screws as required to center the bubble at either the top or the bottom of the model. Turn the tilt adjustment screw as required to cause the bubble to drift slowly from the top to the bottom, or vice versa.

The direction of the gravity field should make a right angle with respect



Influence of reduced gravity on the surface configuration of a fluid

to the line on which the apparatus pivots. When the path of the bubble is properly adjusted, it is parallel to the direction of the gravitational field. If it is not, adjust one of the pivot screws as required to correct the angle of the bubble's path.

Finally, tilt the apparatus just enough to cause the bubble to drift slowly into contact with the tip of the vent needle. By operating the hypodermic syringe remove fluid until the bubble expands to a diameter of one centimeter. Adjust the tilt screw to cause the expanded bubble to drift to the bottom of the model. Next turn the tilt screw to the point at which the bubble drifts toward the top at the rate of .01 centimeter per second (one centimeter per 100 seconds) or slower. When the apparatus is so adjusted, the influence of gravity on the shape of the fluid has been reduced to close to zero.

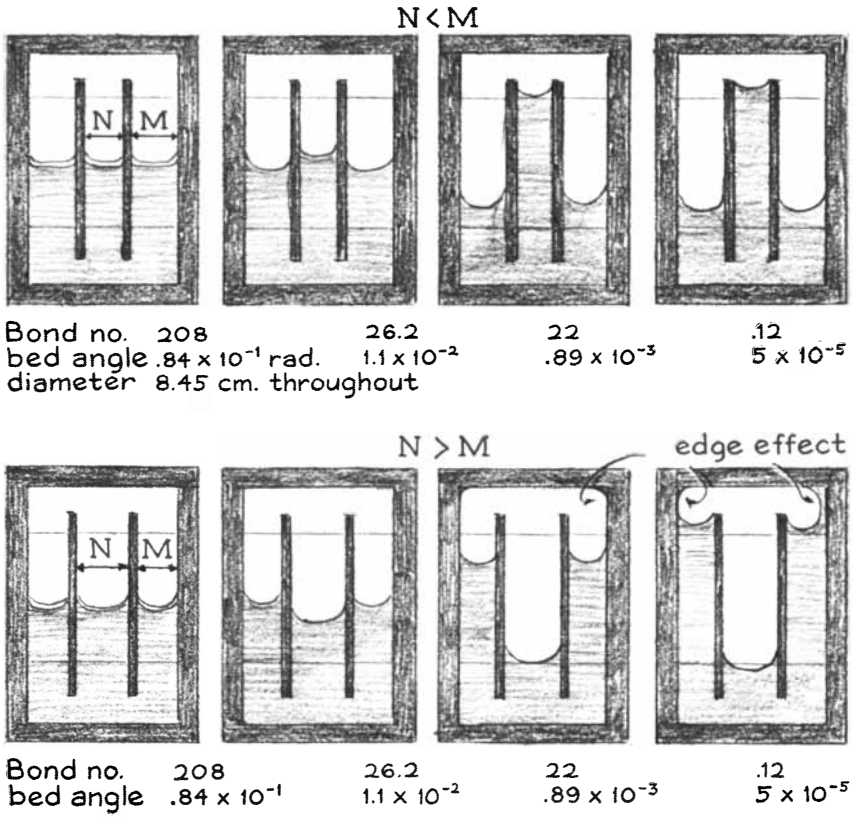
The simulator can now be used for duplicating many of the experiments done with models in drop towers and in aircraft that are flown on zero-gravity trajectories. Remove some of the alcohol, thus making a larger bubble. Tilt the simulator to the vertical plane to observe the shape of the interface at 1 g and return it to the horizontal plane to examine the configuration in free fall.

The simulator can also be used for investigating interface configurations that lie between these two extremes. Most interesting are those that appear when the apparatus is adjusted to small angles with respect to the horizontal plane. All inclinations are usefully described by a quantity known as the Bond number (for W. N. Bond, who investigated capillary effects some 40 years ago); it expresses the ratio of the gravitational force to the surface-tension forces. This quantity can be calculated by timing the velocity at which a bubble moves through the fluid. One must take into account the nature of the fluid, the size of the bubble, the size of the model and the inclination of the apparatus.

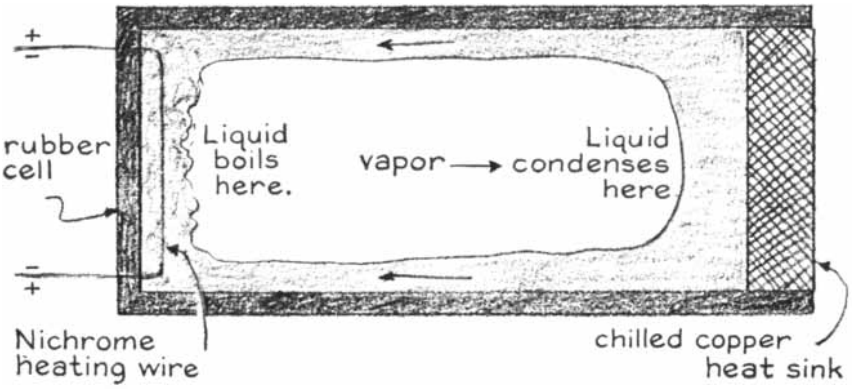
At low inclinations the Bond number of simulators that contain ethyl alcohol is expressed by the equation $B = .69\,VW^2$, in which B is the Bond number, V is the velocity in centimeters per second of a bubble one centimeter in diameter and W is the diameter (in centimeters) of the model tank. Assume that the experimenter wishes to investigate the interface at an inclination such that the force of surface tension is 10 times that of gravity (a Bond number of .1) in a model 3.8 centimeters wide. The corresponding velocity of a bubble one centimeter in diameter would be $V = B/.69\ W^2 = .1/.69 \times 3.8 \times 3.8 = .01$ centimeter per second, or one centimeter in 100 seconds.

In general Bond numbers smaller than 1 indicate that the force of surface tension predominates. Numbers larger than 1 indicate that the force of gravity predominates. The influence of surface tension on the shape of the interface is decreasingly apparent at Bond numbers ranging from 5 to 50. At Bond numbers of 300 and higher the interface is substantially the same as it is at 1 g.

Having adjusted the apparatus for a Bond number of .1 or less, the experimenter can observe phenomena at higher Bond numbers by inserting spacers of selected thickness between the tilt adjustment screw and the metal on which it rests. It is interesting to observe the shape of the interface through a series of Bond numbers as fluid is withdrawn from the model. One can also get an idea of the effects of acceleration on the distribution of a propellant by abruptly altering the inclination of the simulator. For example, increasing the Bond number from .1 to 100 during an interval of a few seconds is approximately equivalent in its effect on the distribution of



Influence of baffles in a cylindrical tank



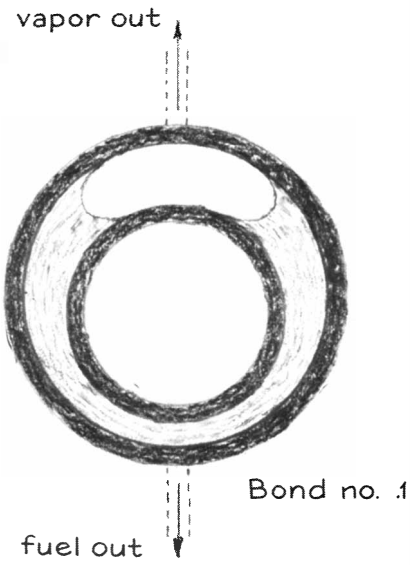
Effect of temperature gradient on a fluid in free fall

fuel to starting the engine of a space vehicle that is in orbit.

To ensure reliable feeding of fuel to a rocket engine, why not store the fluid in an expandable tank, such as a rubber bag? This expedient might work for some fluids. It would not solve the problem of venting, which becomes a major concern in the case of volatile propellants such as liquid hydrogen. Moreover, liquid hydrogen boils at −423 degrees Fahrenheit. At this temperature most organic materials become as brittle as glass. The designer must rely on forces other than stretched rubber to keep the fluid and vapor where they are wanted—forces such as surface tension, inertia, electrostatic fields and heat.

The position at which the fuel tends to collect can be altered by modifying the shape of the tank and installing systems of baffles. For example, fluid in a spherical tank that contains a smaller sphere, which is offset with respect to the center of the tank, tends to collect in the narrowest zone between the inner and the outer sphere [*see illustration below*]. Simultaneously the vapor tends to remain in the widest zone of separation, where the bubble can most closely approach the shape of a sphere: the shape of minimum surface.

The work that must be expended to move the fluid involves the concept of surface energy, which is related to the concept of surface tension. Surface tension is defined as the force a fluid exerts perpendicularly to a line that lies in the surface. It can be likened to the force exerted by a stretched sheet of rubber. Work expended in stretching the "rub-



Baffle for a spherical tank

ber" is conserved in the form of surface energy. Liquids tend to assume shapes of minimum surface and hence of minimum surface energy. It is for this reason that rain falls in the form of spherical drops and that vapor in the spherical tank tends to assume the shape of a spherical bubble, thereby shifting the fluid to the opposite side of the tank.

Baffles can similarly be used to control the position of fluid in a cylindrical tank [*see upper illustration on page 108*]. For example, at low Bond numbers fluid in a cylindrical tank tends to collect inside a coaxial tube if the diameter of the tube is less than one-third of the diameter of the cylinder, as in the upper part of the illustration. In effect the tube fills with fluid by capillary attraction.

The tube can be welded to the bottom of the tank and perforated near the weld to admit fluid to its interior. Fuel can be withdrawn through an opening in the bottom of the tank that is coaxial with the tube. Vapor can be similarly vented at the top of the tubular baffle. If the diameter of the tube is less than one-third of the diameter of the tank, vapor tends to fill the tube at low Bond numbers, as in the lower part of the illustration.

The principle of dielectrophoresis has been suggested as a technique for positioning liquid and vapor in a propellant tank. When liquid that is a poor conductor of electricity is placed in a nonuniform field, it tends to migrate toward the region of highest field strength. For example, if the charged electrode at one end of a cylindrical tank consists of a disk and the electrode at the opposite end consists of a sharp point, liquid would migrate toward the point.

In the simulator the disk could be represented by a length of straight wire at one end of a rectangular rubber gasket. The tip of a hypodermic needle thrust through the middle of the opposite end would serve as the pointed electrode. A suitable dielectric fluid is available in most households in the form of Carbona (carbon tetrachloride), a highly toxic cleaning agent. Use it only in a well-ventilated room and avoid inhaling the vapor. Potential for creating the electric field can be derived from a small electrostatic generator.

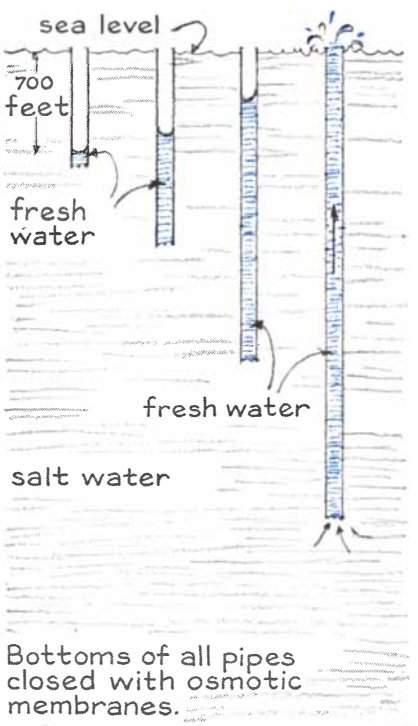
Another physical effect that invites experimentation is the dependence of surface tension on temperature. Surface tension varies inversely with the temperature of the liquid and vanishes at the boiling point. Temperature gradients can exist in fluids in free fall because convection currents largely disappear in the absence of gravity.

A bubble of vapor will move through the temperature gradient. Surface tension is lower on the warmer side of the bubble. Liquid at the warmer interface flows to the region of lower temperature, where surface tension is higher. Simultaneously the bubble drifts from the cooler side to the warmer side.

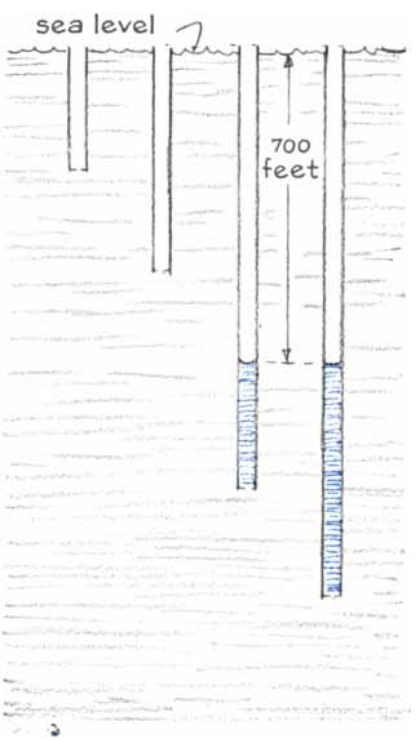
This effect can be investigated in the simulator by suspending a short length of Nichrome wire in a model tank and heating it with electric current from a No. 6 dry cell. The opposite end of the model can be a strip of sheet copper that can be cooled with a piece of ice. In addition to demonstrating how a fluid can be moved by variations in temperature the apparatus can also be made to function as a two-dimensional heat pipe [*see bottom illustration on page 108*].

Numerous other experiments can be undertaken with the simulator, including the observation of currents induced at low Bond numbers by inertial forces, the diffusion of two liquids at an interface and so on. These effects can be observed by placing a few particles of a coloring substance such as potassium permanganate on the lower plate of glass or injecting solutions of differing colors into the model tank.

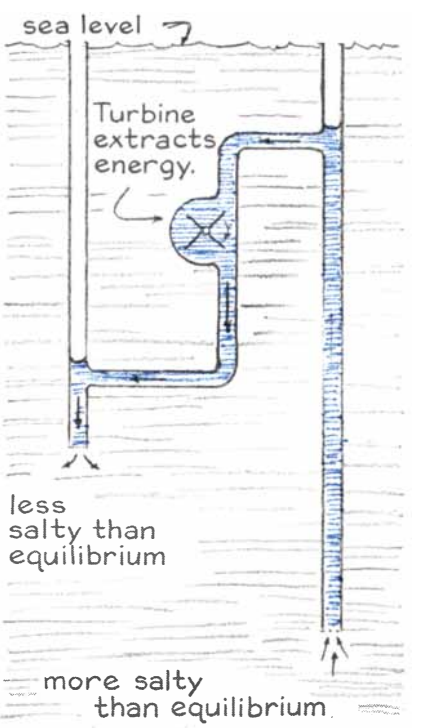
Last December, Octave Levenspiel, who teaches chemical engineering at Oregon State University, described in this department an osmotic pump for creat-



Osmotic-fountain hypothesis



The true situation



Fanciful pump scheme

ing a continuous fountain of fresh water in the ocean. The proposed pump would consist of a vertical pipe closed at the bottom by a semipermeable membrane. At a depth of 700 feet or more external pressure becomes sufficient to induce reverse osmotic flow and fresh water enters the pipe through the membrane.

Levenspiel suggested that in a pipe extending to a depth of five miles the difference in density between salt water outside the pipe and fresh water inside might conceivably develop an additional pressure difference sufficient to lift fresh water above the surface of the ocean, creating a perpetual fountain. He agreed to submit further comments after readers were given an opportunity to appraise the merits of the exotic invention. Many readers responded, including a number of specialists in thermodynamics. I now join Levenspiel in expressing gratitude to the respondents for explaining why man must look to other sources for fresh water. Levenspiel writes as follows:

"When a pipe that is closed at the bottom by a semipermeable membrane is lowered into the ocean beyond the critical depth of about 700 feet, fresh water inside the pipe either rises higher as the pipe goes deeper, or it rises only to the critical depth. If the fresh water were to rise higher as the pipe goes deeper, one could construct the machine I have described [*see illustration at right on opposite page*]. In this machine fresh water would run downhill from one pipe to the other through a turbine that would generate electricity. Simultaneously the machine would partly unmix the salty ocean. Brine near the inlet of the deeper membrane would become increasingly salty; brine in the vicinity of the shallower membrane would become less salty. Clearly this action would violate the second law of thermodynamics, which states in effect that natural events proceed without exception in a preferred direction; that substances when left to themselves may mix but will never unmix; that man grows older, never younger, and that a rolling stone eventually stops rolling.

"Since the assumption of different water levels leads to this absurd result, one must conclude that the water level in the pipe stays unchanged no matter how far the pipe is lowered. If the ocean is not in equilibrium, in either temperature or salinity, the chimney effect could elevate fresh water a little above the critical depth of 700 feet, but not nearly enough to reach the surface. We must consign the osmotic fountain to the historical junk heap of perpetual motion machines."

EINSTEIN...BORN...DOBZHANSKY...FLINT... ASIMOV...DeROPP...MENNINGER...HEISENBERG

Add a few giants to your library. Great time. Great offer.

43990. EINSTEIN: THE LIFE AND TIMES. *Ronald W. Clark.* The definitive work — careful, complete, awe-inspiring. **$15.00**

36770. THE BORN-EINSTEIN LETTERS. The correspondence between Albert Einstein and Max and Hedwig Born, 1916-1955. With commentaries by Max Born. Translated by Irene Born. **$8.50**

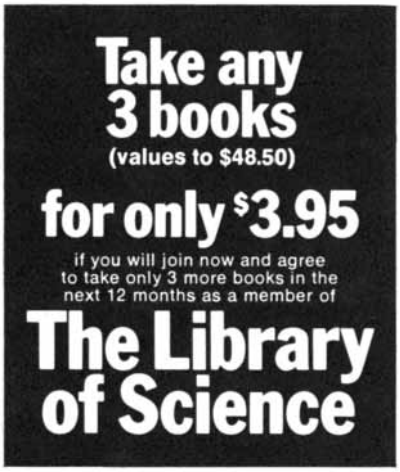
50110. GENETICS OF THE EVOLUTIONARY PROCESS. *Theodosius Dobzhansky.* Perhaps the most critically acclaimed work on our total knowledge of evolutionary theory. **$10.95**

50380. GLACIAL AND QUATERNARY GEOLOGY. *Richard Foster Flint.* The most complete single volume on the subject in any language. *Counts as 2 of your 3 selections.* **$24.95**

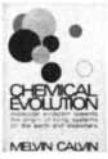
34610. ASIMOV'S BIOGRAPHICAL ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY. New Revised Edition. *Isaac Asimov.* **$12.95**

64580. THE NEW PROMETHEANS. *Robert S. deRopp.* The great discoveries, the evil discoveries, the implications of both. **$7.95**

64910. NUMBER WORDS AND NUMBER SYMBOLS. *Karl Menninger.* Scholarship of the highest order on the history of numbers, and the words and symbols used to express them. **$14.95**



68960. PHYSICS AND BEYOND. *Werner Heisenberg.* The author of the "uncertainty principle" takes you behind the scenes with the men who unravelled the secrets of quantum mechanics. **$7.95**



37530. CHEMICAL EVOLUTION. *Melvin Calvin.* The Nobel Laureate's major work on the biochemical evolution of life. **$9.00**

73180. QUANTUM THEORY AND BEYOND. *Edited by Ted Bastin.* 21 of the finest minds in the world consider the restructuring of quantum theory — and how it is to be done. **$16.00**

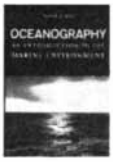
69970. POPULATION, RESOURCES, ENVIRONMENT. *Paul R. Ehrlich and Anne H. Ehrlich.* Very simply, the ecological book of the year. **$8.95**

45990. ESSAYS ON CELLULAR AUTOMATA. *Edited by Arthur W. Burks.* Eight experts in a fascinating overview of one of the fastest-growing and intriguing branches of computer science. **$12.50**

42970. THE DYNAMIC EARTH. *Peter J. Wyllie.* A revolutionary approach to the "new global tectonics". A most exciting study. **$17.50**

85830. THE VIEW FROM SPACE. *Merton E. Davies and Bruce C. Murray.* The main thrust: toward exploration of the entire solar system. **$14.95**

56030. INTRODUCTION TO APPLIED NUMERICAL ANALYSIS. *Richard W. Hamming.* The head of Bell Labs' Computer Science Research Department on the basics of computational mathematics. **$14.95**



65210. OCEANOGRAPHY: An Introduction to the Marine Environment. *Peter K. Weyl.* The geology, chemistry, life and energy processes of the mysterious ocean world. **$12.50**

57900. LASERS. Second Edition. *Bela A. Lengyel.* The best exposition of laser art currently available, emphasizing principles, not mere technical detail. **$14.95**

42150. CHAMBERS' DICTIONARY OF ELECTRONICS AND NUCLEONICS. *Edited by L. E. C. Hughes, R. W. B. Stephens and L. D. Brown.* More than 9,000 individual entries and over 100 pages of data tables. **$14.50**

(retail prices shown)

MEMBERSHIP APPLICATION

THE LIBRARY OF SCIENCE, Riverside, N.J. 08075

2-644

Upon acceptance of this order, please enroll me as a member and send the three books I have indicated. Bill me only $3.95 for all three, plus postage and handling. If not delighted, I will return all books within ten days and this membership will be cancelled.

As a member, I need accept only three more selections during the next 12 months at reduced member prices, plus postage and handling. Savings range up to 30% and occasionally even more.

I understand that I will receive free advance Reviews which fully describe each month's Main Selection and Alternates. If I wish the Main Selection, I need do nothing; it will come to me automatically. If I do not wish to receive the Main Selection, or I wish an Alternate Selection, I will advise you by using the convenient reply card provided with the monthly Reviews. I understand that I may choose a bonus book for every four selections purchased. (Introductory offer counts as first selection toward a bonus book.)

Send no money. Members are billed when the books have been received.

(Offer good in Continental U.S. and Canada only.)

3 Books for $3.95. (write in numbers): Some expensive books count as 2 or 3 selections.

_____ _____ _____

Name _____

Address _____

City _____ State _____ Zip _____

BOOKS

Photography from space; Helmholtz, Michelson and Morley; the history of mammals and man

by Philip Morrison

THE VIEW FROM SPACE: PHOTOGRAPHIC EXPLORATION OF THE PLANETS, by Merton E. Davies and Bruce C. Murray. Columbia University Press ($14.95). AIRBORNE CAMERA: THE WORLD FROM THE AIR AND OUTER SPACE, by Beaumont Newhall. Hastings House, Publishers, Inc. ($10). GEMINI SYNOPTIC PHOTOGRAPHY CATALOG. Technology Application Center, University of New Mexico ($1). APOLLO SYNOPTIC PHOTOGRAPHY CATALOG. Technology Application Center, University of New Mexico ($2). Once the protracted Martian dust storm of last fall cleared, the two slow-scan Vidicon cameras now orbiting the rusty planet in *Mariner 9* at 1,000 miles up have been viewing a new world. The newspapers carry some photographs; in January you could see the Grand Canyon of the Martian tropics, fed by a dozen branching tributaries similar in form to the Little Colorado itself, although presumably they are not gorges worn by water but slumps sharpened by the fierce, thin winds and volcanic gases of dry Mars. Few pieces of scientific data have the immediate impact of such photographs. Our eye-minded age, nurtured on halftone and glowing screen, can feast on photographic views from space, from those moving glimpses of our own sky blue earth to the moon mapped fully from orbit. One day there will be an atlas for every planet in the sun's family.

Merton Davies is a veteran photographic scientist with the Rand Corporation; Bruce Murray is a Cal Tech planetologist. Their *View from Space* is a crisp general account of space photography for the earth, the moon and the planets, with an interest as distinctly in the "engineering, fiscal and historical constraints" as in the scientific merits of the discipline. The last quarter of the text is a dense, useful technical summary of the basic information theory behind picture-winning, of camera design and of the systems actually flown, both Russian and American. The Mars orbiter stores its digital television picture electrically; its two tape recorders are ingeniously arranged to talk with each other to improve picture effectiveness by digital processing, both internally near Mars and after reception in California. The Lunar Orbiters, on the other hand, stored their images on film, which is definitely more compact than magnetic tape, and they scanned the film to read out pulses for transmission. Mars will be mapped this season from near orbit at a resolution of better than a kilometer over about a thousandth of its surface. We shall have nearly 100 million bits (one encyclopedia's worth) at that detail. All our telescopic study of Mars—Percival Lowell's "biased and misleading observations" and all—has yielded only about the same number of bits, but at a resolution never better than 100 kilometers.

The style of the book is refreshingly unusual; the authors do not hesitate to criticize scientific policy and planning in their broadly technical account. They hold that "Lowell's legacy," a self-deluding belief in life on Mars, was a misinterpretation of Martian conditions inferred even from modern ground-based observations, and that it resulted in what they see as a hasty American overemphasis on the study of Mars among all the planets. Plans for Mars have been overambitious. In the end they were not supported, and more modest and more balanced programs might have been carried out without treating Venus as "a woman scorned" or ignoring singular Mercury, the densest chunk in the sun's family. These views are well argued, but it is by no means clear that the authors are correct. They themselves view the far more costly and single-minded man-on-the-moon program as an essential and wise national stimulus toward space.

The book contains a good deal—even the verbatim text of a proposal tabled by the U.S. at Geneva in 1958—about the prospective uses of orbital reconnaissance in terrestrial arms control. It is therefore fitting to point out that the authors are intolerably coy about the present state of earth reconnaissance. They refer two or three times to the military uses of space; one table implies that such operations consume about a fourth of the total U.S. space budget. The foreword, by a Rand vice-president, goes so far as to mention the name of the Discoverer series—several score of Rand-initiated satellites sending back recoverable film capsules—and to observe a little primly that "the results in some instances are necessarily obscured by the requirements of national security." No more is said about the scale, methods or broad results of what is clearly the largest and most advanced part of space photography, from which a good deal of hardware for planetary exploration has probably been derived. Once again the shortsighted Rand–Air Force security principles strain the lines of trust; the other side of such disingenuous "open" publication may appear one day as unauthorized guerrilla release by Xerox.

The scale and viewpoint of the orbital photographs of course have a long precedent. That story, together with a good simple account of the basis of map making and of photointerpretation, is the theme of Beaumont Newhall's *Airborne Camera*. He is a well-known historian of photography who appreciates photography as a craft, as a science and as an art form. He retains his touch in his chronicle of the airborne photograph. It evidently all began in the world of the artist's imagination. For example, in 1827 the painter George Catlin painted "Topography of Niagara," a view of the falls from straight above as fine, with a little license, as an excursion in a light plane could show you today. The lithograph he made of his painting is reproduced, with a portfolio of black-and-white photographs from balloon, kite, rocket, carrier pigeon, airplane and orbiting capsule. It was split-second exposure first achieved with wet collodion plates about 1859 that began aerial photography. The ruins of San Francisco taken with a big view camera, holding a plate that mea-

sured four feet by a foot and a half and suspended from a train of 17 kites sent aloft from a ship in the bay, are one subject shown here that was topical in 1906 and is still hard to beat.

In 1914 the Royal Flying Corps had a pair of talented young pilots over France who brought back snapshots they had made from the air, only to be told it was "a most disgraceful thing to have attempted." By the following season, however, special cameras were fixed to the side of a few wood-and-canvas two-seater biplanes. Hand-held cameras have been flown even recently, not of course in aircraft but in a dozen earth-orbiting manned flights of Gemini and Apollo. It was almost the only vestige of improvisation. Not until *Apollo 13* did up-to-date aerial cameras appear.

The spectacular color photographs made by the astronauts from 100 miles up are familiar from a score of magazines, including this one. We have earlier described in these reviews the bargain NASA volumes from the Government Printing Office, and the equally remarkable atlas-like volume by Rand McNally & Company, that reproduce many of those beautiful and fascinating pictures, resembling good maps magically turned into reality. Now one can obtain one's choice of nearly 4,000 shots taken at hyperaltitude from a dozen orbital missions. They are catalogued by location, with a very brief description of subject and quality, in the two paperbound typescript catalogues for the photography of Gemini and Apollo. The New Mexico Center will promptly send you—within a week or so in one test order—any frame you choose at about a dollar per color transparency, and a variety of larger formats and of preselected thematic sets (sea state, Peru or hydrology, for example) at other prices after a longer wait.

This is a remarkable photographic resource; travelers anywhere in the Tropics or subtropics can now see their own voyages from the vantage of the angels. Lake Chad or Mount Everest, the Irrawaddy or the Congo, even the lakes near Hangchow before Professor Kissinger, are listed. The Center is a nonprofit representative of NASA; you taxpayers might as well order some copies, because you have already paid the cost of the original exposures.

SELECTED WRITINGS OF HERMANN VON HELMHOLTZ, edited, with an introduction, by Russell Kahl. Wesleyan University Press ($25). THE ETHEREAL AETHER: A HISTORY OF THE MICHELSON-MORLEY-MILLER AETHER-DRIFT EXPERIMENTS, 1880–1930, by Loyd S. Swenson, Jr., with a foreword by Gerald Holton. University of Texas Press ($10). When Benjamin Franklin went to Paris, he was welcomed as natural philosopher no less than as representative of a new political order. A hundred years later American physics boasted no such self-taught masters. J. Willard Gibbs, the deepest American theorist, like the virtuoso experimenters Henry A. Rowland and A. A. Michelson, did postgraduate work abroad. The Berlin school of Hermann von Helmholtz received all three young men, and it demonstrably gave them much.

Helmholtz is represented in our memories and our libraries by his big books on physiological optics and on hearing, still useful in their special domains even though the newer of them was first published during our Civil War. But he was a scholar-scientist in a tradition far from the cliché of the German specialist professor; he was a man of sensibility in literature and in poetry, a talented musician, a first-class experimenter, a powerful mathematical physicist and a trained and explicit philosopher. *Selected Writings* presents 20 shorter papers that "seem most clearly to deserve a permanent place in our libraries."

The man comes through plainly in these lectures and papers. (One of the more mathematical ones, say on hydrodynamics, ought to have been included.) The translation and editing are mainly new and generally excellent, although for the price we have the right to grumble at the absence of an index. Helmholtz of course writes very well indeed on the questions of perception, and he is also surprisingly good reading on the solar system, on medicine, on electrolysis and its implications of an atom of electric charge. One pervasive theme is the deep question of the origin of the ideas of mathematics and their apparent control of the world. He claims to be a Kantian, but in fact he is more modern; for Helmholtz mathematics was a rich storehouse of models, not an a priori system to which we find marvelous correspondences. "We must not be surprised if the axioms of arithmetic turn out to be true of natural processes, for we recognize as additive only such physical combinations as conform to the axioms of arithmetic." Resistances add in series, we say, but capacitors add only in parallel.

But the note that rings most strikingly in all these warm and graceful essays is Helmholtz' recognition that we do not dwell in Euclid's space a priori, as his master Kant taught. No, space might well be curved; we can tell only by measurement. It was Helmholtz' study of the color space and the space of the visual image that convinced him. He even alludes to non-Riemannian spaces, where a figure after rotation is similar to but not congruent with its old self.

Professor Swenson's meticulous history of the Michelson-Morley-Miller experiments makes close contact with Helmholtz. Young Michelson, five years an ensign in the regular Navy and newly promoted to lieutenant junior grade, a physics instructor at Annapolis with one measurement of the speed of light already behind him, went off to Europe for postgraduate study. He had leave with pay under the patronage of Simon Newcomb, then superintendent of the *Nautical Almanac* and a man of international standing. In Paris he studied with Alfred Cornu and probably encountered Hippolyte Fizeau. His interferometer was conceived in Paris. In Berlin during the ensuing months he described the scheme to Helmholtz in "quite a long conversation."

Helmholtz could see no objection to Michelson's plan for "finding the motion of the earth relative to the ether ... except the difficulty of keeping a constant temperature." The first trial—with an error in the theory of a factor of two—was made with a brass instrument built to order by a Berlin firm with grant money from Alexander Graham Bell. It was no good making a run in the city laboratory; even during the small hours vibrations of the stone pier blurred the fringes. (Michelson loved the delicacy of his interfering beams. This reviewer can sympathize, recalling his own first sight 40 years ago of such an interferometer, built on a length of railroad rail, able to show the deflection produced by a piece of paper loading the heavy rail.) But a circular cellar under the big telescope in suburban Potsdam was quiet enough to show that the fringes did not shift when the crossed meter-long arms were rotated with respect to the direction of the presumed motion of the earth and the sun, a conclusion that, Michelson wrote, "directly contradicts the explanation of ... aberration ... which presupposes that the earth moves through the ether."

Six years later, back at the Case School and Western Reserve College in Cleveland, a much better device—the famous sandstone slab floating on mercury—gave an even clearer null result. But Michelson and his colleague Morley were not anxious to draw profound conclusions. The ether is dragged by the heavy opaque surroundings, they

thought (with George Stokes). Michelson went off to Clark University and thence to the University of Chicago. Ten years later he looked very sensibly for a vertical velocity gradient in the ether: maybe the ether wind would blow more strongly 50 feet up on the roof of Ryerson Laboratory. No shift. He did not return to the problem again until 1925.

Professor Swenson is a historian, not a physicist. He succeeds very well in making it plain that the familiar naïve pseudo-history, in which Einstein's special theory was created to explain away the Michelson-Morley null result, is not part of the real time sequence at all. Einstein was mining a far richer lode. It was rather the pedagogy of scores of popular and textbook accounts written mostly after 1915, each seeking a simple line of exposition, that gave crucial importance in hindsight to the two crossed light paths. This readable book includes the famous table from the modern textbook by W. K. H. Panofsky and Melba Phillips, which does something like justice to physicists' views. Thirteen experimental results of many kinds are tabulated for their agreement or disagreement with seven variant theories: the emission of light with several kinds of particle-like velocity sharings, and ether theories, passive or active. Only special relativity provides a unique prediction for every test and passes them all. That is closer to science than any more restricted viewpoint can be.

The author tries learnedly to place the issue in a philosophical arena that is very popular these days. The notion that any experiment led to Einstein he sees as a bias, sheer positivism justly out of date; actually it was the elegance of the symmetries that led Einstein on. That is the artistic rationalism now popular, in proper cyclical reaction to the overreduced austerities out of Vienna between the world wars. But physicists know that Newton's theory claims its invariants too, just as rational as those of H. A. Lorentz that Einstein perceived. The difference is in their fertility; dynamics, not electromagnetism, was an unexpected arena of success for special relativity. E equals mc^2, after all. Einstein himself said it: A physicist ought not to be so consistent a philosopher; he is an eclectic, an opportunist, an exploiter of tensions. One must watch what such people do but not mind so closely what they say. (Historians might also attach to every remark they quote the age of the physicist who said it.)

The figure of Dayton C. Miller colors a sadder background in most of the book. It was the young and energetic Miller who succeeded Michelson in Cleveland. He joined Morley and carried on the work long afterward, supporting the antirelativism of a stubborn and by no means vanished minority. He knew there was a small residual fringe shift; it persisted through many redesigns in wood and steel and concrete, through trials at sea level and on mountaintop. His results received repeated attention and induced many repetitions and improvements by others; the Jena physicist Georg Joos, working with a huge automated helium-filled instrument made by Zeiss, reduced the signal by 1930 to less than a thousandth of a fringe. It was convincingly concluded by R. S. Shankland of Case in 1954 that Miller's definite Mount Wilson effect, found in a laboratory made like an airy garden shed in order not to trap the ether, arose from the temperature effects Helmholtz had warned about so long before.

It is not the ether as a medium that is denied by Einstein, in spite of the entire emphasis of this book and of many discussions cited in it. We can deny only the Newtonian properties of the ether, in particular the linear addition of velocities. Quantum electrodynamics builds a real plenum in space, to use a still older language. There they are, those fluctuating photons and particle pairs in God's plenty. But you cannot assign them any unique inertial frame; that rich vacuum or ether or space or the plenum is what looks alike to all inertial observers, call it what you will.

THE AGE OF MAMMALS, by Björn Kurtén. Columbia University Press ($9.95). NOT FROM THE APES, by Björn Kurtén. Pantheon Books ($5.95). A Helsinki paleontologist, Professor Kurtén is a fluent writer in both Swedish and English. In Swedish he has written novels; in English he has presented the general reader with five volumes spanning a wide range of paleontological topics. These two, at once graceful and decisive, are the latest.

We readers are mammals, "tetrapods with hair that suckle their young." Nothing mammalian among the strange beasts with which evolution has peopled the ark of the earth is quite alien to us. The reptiles ruled land and sea for a very long time. We mammals are descended from a form of somewhat warm-blooded reptiles that appear to have had the specialized forms of teeth whose efficiency has always marked our hungry kind, which need food to fuel perpetual warmth. For more than 100 million years mammals were mainly small opossum-like animals or little insect-eating shrews, beneath the notice of the fearsome *Tyrannosaurus* that stalked the plains to prey on the great horned-dinosaur herds. By the end of the Cretaceous the Age of Reptiles had passed. The drifting continents had broken up the land masses into eight separate areas, "each one a private Noah's Ark"; never before or since has the land been so fragmented. That fact ensured mammalian diversity; each ecological niche, say of giant herbivores on the plains, is filled by performers that have no single lineage but come from a succession of unrelated groups. "The roles persist, unchanging like Columbine, Pantaloon and Harlequin," but new actors play the roles, entering from other areas as land masses part and join, each with its local talent, now the one and now the other better fitted for a wider stage. "Natural selection is judge and jury."

In this vein the volume proceeds from the opening of the Age of Mammals 65 million years ago through the six stages of the Cenozoic until recent times only a hundredth of a million years ago, when modern man began a new form of selection, for other species as well as for himself. The book takes on a kind of epic tone: variations on a great tragic theme. With each epoch there is a new map, a new set of transient solutions to the problems of gaining a living from the distant sun through the chain of life. There is also generally a new fossil location, where chance has preserved an entire sample fauna to illuminate how life was led.

This kind of treatment, far removed from the evidence but so much warmer than the expert's treatise, comes off remarkably well in Professor Kurtén's hands. He has a way of being full and explicit without being tedious; the connections he suggests supplant the bone shapes of the primary documents. (Plant life is somewhat slighted throughout, even the decisive grasses.)

Meet a few characters and events of the rich narrative. There are the terror cranes, *Diatryma*. These flightless "imperious effigies of the roc bird" stood eight feet high, with a huge head, a deep, powerful beak and great four-clawed feet; they could kill any mammal of the Paleocene. In them one can see "the ancient reptile line... making a final bid for supremacy on land; for birds, of course, are descended from reptiles, and are cousins to the dinosaurs." Then there is a muster of real giants, mainly back a couple of million years in the Pleistocene. There is a rhinoceros with a horn two meters long, a wild horse as

heavy as a Clydesdale, a marsupial hippopotamus 11 feet long, a gigantic North American cat with bone-sheathed saber teeth ("the most grotesque and aberrant felid known"), a lion-sized hyena, a bear-sized beaver and (a little earlier, to be sure) a rodent the size of a hippopotamus.

For contrast, on the islands there is a donkey-sized elephant on Malta, along with giant rats and dormice on other Mediterranean islands. No big carnivores can live on islands; they need too large a range. Therefore the bigger herbivores can become small without much danger, and the little timorous ones can "grow big and fat with impunity." Certainly on the islands, Kurtén says, and very likely in Africa and America as well, it was predator man who caused the extinction of a large fraction of the bigger species of mammals, by fire and the hunt. Man has even pressed the survivors to smaller size: five carnivores of the eastern Mediterranean coasts show this effect toward the end of the last glaciation: wolf, hyena, bear, wildcat and marten. Only the red fox grew bigger. Did he prosper somehow around the camps of the Paleolithic hunters?

The Age of Mammals is a compendious background work for the biological amateur. *Not from the Apes,* a brief, personal book, with a glossary of definitions instead of the careful indexes and references of the bigger book, is aimed at a much wider audience; it puts forth explicit conclusions about humankind. It is the work of an evolutionist with imagination and convictions. He states both unequivocally, arguing his points engagingly, sometimes from the evidence described and sometimes by learnedly referring to such studies as those of F. E. Koby, who showed that Neanderthalers were right-handed. Koby found scratches on their front teeth that had been made, it is said, by the flint blades they had used to cut off the morsel they had stuffed into their mouth. Kurtén gives us no way to pursue the matter. He is a bone-oriented scientist, by his own statement unenthusiastic about more complex modes of theorizing on human behavior. Certainly his matter-of-factness, however much it owes to unstated assumptions, is a welcome relief from the stream of writers who have woven the whole cloth of the past out of some partial insight, say into territoriality.

Kurtén's ascent of man begins in the deep African past, with four or five primate forms found in fossil forest beds deposited 35 million years ago beside the proto-Nile. One of those forms—we have only a few jawbones—already has some manlike traits; other quite distinct forms are ancestral to the great apes. Thus we share our cousins' forebears but we are indeed not from the apes. They arose from our early ancestors. Kurtén believes that ever since the time of *Australopithecus* (of which there are two distinct strains, the robust and the gracile) men have been split into "hominids that did not breed with each other and presumably did not regard [each other] as 'men.'" Neanderthal man appeared and vanished, under competition or by direct violence, in just that way.

So Kurtén holds, and he sees in that past an evil omen for our future. Our most abhorrent trait—ceasing to look on one another as being members of the same species—now implies overall disaster; today's weapons are deadlier than flint. Kurtén hopes for some modestly successful compromise with "enough love and compassion ... to carry us through." Given those we shall yet evolve.

These are two distinct and quite interesting books. Professor Kurtén is not always convincing on man; he gives too few numbers. He sees the avoidance of predation as being maximally important among survival paths, but there is no great support for that in many cases. He thinks natural selection among mankind is more important than ever. That might be true in some absolute way, but in a relative sense it is surely impossible. It is culture that has caused our population to grow several thousandfold in four million years. The hominids increased by only one power of 10 in all the time before the rise of *Homo sapiens.* They have advanced more than two powers of 10 within the brief life-span of that single species. Bone-based evolutionary thought and qualitative population genetics are not enough to explain social humanity, although we can never understand ourselves without them.

FINGERPRINT TECHNIQUES, by Andre A. Moenssens. Chilton Book Company ($7.95). Apes and men alike have an intricate set of ridges on the palmar surfaces of their hands and on the soles of their feet. Along these ridges, which are presumably nonskid adaptations, like the tread of a tire, are irregularly scattered the ends of the tiny ducts that discharge perspiration from the sweat glands a millimeter or so down in the living dermal layer. The ridges form in the third or fourth fetal month, and they persist from birth to death. They change in size with growth, even change somewhat as the fingers press lightly or firmly and so are distorted under pressure. The topology—the whorls, the loops, the deltas, the arches and the rest—never changes.

The systematic use of fingerprints as identification in criminal cases is a modern phenomenon. A Scottish physician working in Japan first published the suggestion in 1880; the geneticist Francis Galton published the first book on the topic and pressed for the use of fingerprints by the British police. The modern system, which among other things has deposited some hundreds of tons of fingerprint cards in the files of the FBI in Washington, is the product of Bengal. It was the British inspector general of police there at the turn of the century, with his mathematically minded Bengali assistants, Azizul Haque and Hem Chandra Bose, who developed the system still used as the basis of fingerprinting in many countries, including the U.S. today. The bibliography of this unusual textbook lists recent texts from Dacca and Calcutta; interest remains high in Bengal.

The one-to-one identification of fingertips is made by direct comparison of details of ridge endings, bifurcations and other small features of the prints on a submillimeter scale. A dozen such points must be found to agree in topology and direction, although not in measurement, before an identity is firm. The author, a Chicago criminologist who writes in a clear, disarming style, is plainly an honest scientific associate of and instructor to the working police. He says that he has never been able to find as many as "four similar ridge characteristics in identical interrelationships" in different finger impressions. Identical twins and triplets are no exception, in spite of repeated newspaper stories to the contrary. The overall type of pattern tends to be inherited; it is strongly correlated in identical siblings, but the details of the print are not. John Dillinger did not understand this difference. He obliterated the key "delta" patterns on all 10 fingers with acid, because he knew that prints were classified by such patterns. His scarred prints were of course even easier to classify.

It is true that the great files of prints stand mute unless a search process can be limited by some classification easier than the print-by-print comparison of the details used for final identity. Much of the text deals with clever and painstaking schemes for assigning unambiguous codes to classes of prints. Even with the best of them a print-by-print search may well require looking at thousands of cards. Machine searching is already used widely, although by no

means universally and not for the largest national files. The New York State method involves counts of ridges on each finger and some coded assignment of patterns. Then a punched-card file can be searched for the entire set of numbers, at a rate of seven or eight cards per minute. Fully automatic methods, without the human assignment and count, are latent in the technology; like most pattern-recognition problems, this one is not yet solved as a practical matter.

The use of fingerprints in crime involves matching the inked print in some file—or extracted from the suspect—with a piece of field microanalysis: the couple of hundred micrograms of residue left behind by evaporating sweat somewhere on a gun or a glass or a doorknob in the true ridge pattern. (Bloody prints, or prints in dust or clay, are easier to see.) There is an interesting technology for revealing such latent prints. The most direct method is the use of an adhering fine powder. Lampblack and aluminum dust will do, but there are many variants, from fluorescent powders to mercury and chalk. (One British police department gave up the classic gray mercurial mix when seven of their men developed mercury poisoning.) The print contains a fatty fraction, so that it can also be made visible by iodine reactions. It contains chloride ion, so that silver nitrate blackens it too. It contains amino acids, so that a complex organic reagent, ninhydrin, shows it up as a pink pattern. Ninhydrin works best on old prints, since the amino acids diffuse much more slowly than salts: "The author developed several identifiable latent impressions in a college textbook that he had not used or touched for at least nine years."

The ridge pattern can be removed from the fingertips. It hurts. The dermatologist planes the skin a millimeter or more deep with a wire brush. The new skin, six weeks later, has no ridges, leaving a bland pattern of pores and creases. Prints can probably be forged too; the most subtle fraud is the lifting of a real powder-developed latent print in the normal Scotch-tape way but falsely placing it in the record book. Instead of taking the print from, say, a gun, a dishonest technician can lift it from some quite innocent spot. Only the probity of the experts, the use of witnessed procedures and similar schemes of the advocacy process can guard the law from that kind of fraud. The same author has written an earlier book with three chapters of questions and answers to show the honest expert how to make his actions manifest in court.

BIBLIOGRAPHY

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

POLITICAL FACTORS IN ECONOMIC ASSISTANCE

ASIAN DRAMA: AN INQUIRY INTO THE POVERTY OF NATIONS. Gunnar Myrdal. Pantheon Books, 1968.

THE CHALLENGE OF WORLD POVERTY: A WORLD ANTI-POVERTY PROGRAM IN OUTLINE. Gunnar Myrdal. Pantheon Books, 1970.

THE ARMS TRADE WITH THE THIRD WORLD. SIPRI: Stockholm International Peace Research Institute. Humanities Press, 1971.

COLLECTIVE-EFFECT ACCELERATORS

SYMPOSIUM ON ELECTRON RING ACCELERATORS: HELD AT LAWRENCE RADIATION LABORATORY, UCRL-18103. University of California at Berkeley, 1968.

THE ACCELERATION OF PARTICLES BY COLLECTIVE FIELDS: I. Andrew M. Sessler in *Comments on Nuclear and Particle Physics,* Vol. 3, No. 4, pages 93–100; July–August, 1969.

RESEARCH ON THE ELECTRON RING ACCELERATOR. D. Keefe in *Particle Accelerators,* Vol. 1, No. 1, pages 1–13; March, 1970.

THE ACCELERATION OF PARTICLES BY COLLECTIVE FIELDS: II. L. Jackson Laslett and Andrew M. Sessler in *Comments on Nuclear and Particle Physics,* Vol. 4, No. 6, pages 211–218; November–December, 1970.

COLLECTIVE AND COHERENT METHODS OF PARTICLE ACCELERATION. J. D. Lawson in *Particle Accelerators,* Vol. 3, No. 1, pages 21–33; January, 1972.

AN EARLIER AGRICULTURAL REVOLUTION

SOUTHEAST ASIA AND THE WEST. Wilhelm G. Solheim II in *Science,* Vol. 157, No. 3791, pages 896–902; August 25, 1967.

THE HOABINHIAN AND AFTER: SUBSISTENCE PATTERNS IN SOUTHEAST ASIA DURING THE LATE PLEISTOCENE AND EARLY RECENT PERIODS. Chester Gorman in *World Archaeology,* Vol. 2, No. 3, pages 300–320; February, 1971.

EXCAVATIONS AT NON NOK THA, NORTHEASTERN THAILAND. Donn T. Bayard in *Asian Perspectives,* Vol. 13, in press.

TIDES AND THE EARTH-MOON SYSTEM

HISTORY OF THE LUNAR ORBIT. Peter Goldreich in *Reviews of Geophysics,* Vol. 4, No. 4, pages 411–439; November, 1966.

ONCE AGAIN—TIDAL FRICTION. Walter Munk in *The Quarterly Journal of the Royal Astronomical Society,* Vol. 9, No. 4, pages 352–375; December, 1968.

ANCIENT ASTRONOMICAL OBSERVATIONS AND THE ACCELERATIONS OF THE EARTH AND MOON. Robert R. Newton. Johns Hopkins Press, 1970.

DYNAMICAL ASPECTS OF LUNAR ORIGIN. William M. Kaula in *Reviews of Geophysics and Space Physics,* Vol. 9, No. 2, pages 217–238; May, 1971.

THE STRUCTURE AND HISTORY OF AN ANCIENT PROTEIN

THE STRUCTURE AND ACTION OF PROTEINS. Richard E. Dickerson and Irving Geis. Harper & Row, Publishers, 1969.

THE STRUCTURE OF CYTOCHROME *c* AND THE RATES OF MOLECULAR EVOLUTION. Richard E. Dickerson in *Journal of Molecular Evolution,* Vol. 1, No. 1, pages 26–45; 1971.

FERRICYTOCHROME *c:* I, GENERAL FEATURES OF THE HORSE AND BONITO PROTEINS AT 2.8 Å RESOLUTION. Richard E. Dickerson, Tsunehiro Takano, David Eisenberg, Olga B. Kallai, Lalli Samson, Angela Cooper and E. Margoliash in *The Journal of Biological Chemistry,* Vol. 246, No. 5, pages 1511–1535; March 10, 1971.

CONFORMATIONAL CHANGES UPON REDUCTION OF CYTOCHROME *c.* T. Takano, R. Swanson, O. B. Kallai and R. E. Dickerson in *Cold Spring Harbor Symposium on Quantitative Biology,* in press.

LANGUAGE AND THE BRAIN

CEREBRAL DOMINANCE AND ITS RELATION TO PSYCHOLOGICAL FUNCTION. O. L. Zangwill. Oliver and Boyd, 1960.

DISCONNEXION SYNDROMES IN ANIMALS AND MAN: PART I. Norman Geschwind in *Brain,* Vol. 88, Part 2, pages 237–294; June, 1965.

DISCONNEXION SYNDROMES IN ANIMALS AND MAN: PART II. Norman Geschwind in *Brain,* Vol. 88, Part 3, pages 585–644; September, 1965.

HUMAN BRAIN: LEFT-RIGHT ASYMMETRIES IN TEMPORAL SPEECH REGION. Norman Geschwind and Walter Levitsky in *Science,* Vol. 161, No. 3837, pages 186–187; July 12, 1968.

TRAUMATIC APHASIA: ITS SYNDROMES, PSYCHOLOGY AND TREATMENT. A. R. Luria. Mouton & Co., 1970.

SUPERCONDUCTORS FOR POWER TRANSMISSION

UNDERGROUND POWER TRANSMISSION: A REPORT TO THE FEDERAL POWER COMMISSION. Commission's Advisory Committee on Underground Transmission. U.S. Government Printing Office, April, 1966.

THE 1970 NATIONAL POWER SURVEY. Federal Power Commission. U.S. Government Printing Office, 1971.

THE 1970 NATIONAL POWER SURVEY, PART II: ADVISORY REPORTS. Federal Power Commission. U.S. Government Printing Office, 1971.

SUPERCONDUCTING POWER TRANSMISSION. R. W. Meyerhoff in *Cryogenics,* Vol. 11, No. 2, pages 91–101; April, 1971.

ENVIRONMENTAL CONTROL IN THE BEEHIVE

THE BEHAVIOUR AND SOCIAL LIFE OF HONEYBEES. C. R. Ribbands. Bee Research Association Limited, 1953.

WORLD OF THE HONEYBEE. Colin G. Butler. Macmillan Co., 1954.

ANATOMY OF THE HONEY BEE. R. E. Snodgrass. Cornell University Press, 1956.

COMMUNICATION AMONG SOCIAL BEES. Martin Lindauer. Harvard University Press, 1961.

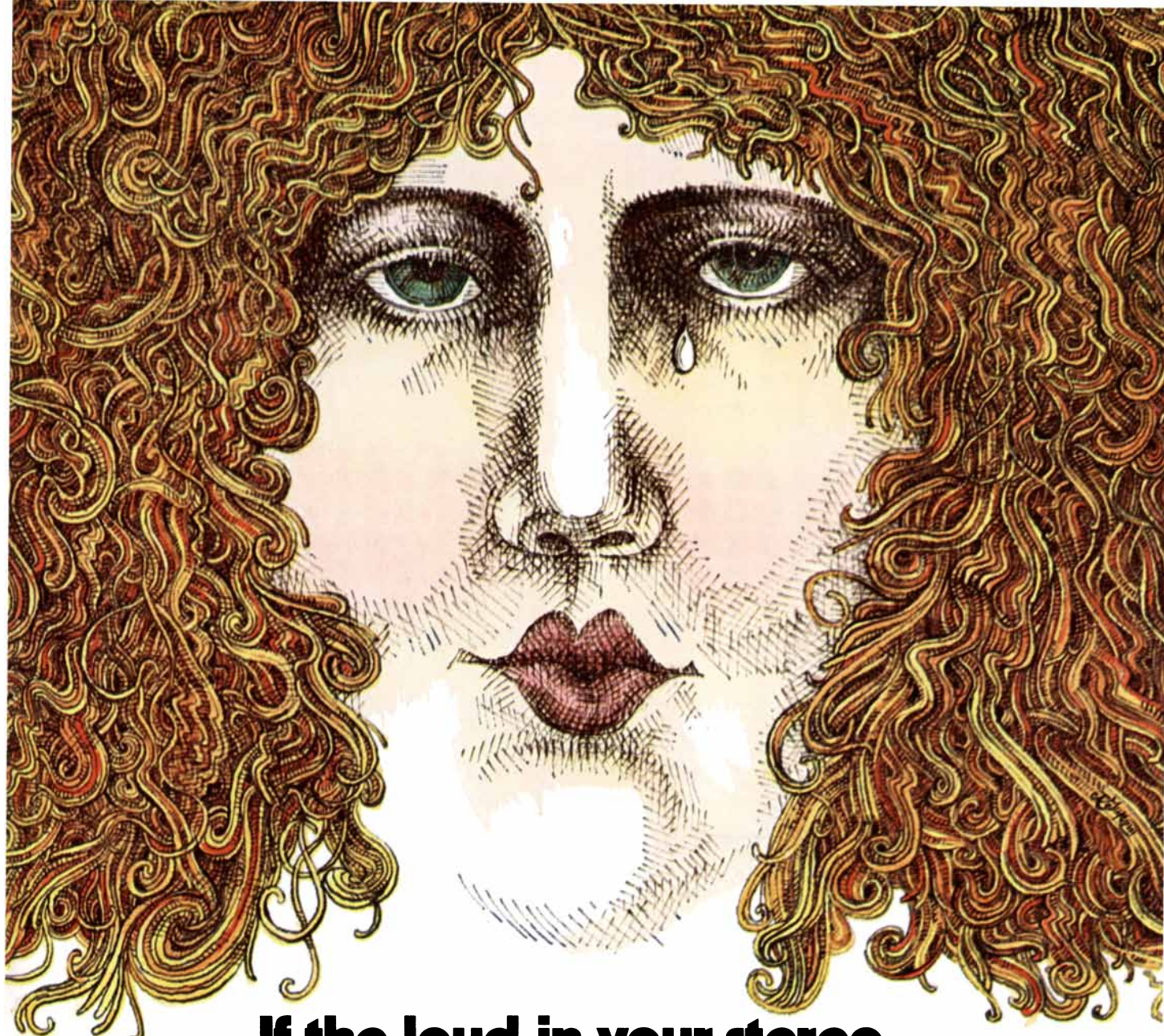
BEES: THEIR VISION, CHEMICAL SENSES, AND LANGUAGE. Karl von Frisch. Cornell University Press, 1971.

THE AMATEUR SCIENTIST

SIMULATOR FOR STATIC LIQUID CONFIGURATION IN PROPELLANT TANKS SUBJECT TO REDUCED GRAVITY. William A. Olsen in *NASA Technical Note D-3249.* Scientific and Technical Information Division, National Aeronautics and Space Administration.

ELEMENTARY MECHANICS OF FLUIDS. Hunter Rouse. John Wiley & Sons, Inc., 1946.

STATIC AND DYNAMIC BEHAVIORS OF THE LIQUID-VAPOR INTERFACE DURING WEIGHTLESSNESS. E. W. Otto in *A.I.Ch.E. Chemical Engineering Progress Symposium Series,* Vol. 62, No. 61, pages 158–177; 1966.



If the loud in your stereo gives her a pain in the head get a Marantz.

It's not that she's got super-duper sensitive hearing (like that spotted beagle two doors down), it's because most women hear better than men, so when she screams turn down the sound what she really means is turn down the damn distortion because the distortion is driving her bananas.

Not so with Marantz stereo. Take the Marantz 2270 stereo AM/FM receiver for $549.95. It delivers a walloping 140 watts RMS power at less than 0.3% distortion...which means virtually NO DISTORTION. And because Marantz measures distortion at continuous full power through the whole listening range—it won't bring tears to her eyes or a pain to her head. NO MATTER HOW LOUD. Other companies measure power and distortion only in middle listening range. So they quote their highest power and lowest distortion only in the middle. But on the low and high sides their power is down and their distortion is up. So if you pay for 140 watts be sure you get 140 watts at both ends and in the middle.

With Marantz you get exactly what you pay for. If all you need is 30 watts, take our model 2215 for $249.95. Want twice as much power? Our model 2230 gives you 60 watts RMS for $349.95. And our model 2245 at $449.95 delivers 90 watts RMS. ALL FULL RANGE CONTINUOUS POWER.

No matter which model you choose, remember this. You're getting Marantz quality. The same quality that goes into the magnificent Model 19 FM stereo receiver (shown) that costs $1200. YES. That's right, a cool $1200! But it is the absolute, ultimate, very best there is.

Visit your Marantz dealer. And take your wife along. (Or that spotted beagle two doors down.)

marantz.®
We sound better.

The Literary Guild is the best $1 investment you'll ever make!

Joining the Literary Guild is really a guaranteed investment. You get the best books, the finest authors, greatest savings. In fact, during your membership **you can expect to save more than 70%** off the prices of publishers' editions! That's more than any other book club offers on current best sellers and other fine books. Here's how the Guild does it:

- You get big savings on the introductory offer by taking any 4 books shown here for only $1, plus shipping and handling.
- Month after month, you continue to save on the books you really want. What makes the Guild unique, however, is that the new best sellers offered as selections and alternates **average 30% below** the prices of publishers' editions. (Plus a modest charge for shipping and handling.)
- You save again with the Guild's generous bonus book plan. The Guild gives a bonus certificate with the very first book you buy — and for every one thereafter. You choose from a wide selection of valuable Bonus Books — many at only $1.

So go ahead. Build the library you've always wanted. Delight in overall savings that can amount to 70% — and more — just by joining right now. As a member, you need choose only 4 books during the entire year — out of more than 40 listed each month in the free Literary Guild magazine. Join today. Send no money. Just mail the coupon.

Any 4 books for $1.

when you join the Literary Guild. You merely agree to accept only 4 selections or alternates during the coming year.

The Literary Guild

Dept. AL 652 Garden City, N.Y. 11530

Please accept my application for membership in the Literary Guild and send me the 4 books or sets whose numbers I have printed in the boxes below. Bill me only $1 plus shipping and handling, for all 4.

About every 4 weeks, send me the Literary Guild Magazine describing the Featured Selection plus at least 50 Alternate choices. If I wish to receive the Featured Selection, I need do nothing; it will be shipped to me automatically. Whenever I prefer an Alternate, or no book at all, I will notify you by the date specified by returning the convenient form always provided. I need take only 4 Selections or Alternates during the coming year, and may resign any time thereafter. The prices of books offered will average 30% below the prices of publishers' editions, plus a modest charge for shipping and handling. For each monthly Selection or Alternate I accept, I may choose a bonus book from the Guild's Bonus Catalogs at a fraction of the already-low Guild price — often for as little as $1. NO-RISK GUARANTEE: If not delighted, I may return the introductory package within 10 days. Membership will be canceled — I will owe nothing.

Mr.
Mrs.
Miss (Please Print)

Address

City

State Zip

Members accepted in U.S.A. and Canada only. Canadian members will be serviced from Toronto. Offer slightly different in Canada.

50-G852

8235. **THE WINDS OF WAR** Herman Wouk. (Pub. edition, $10.00)

1578. **MISS TALLULAH BANKHEAD** Lee Israel. (Pub. Edition, $7.95)

3046. **WHEELS** Arthur Hailey. (Pub. edition, $7.95)

3467. **HONOR THY FATHER** Gay Talese. (Pub. edition, $10.00)

1446. **THE EXORCIST** William Peter Blatty. (Pub. edition, $6.95)

1560. **BRING ME A UNICORN** Anne Morrow Lindbergh. (Pub. Edition, $6.95)

8227. **TRACY AND HEPBURN** Garson Kanin. (Pub. edition, $7.95)

1024. **THE BLUE KNIGHT** Joseph Wambaugh. (Pub. edition, $7.95)

3350. **INTIMATE BEHAVIOUR** Desmond Morris. (Pub. edition, $6.95)

0653. **OUR GANG** Philip Roth. (Pub. Edition, $5.95)

1156. **MONDAY THE RABBI TOOK OFF** Harry Kemelman. (Pub. Edition, $5.95)

3129. **THE MOON'S A BALLOON** David Niven. (Pub. edition, $7.95)

0463. **POOR COUSINS** Ande Manners. (Pub. edition, $8.95)

4135. **THE NEW YORK TIMES INTERNATIONAL COOKBOOK** Craig Claiborne. (Pub. edition, $12.50)

0612. **IN THE SHADOW OF MAN** Jane Goodall. (Pub. edition, $10.00)

1453. **LIMBO** Joan Silver and Linda Gottlieb. (Pub. Edition, $6.95)

0505. **BEAR ISLAND** Alistair MacLean. (Pub. edition, $5.95)

4168. **CONFESSIONS OF A STOCKBROKER "Brutus."** (Pub. edition, $6.95)

3335. **CIVILISATION** Kenneth Clark. (Pub. edition, $15.00)

0026. **HAMMOND CONTEMPORARY WORLD ATLAS** (Pub. edition, $12.50)

1271. **"I AM THIRD"** Gale Sayers. (Pub. Edition, $6.95)

1222. **THE RA EXPEDITIONS** Thor Heyerdahl. (Pub. edition, $10.00)

8292. **THE FATAL FRIENDSHIP** Stanley Loomis. (Pub. edition, $9.95)

2022. **THE CHANDLER HERITAGE** Ben Haas. (Pub. edition, $8.95)

0240. **VOYAGE TO THE FIRST OF DECEMBER** Henry Carlisle. (Pub. Edition, $6.95)

0042. **THE ARNHEITER AFFAIR** Neil Sheehan. (Pub. edition, $7.95)

1586. **THEIRS WAS THE KINGDOM** R. F. Delderfield. (Pub. edition, $8.95)

2279. **THE WORD** Irving Wallace. (Pub. edition, $7.95)

0448. **MUTANT 59: THE PLASTIC EATERS** Kit Pedler and Gerry Davis. (Pub. Edition, $5.95)

2402. **F. Scott Fitzgerald: TENDER IS THE NIGHT, THIS SIDE OF PARADISE, THE GREAT GATSBY, THE LAST TYCOON** 4 vols. count as 1 choice. (Pub. editions, $16.35)

0182. **THE GREAT NOVELS OF ERNEST HEMINGWAY** The Sun Also Rises, For Whom the Bell Tolls, A Farewell to Arms. 3 vols. count as 1 choice (Pub. editions. $13.95)

8201. **THE PEACEABLE KINGDOM** Jan de Hartog. (Pub. edition, $10.00)

0513. **SHADOWS IN PARADISE** Erich Maria Remarque. (Pub. edition, $6.95)

The Guild offers its own complete, hardbound editions, sometimes altered slightly in size to fit special presses and save members even more.

1972 Pinto Wagon, shown at top with Squire option, optional whitewall tires, and luggage rack; shown at bottom, with optional whitewall tires and Base Accent Group.

The Pinto Wagon. For people who'd like an economy car if it carried more. Or a wagon if it cost less.

It's a solid Ford Pinto.

Under the hood, you'll find a tough little 2000cc overhead cam engine as standard equipment.

The body is welded solid. Then electrocoated against corrosion. Then painted five more times.

There's rack and pinion steering like on some expensive sports cars.

In normal use, recommended maintenance is only every six thousand miles or every six months.

So if it's economy you're after, you'll find plenty in this new Pinto.

It's a convenient Ford Wagon.

You can get it as a basic wagon (below), or with the Squire option (above).

It's only 10 inches or so longer than our Pinto Sedan: easy to handle and park. But put the rear seat down and you get over 60 cubic feet of cargo space. VW Squareback and Vega Kammback both give you about 50 cubic feet.

The lift gate swings up out of the way. The spare tire has its own well under the floor. The rear seat passenger windows flip open for ventilation.

In short, the Pinto Wagon gives you the kind of convenience and usefulness you'd expect from any Ford Wagon. Only it gives it to you in a basic little Pinto size.

When you get back to basics, you get back to Ford.

FORD PINTO

FORD DIVISION Ford

We can turn on what we connect.

A switch for us? Exactly. And what it means is new and better switches for you.

Electrical switches and electrical connectors have a lot in common. Both use contacts . . . we've made them by the billions. Both need creative engineering for economical mass production to high quality standards. That's our specialty.

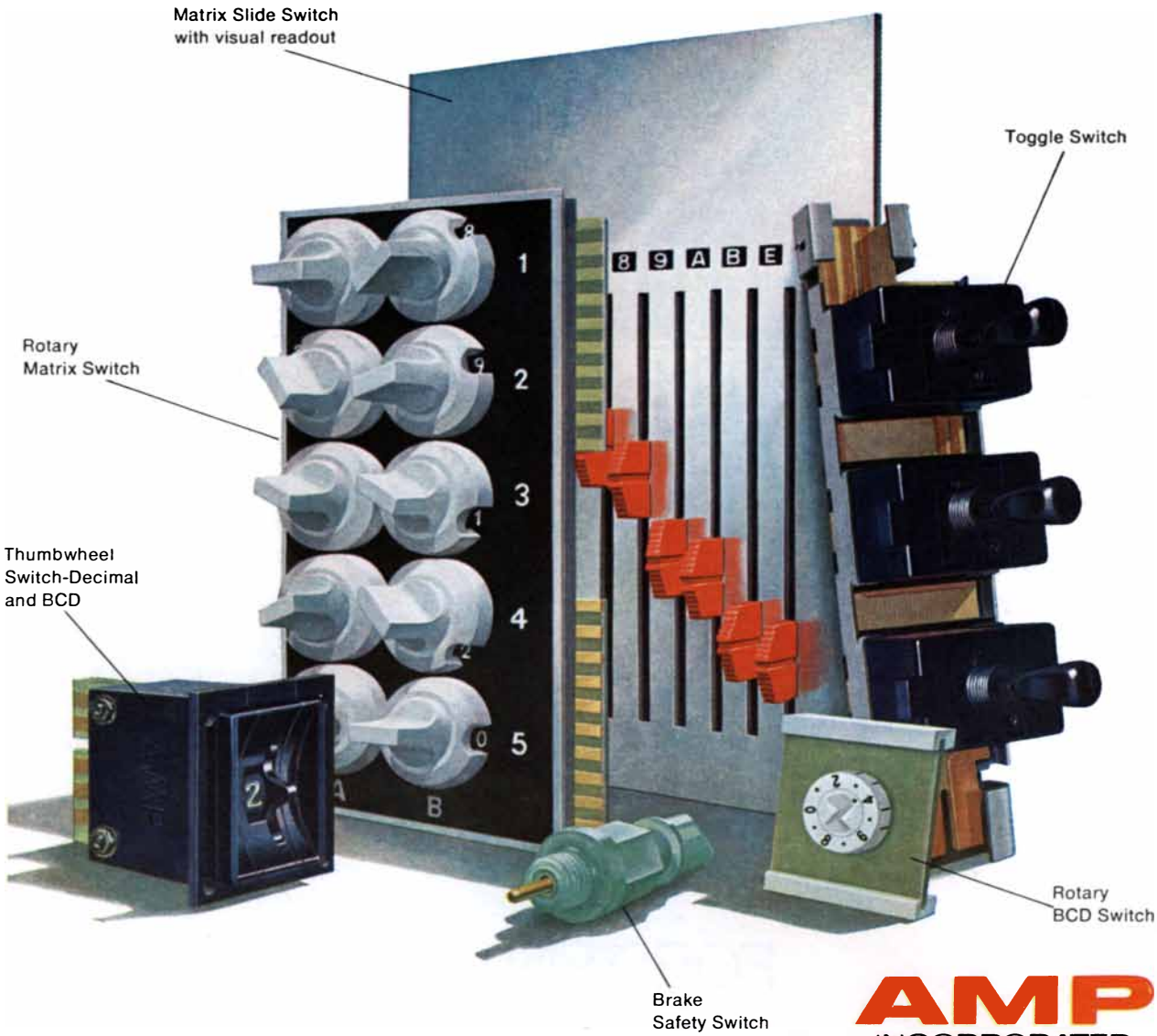
We have always had these skills, now we've applied them to switches.

Some of our switches appear to be standard . . . and some unique. The standard switches we made more reliable yet economical. The unique switches solve unique problems.

For complete information on our growing switch line write to

Industrial Division,

AMP Incorporated, Harrisburg, Pa. 17105



AMP INCORPORATED

Manufacturing and direct sales facilities worldwide: Barcelona, Brussels, Buenos Aires, Frankfurt, London, Mexico City, Paris, Puerto Rico, Sao Paulo, s'Hertogenbosch (Holland), Sydney, Stockholm, Tokyo, Toronto, Turin, Vienna.