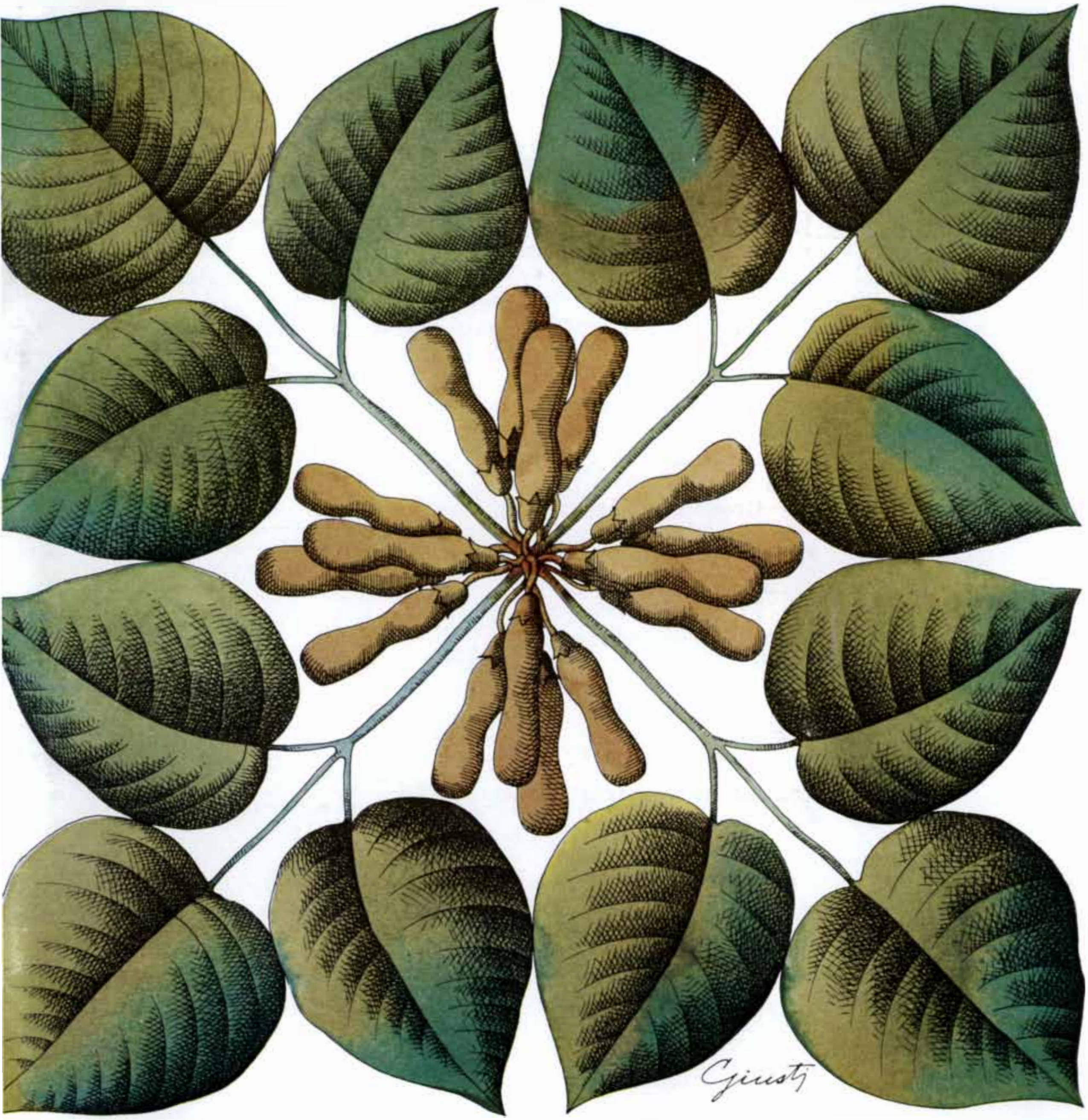
SCIENTIFIC AMERICAN



SOYBEANS

ONE DOLLAR

February 1974

What America needs is a new balance of power.

Energy shortages affect more than the corner gas pump. The petroleum squeeze threatens our petrochemical industry, raw material producer for products as diverse as bowling balls, luggage, tires and textiles.

Also threatened is our electric generating capacity, critical to industry, commerce and labor—as well as the consumer.

But these critical problems can be solved.

New exploration initiatives and technology will extend our coal, oil and gas reserves.

Our environment can accommodate new electric generating capacity from nuclear and coal-fired plants until we can harness nuclear fusion and solar power.

But one energy priority requires greater awareness: the need for a new equilibrium in our use of energy . . .

The piston engine, barely 25% efficient, consumes ⅔ of our petroleum.

Natural gas forms 1½% of U.S. energy reserves but 32% of consumption.

The energy equivalent of domestic coal reserves is twenty times that of oil and gas combined.

Greater use of electricity, derived from coal and nuclear fuels, appears both prudent and inevitable.

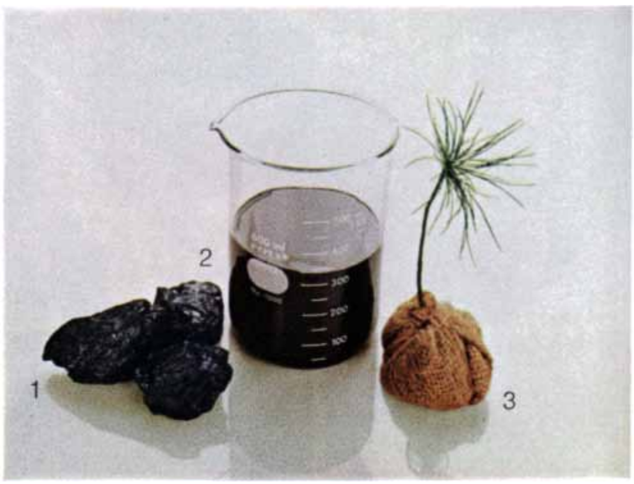
Future shifts to synthetic fuels will be highly desirable.

Integrated environmental policy is needed. (Consider automobile emissions: While new controls decrease emission rates, increased gas consumption negates much of the gain.)

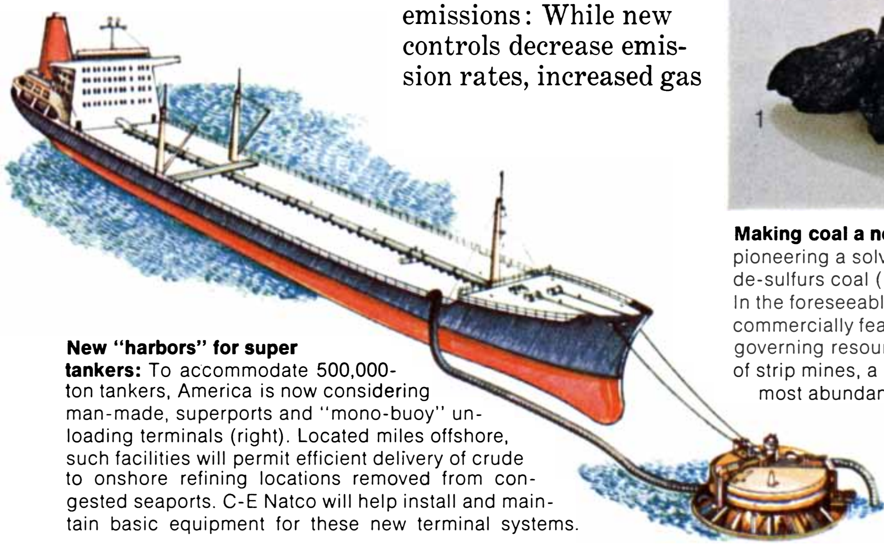
Combustion Engineering's stake in energy is as broad as the field itself. C-E manufactures fossil fuel and nuclear steam generating equipment . . . designs and builds refineries, chemical plants, natural and synthetic gas facilities . . . and produces a growing range of anti-pollution devices.

Our interests are those of the nation: achieve the highest and best energy balance for the economy . . . and the ecology.

If you'd like more facts on today's energy challenge please write Combustion Engineering, Inc., Dept. S, 900 Long Ridge Road, Stamford, Connecticut 06902.

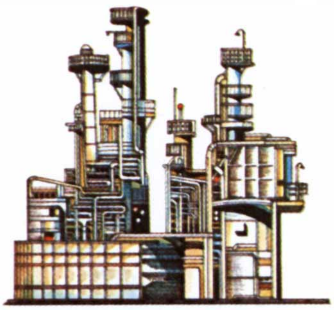


Making coal a new fuel for tomorrow: C-E Lummus is pioneering a solvent refining process which de-ashes and de-sulfurs coal (1) into clean-burning liquid fuel (2). In the foreseeable future, the process promises to be commercially feasible. In concert with new measures governing resource extraction, such as reclamation (3) of strip mines, a major new energy source from our most abundant fossil fuel is at hand.



New "harbors" for super tankers: To accommodate 500,000-ton tankers, America is now considering man-made, superports and "mono-buoy" unloading terminals (right). Located miles offshore, such facilities will permit efficient delivery of crude to onshore refining locations removed from congested seaports. C-E Natco will help install and maintain basic equipment for these new terminal systems.

The future fuels: Man is making great strides in discovering how to economically harness the "inexhaustable" fuels. By the turn of this century, we expect that both nuclear fusion and solar power will have been demonstrated practical. (Left: prototype solar furnace near Odeillo, France.) But even solar power will present environmental problems we must learn to accommodate.



Return of the prodigal refinery: Having to refine oil imports abroad aggravates an already serious balance of payments predicament. Previous import quotas and environmental concerns helped bring refinery construction to a virtual standstill in recent years. The need for dozens of new facilities, many to be built by C-E Lummus, is predicted for the years immediately ahead.

We must balance today . . . to provide for tomorrow: "Maine Yankee"—recently completed for Maine Yankee Atomic Power Company with C-E nuclear steam generating system. An excellent example of environmentally-compatible electric generating facility. Along with coal-fired plants, new nuclear capacity will help relieve the squeeze on critical fuels.

"Using less" is part of the answer: If cars getting 30 mpg were used for all trips within a 30-mile radius of "home," the nation's airlines could fly on the petroleum we'd save. Conservation and shifts to energy sources other than short-supply fossil fuels are vital factors in our new energy policy.

COMBUSTION ENGINEERING, INC.

THE ENERGY SYSTEMS COMPANY

VEGAVERSATILITY.

1. The '74 Vega Hatchback GT. This year's version of last year's *Motor Trend* Economy Car of the Year.

2. Vega Hatchbacks are the little cars you can use sort of like little wagons.

3. Our Vega truck, the Panel Express, has lots of space for the money.

4. And you can get some extra room and still enjoy economy with one of three Vega wagons. This one, with the wood-grain vinyl trim, is the Kammback Estate.

5. Vega has an improved bumper system this year and cooling slots in place of a grille. Functional as well as good looking.

6. Take our lowest priced Notchback and spiffy it up (available vinyl roof cover, Custom Interior trim, etc.) and you graduate into the Notchback LX.

7. The Vega Hatchback Hutch. With the addition of a tent-like accessory you can turn your Vega into an economy camper wherever you stop.

8. The Kammback GT. What you don't see are the special ride and handling components, the instrumentation, the wood-grain vinyl trim on the instrument panel, the GT steering wheel, the adjustable driver's seat back and a few other classy items.

9. This is what's waiting for you at your Chevy dealer's—the front seat of a Vega to try out for size.

VEGA

Chevrolet

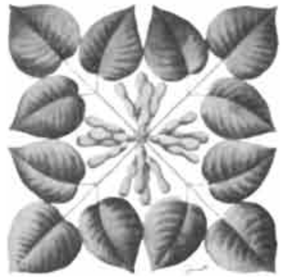
CHEVROLET MAKES SENSE FOR AMERICA.

Established 1845 **SCIENTIFIC AMERICAN** February 1974 Volume 230 Number 2

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, U.S., POSSESSIONS AND CANADA; $12.50 PER YEAR, ALL OTHER COUNTRIES. EUROPEAN SUBSCRIPTION COPIES MAILED C/O PUBLICATIONS DISTRIBUTION SERVICE, ST. KATELIJNEVEST 14, B2000 ANTWERP, BELGIË.



THE COVER

The painting on the cover is a stylized representation of a soybean plant as seen from above. A photograph of a mature soybean plant from the same vantage point would show few if any pods, and the pods would hang downward. A mature soybean plant stands more than three feet high and has a large number of pods. In recent years soybeans have become a major agricultural crop in the U.S. because the oil in the bean is in demand for many products and the residue after the oil is extracted is rich in protein (see "Soybeans," page 14). Soybeans have also become a major export and thus play an important role in balancing payments for the nation's many imports.

THE ILLUSTRATIONS

Cover illustration by George Giusti

Page	Source
15	American Soybean Association
16–20	Lorelle M. Raboni
21	U.S. Department of Agriculture (*top*), Harris H. Barnes, Jr. (*bottom*)
22–23	Fairchild Semiconductor
24–30	Ilil Arbel
31	Fairchild Semiconductor
32	Ronald Leavitt, University of California at San Diego
34–35	Bunji Tagawa
36	Garth Nicholson, Salk Institute for Biological Studies (*left*); D. Stuart and M. Sturm, from "Mammalian Cells in Culture Frequently Release Type C Virus," Lieber, M. M. *et al.*, *Science*, Volume 182, pages 56–59, Figure 1 D, 5 October 1973. Copyright American Association for the Advancement of Science (*right*)
37–40	Bunji Tagawa
48	R. B. Minton, Lunar and Planetary Laboratory, University of Arizona
50	Lick Observatory (*left*), Hale Observatories (*right*)
51	Hale Observatories
52	Tom Prentiss
53	R. B. Minton, Lunar and Planetary Laboratory, University of Arizona
54–55	Tom Prentiss
56	Jesse L. Greenstein, California Institute of Technology
57	California Institute of Technology and Carnegie Institution of Washington
58	H. E. Huxley, Medical Research Council Laboratory of Molecular Biology
59–70	George V. Kelvin
72	National Accelerator Laboratory
73–82	Allen Beechel
83	National Accelerator Laboratory
85–90	Alan D. Iselin
92–95	Dan Todd
96	Carl W. Condit, Northwestern University
97	Engineering Societies Library
98–99	Dan Todd
100	Arthur Lavine, Chase Manhattan Bank
101–102	Dan Todd
105	Skidmore, Owings & Merrill
106–108	Ilil Arbel
110–111	Gary S. Thorpe
112–115	Jerome Kuhl

SCIENCE/SCOPE

<u>The best pictures ever taken of Jupiter</u> were made by the imaging photopolarimeter (IPP) aboard Pioneer 10 and sent a half-billion miles by radio to NASA's Ames Research Center at Mountain View, CA. They clearly show the Great Red Spot and -- because the immense gravity caused Pioneer 10's trajectory to curve around Jupiter as it sped to escape the solar system -- include views from angles that cannot be seen from earth. Pioneer 10's infrared radiometer returned data on Jupiter's net energy flux and the thermal structure and chemical composition of its atmosphere. Both instruments were built by Santa Barbara Research Center, a Hughes subsidiary.

<u>A computerized pattern grading and marker making system</u> for apparel manufacturing, developed by Autographics, Inc. of Baltimore, will be marketed by Hughes. The system is available in both semiautomatic and full automatic configurations; the latter can lay out a men's or ladies' suit marker in three minutes. It can also be used to provide computer tapes for the Hughes Lasercutter, an automated cloth-cutting system. While Autographics will continue to market its system directly, the arrangement with Hughes will provide additional system and service capability in the U.S. and abroad.

<u>A new liquid crystal pictorial display system</u>, developed by Hughes for the U.S. Air Force Avionics Laboratory, promises performance superior to that of the cathode-ray tube for displaying symbolic, graphic, and pictorial television images in real time. The liquid crystal display produces no light of its own, but is viewed by natural or artificial light. The brighter the ambient light the more brilliant the display -- a distinct advantage for airborne systems. It consists of 10,000 elemental liquid crystal cells per square inch of display. A cell appears black when no voltage is applied; increasing voltage produces tones ranging from black to white. The liquid crystal display offers high resolution, is compact and lightweight, requires little power, and needs only a simple electrical interface with sensors or video signals.

<u>Circuit Design Engineers with project engineering experience</u> are needed for our commercial and government programs. Responsibilities: design and development of complex high- and low-voltage power conditioners, including command signal conditioning, logic sequencing, and analog telemetry instrumentation. U.S. citizenship required. Please send your resume to: B.E. Shryack, Hughes Electron Dynamics Division, 3100 W. Lomita Blvd., Torrance, CA 90509. An equal opportunity M/F employer.

<u>Two-way cable television has been tested for the past year</u> in El Segundo, Calif., by Theta-Com of California, a Hughes subsidiary. The Theta-Com Subscriber Response System (SRS) includes a central computer and 30 prototype terminals in subscribers' homes. Among the various types of services available: purchase of first-run movies and blacked-out sporting events, theater and airline reservations, and merchandise. Computer confirms order on the terminal's paper-tape readout and automatically bills subscriber. Subscribers can also request information, participate in games and educational programs, respond to opinion polls, and be provided with emergency intrusion, fire, or medical alarm service. The El Segundo tests will be expanded in mid-1974 to include 1,000 terminals in a comprehensive market test.

<u>Energy conservation is a potential future benefit of the SRS system</u>, which could monitor and manage power loads from a central control and substitute two-way communication from home or office for physical transportation.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

LETTERS

Sirs:

I enjoyed reading Manfred Korfmann's fine article "The Sling as a Weapon" [Scientific American, October, 1973]. As a physicist, however, I was surprised at his statement that the initial speed of the projectiles was "more than 60 miles per hour," and that they "can easily exceed 100 kilometers per hour." These figures are quite conservative.

If one neglects air resistance, the initial speed v of a projectile launched at an angle θ above the horizontal and having a range of x is given by $v = (xg/\sin 2\theta)^{1/2}$, where g is the acceleration of gravity, 9.8 meters per second squared. If we assume that 230 meters is a typical range, and if θ is 45 degrees, the minimum initial speed required is 171 kilometers per hour, or 106 miles per hour. For a maximum range of 400 meters, v must be at least 225 kilometers per hour, or 140 miles per hour. In actual practice θ may not be exactly 45 degrees and air resistance will be important, so that the initial speeds required to achieve these ranges must be even greater.

Carl T. Rutledge

Assistant Professor of Physics
Southern State College
Magnolia, Ark.

Sirs:

"The Sling as a Weapon" is very interesting, and treats a weapon not often fully appreciated. As a mechanical engineer and a lifelong sling enthusiast (more noted for distance than for accuracy), I should, however, like to take exception to the statement that the velocity of the projectile "can easily exceed 100 kilometers per hour." Even baseball pitchers can attain 160 kilometers per hour. If slings could not do better, nobody would ever have used them.

I have never measured the velocity, but with steel ball bearings of a diameter of about 2.2 centimeters (about 45 grams weight) I can probably reach at least 300 meters at an initial angle of 30 degrees, for a calculated (no wind resistance) initial velocity of about 210 kilometers per hour. The range quoted in the article of some 400 meters, even at an optimum initial angle of 45 degrees, yields an initial velocity of about 226 kilometers per hour. I can testify that it is quite difficult to shoot as steeply as 45 degrees with a lightweight projectile when one is trying to achieve full power. Because of this I am quite sure that the maximum attainable velocity must be higher than 226 kilometers per hour.

One has to be in good practice and in good shape for a full-power follow-through required for maximum distance. Whenever I try to pile in more momentum than my physical shape will allow me, I immediately get a sore shoulder.

I can also attest to the difficulty of developing good aim. Beginners invariably miss things by anywhere up to 180 degrees. I have occasionally hit what I was aiming at, but on the average I am probably no more accurate than a few degrees.

Francis de Winter

Pasadena, Calif.

Sirs:

Having been a fairly expert slinger in my youth (some 55 years ago), I was greatly interested in "The Sling as a Weapon," by Manfred Korfmann. As a boy I lived in an area near Vancouver heavily covered with second-growth forest. The area was sparsely inhabited, and so it was quite safe to use a sling or even a rifle.

Dr. Korfmann's article does not make quite clear the peculiar muscular action of slinging. It is this action that, if properly effectuated, gives the stone its high velocity (which is, I would think, well above 100 kilometers per hour).

Here is the drill. One holds the sling hanging down by one's right hand. It is swung a little like a pendulum, and then flipped into a circular path horizontally around the head, with the arm about half-extended. The motion is generated not so much by the wrist as by the entire arm, acting through the biceps.

As the stone comes around to about 11 o'clock on the second or third turn (considering the turns to be clockwise as seen from below, with 12 o'clock behind the head) one exerts pull on the stone with the entire arm (acting against the centrifugal force) so as to, as it were, snap the sling through to about the four o'clock position on the next turn, when the stone is released.

It is this last muscular heave that generates the speed, which is sufficient to make the stone go through the air with a whining buzz, even if it is quite round and smooth. (I think the stone may rotate in passage.)

As to force, we used sometimes to shoot at the walls of an old empty house sheathed with three-quarter-inch cedar. One-inch stones would go right through such wood.

We also did a good deal of archery, but my recollection is that slingstones went far, far beyond any arrow, and had far more force.

John Stanley

Professor of Population Dynamics
(Retired)
McGill University

Sirs:

Being a prehistorian and an archaeologist rather than a natural scientist, I was happy that several readers found my estimate ("that the speed of a missile leaving a sling can easily exceed 100 kilometers per hour") conservative. They have provided me with data that will be very useful.

I found Mr. de Winter's and Professor Stanley's firsthand experiences with the force of the throw most interesting. I did not mean to imply in my article that the wrist does all the work in the throw;

Scientific American, February, 1974; Vol. 230, No. 2. 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; Arlene Wright, assistant 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 Harry T. Morris, Advertising Director, 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.
U.S. and Canada:
☐ 1 year, $10 ☐ 2 years, $18 ☐ 3 years, $25
All other countries:
☐ 1 year, $12.50 ☐ 2 years, $23 ☐ 3 years, $32

NAME

NEW ADDRESS

OLD ADDRESS

the arm, the biceps and the back muscles play an important part, particularly in the final cast.

I should like to take this opportunity to correct two errors of wording in the article.

On page 38 the text should read: "In the Near East the first manufactured sling missiles were spherical. Biconical projectiles were the next to appear. They made their appearance shortly before the beginning of the sixth millennium B.C."

On page 39 it should be stated that the advantages of clay material for pellets were recognized earlier than the pottery phase of the Neolithic and not, as the article has it, "earlier than the prepottery phase."

MANFRED KORFMANN

Deutsches Archäologisches Institut
Istanbul

Sirs:

I am pleased to learn of Michael Crichton's interest in the dynamic aspects of block pictures ["Letters," SCIENTIFIC AMERICAN, January]. Studies of various perceptual effects of moving images that have been coarsely quantized are important to achieving improved understanding of visual perception.

Although my article reporting face-recognition experiments [SCIENTIFIC AMERICAN, November, 1973] was restricted to single-frame pictures, I have been studying dynamic aspects for some time now. Lateral motion, as Dr. Crichton observes, enhances recognition. Additionally I have found that two-dimensional expansion and contraction, as in zooming, is particularly effective. Color contrasts and gradients also produce some strange and wonderful effects leading to image perception that is surprisingly detailed.

Rather than using the older techniques of computer processing, I now employ a real-time optical system that is fast, inexpensive and works as easily with color as with black and white. (The George Washington and Mona Lisa pictures in the November issue were produced this way.) This new tool readily allows motion-picture production in large quantities and is turning out to provide a technique that is both instructive and quite fascinating.

LEON D. HARMON

Department of Biomedical Engineering
Case Western Reserve University
Cleveland

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

FEBRUARY, 1924: "The recent discovery of hafnium, the latest addition to the list of chemical elements, was the result of investigations based on the latest and most advanced conceptions of atomic structure. The list of elements arranged in the order of their atomic number showed a break after No. 71, the element lutecium. No. 72 was lacking. D. Coster and George de Hevesy, working in Copenhagen, deduced that the unknown element would probably show great resemblance to element No. 40, zirconium, to which, according to the theory of Niels Bohr, it must be closely related. The two investigators examined the X-ray spectra of all zirconium minerals and in each case found, in addition to the characteristic lines of the known element, lines of another, unknown element in the position where the lines of element No. 72 should be. The two scientists succeeded in separating the new element from the zirconium and named it in honor of Copenhagen (called *Hafnia* in its Latin form)."

"The results of two expeditions, one from Canada and the other from the Lick Observatory, to Wallal, Australia, for the recent solar eclipse have now come to hand. Both report in favor of the Einstein shift of starlight. In each case the number of stars measured was very large, exceeding 80. The measures were sufficiently exact to give a decisive result. The values for the shift at the limb of the sun deduced from the individual plates ranged from 1.59 inches to 1.86 inches, the mean of all being 1.74 inches, which is only .01 inch less than Einstein's predicted value."

"A discovery recently made in southern France shows that if Stone Age men were still ignorant of the art of making pottery by baking earth, they had at least discovered its plasticity and used it to model figures of animals. A young French archaeologist, M. Norbert Casteret, entered what appeared to be the mouth of a cave and swam through a subterranean stream until he reached a great gallery containing numerous works of prehistoric art. The sculptures–figures of animals modeled in clay–included a bear, three tigers, three horses and 20 modelings of uncertain character. On the rocky walls are engravings, made with filing implements, of prehistoric animals and of others that disappeared from France thousands of years ago. A study of the sculptures points to their being the work of the Aurignacians, who lived 25,000 years ago. Hitherto, only one discovery of prehistoric sculpture had been made, in 1912, and also in France."

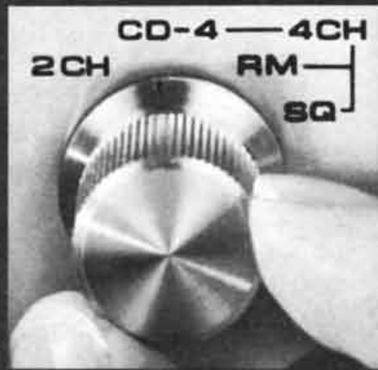
SCIENTIFIC AMERICAN

SEPTEMBER, 1874: "If we look around at the condition of the planetary system, we find much to lead us to the belief that it grew to its present state, that there was a process of development. There are eight primary planets and 134 asteroids, and all of these bodies travel in the same direction around the sun. Every one of the bodies whose motion has been determined turns in the same direction. There are in fact so many similarities that we are bound by the laws of probability to believe in the evolution process. An explanation of this motion is that there was a great nebulous mass extending far beyond the present path of the outermost planet. The mass was rotating, and as it contracted it began to rotate more rapidly. A ring was thrown off and in the course of millions of ages was amalgamated into one mass, having the same direction of motion that the nebulous mass had and traveling around the center, which was the sun. This process went on until one planet after another was formed."

"Information has recently reached England of the decease of Dr. Livingstone, the celebrated African explorer, during June last. In his death geographical science loses one of its most persevering students. It may be truly said that for a blank spot on the map of Africa he has substituted a country rich, fertile and productive, which before many years will exercise no small effect on the commerce of the world. In 1868 Dr. Livingstone entered a region totally unknown to civilization. Until found by the *Herald* reporter Stanley some two years ago, little was heard from him. After Stanley's departure he continued his exploration, but no news of him was received until the present time, when the British officials at Zanzibar transmitted the intelligence of his death."

"Studying the solar spectrum around the beginning of this century, the elder Herschel passed a sensitive thermometer through the successive colors and observed that the greatest heating effect was not coincident with the brightest illumination but at a distance beyond the red, where no rays were visible. The inference he drew from these observations was that the heating rays were distinct from the luminous rays. By the use of photographic papers it was subsequently ascertained that the chemical action of the sun ray appeared to be greater toward the violet end of the spectrum. Hence arose the belief, which the scientific world has generally entertained of late, that the solar radiation was triple in constitution, being separated into a visible spectrum, a heat spectrum and a chemical spectrum. Dr. Draper, in a memoir, points out Herschel's error and also proves conclusively that every part of the spectrum can produce chemical changes. As a consequence the supposed triple constitution of the sun ray must be dropped. There is in the sun ray neither light nor heat nor chemical power as such but simply vibrations, which, when stopped, may manifest themselves in one or another of these phenomena. 'The evolution of heat, the sensation of light, the production of chemical changes are merely effects, manifestations of the motions imparted to ponderable atoms.' "

"The celebrated Siamese twins, who for the past half century have been the foremost of living curiosities, recently died at Salisbury, N.C. These remarkable personages were born in Siam in 1811. Captain Abel Coffin of Newburyport, Mass., found them in the city of Meklong and bought them from their mother. In 1850 Barnum exhibited them throughout the country. Chang and Eng were linked together by a fleshy ligature, an extension of the sternum, about four inches in length and two in breadth. The two bodies had but one navel, which was in the center of the band. Chang's death occurred first; the shock, or more probably the cessation of circulation, affected Eng so strongly that delirium, followed by stupor, almost immediately set in. At the end of two hours he also expired. The brothers had married, Chang having six children and Eng five. The autopsy of the bodies has been consented to by the relatives, and the remains have been transported to Philadelphia."



THE AUTHORS

FOLKE DOVRING ("Soybeans") is professor of agricultural economics at the University of Illinois. The son of a Swedish poet, he studied first in the humanities. As an economist he has served the United Nations and other international agencies, both before and since going to the University of Illinois in 1960. The many books and articles he has written range in subject matter from the land systems of medieval Europe to land reforms in present low-income countries. His most recent book, *The Optional Society*, treats what he regards as a basic socioeconomic problem of modern affluent nations, the lack of firm constraints to tell people what to do. "Whether my research concerns India or Mexico, the U.S.S.R. or the U.S.," he writes, "I always view economic problems in their context—the total economic system but also the cultural background of a people and their ways of communicating. My main hobby is studying languages; being familiar with most of those in Europe, including Russian, I now work at Chinese. I want to face the whole width of the human phenomenon."

GILBERT F. AMELIO ("Charge-coupled Devices") is director of the development of charge-coupled devices in the Research and Development Laboratory of the Fairchild Camera and Instrument Corporation. He received his bachelor's, master's and doctor's degrees in physics at the Georgia Institute of Technology in 1965, 1967 and 1968 respectively. From 1962 to 1965 he worked for Information Sciences, Inc., of Atlanta. From 1966 to 1968 he was with the physical sciences division of Georgia Tech's Engineering Experiment Station. As a private consultant during that period he participated in the development of medical electronics for heart research and of scanning spectrophotometers. From 1968 to 1971 he was on the staff of the Bell Laboratories.

JOHN J. HOLLAND ("Slow, Inapparent and Recurrent Viruses") is professor of biology at the University of California at San Diego. Before going there in 1968 he had taught at the University of Minnesota, the University of Washington and the University of California at Irvine, where he was chairman of the department of molecular and cell biology. During a sabbatical leave in 1970–1971 he was a visiting scientist at the Institute of Molecular Biology in Geneva. Holland was graduated from Loyola University in 1953 and obtained his Ph.D. (in microbiology) from the University of California at Los Angeles in 1957. He writes of his pleasure in "California's outdoor opportunities," which he enjoys by "skiing in the high Sierras in winter and backpacking in the summer into isolated high-altitude lakes off marked trails." Closer to home he enjoys "body surfing in the Pacific and hiking in the high deserts."

FRED L. WHIPPLE ("The Nature of Comets") is Phillips Professor of Astronomy at Harvard University and senior scientist at the Smithsonian Astrophysical Observatory, of which, until he retired recently, he had been director since 1955. He went to Harvard in 1931, the year he obtained his Ph.D. from the University of California at Berkeley, as a member of the staff of the Harvard College Observatory, with which he is still associated. "My research goal," he writes, "centers on physical processes in the evolution of the solar system." His recent work, in addition to comets, has involved the origin and development of the moon and has utilized aerodynamic and dynamic theory coupled with the most recent direct measurements of lunar material from the Apollo program. Whipple is writing on comets for the second time in SCIENTIFIC AMERICAN; his first article, "Comets," was published in July, 1951.

JOHN M. MURRAY and ANNEMARIE WEBER ("The Cooperative Action of Muscle Proteins") are at the School of Medicine of the University of Pennsylvania; Murray is completing the requirements for the M.D. degree and Weber is professor of biochemistry. Murray, who has his doctorate in biochemistry, writes that he enjoys both research and clinical medicine and has not settled on one or the other as a career. "This indecision," he writes, "is compounded by the even greater satisfaction I obtain from all sorts of outdoor activities. Lately I have been particularly fascinated by scuba diving, which my wife and I took up several years ago." Weber received her M.D. in Germany in 1950 and has since been involved in research and teaching. She was at the Institute for Muscle Disease in New York from 1960 to 1965 and then at St. Louis University until her recent move to Pennsylvania. "Originally I had some difficulty deciding whether I wanted to go into science or history," she writes. "In family tradition (my father is a physiologist known for his research in muscle) I settled for science. My other interests have contracted and assert themselves only in my choice of books."

R. R. WILSON ("The Batavia Accelerator") is director of the National Accelerator Laboratory. After receiving his Ph.D. from the University of California at Berkeley in 1940 he went to Princeton University, where in association with Enrico Fermi he made some of the early measurements of the neutron-absorbing properties of uranium 235. At Princeton he invented the isotron method of separating isotopes of uranium and soon was put in charge of the university's atomic energy project. The group moved to the Los Alamos Scientific Laboratory in 1943, and there Wilson was director of the cyclotron group and the research division. For 20 years until receiving his present appointment in 1967 he was director of the Laboratory of Nuclear Studies at Cornell University, where he and his colleagues built a progression of electron synchrotrons.

JOHN D. FERNSTROM and RICHARD J. WURTMAN ("Nutrition and the Brain") are in the department of nutrition and food science at the Massachusetts Institute of Technology; Fernstrom is assistant professor of nutritional biochemistry and metabolism and Wurtman is professor of endocrinology and metabolism. Fernstrom received both his bachelor's degree and his Ph.D. at M.I.T. Thereafter he was a research fellow at the Roche Institute of Molecular Biology for a year before returning to M.I.T. Wurtman, who received his M.D. degree from the Harvard Medical School in 1960, spent two years on the medical service at the Massachusetts General Hospital and then worked at the National Institute of Mental Health before going to M.I.T. in 1967.

CARL W. CONDIT ("The Wind Bracing of Buildings") is professor of history, art history and urban affairs at Northwestern University. "The triple appointment," he writes, "arises from the fact that my courses are part of the major offerings of the three departments." He teaches courses in building and planning in Chicago, and in the history of science, of building techniques and of the city. His education was similarly diversified: he studied engineering at Purdue University, English and intellectual history at the University of Cincinnati and the history of science at the University of Wisconsin. Condit is a founder of the Society for the History of Technology.

Now! Preview the great new

Deutsche Grammophon

recording of the complete

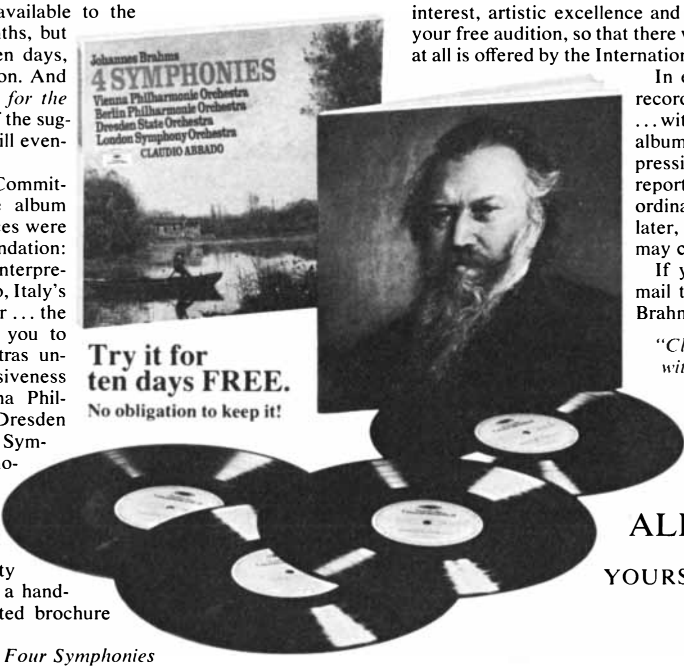
FOUR SYMPHONIES OF BRAHMS

superbly conducted by CLAUDIO ABBADO

ALL 4 RECORDS FOR THE PRICE OF ONE

plus free auditions of important new recordings months before they are available to the public!

BRAHMS' *Four Symphonies*—most magnificent since Beethoven's! The extraordinary new recording of them from Deutsche Grammophon offers Claudio Abbado conducting *four* of the world's greatest orchestras. This album will not be available to the American public for many months, but you may preview it now for ten days, without charge, without obligation. And you may keep all *four records for the price of one*—yes, a full 75% off the suggested list price at which they will eventually be sold in stores!



The Carnegie Hall Selection Committee's report accompanying the album points out why these performances were singled out for their recommendation: freshly conceived, deeply felt interpretations by young Claudio Abbado, Italy's foremost present-day conductor . . . the unprecedented opportunity for you to enjoy and compare four orchestras unsurpassed in precision, responsiveness and tonal splendor—the Vienna Philharmonic, Berlin Philharmonic, Dresden State Orchestra and London Symphony . . . plus Deutsche Grammophon's incredibly lifelike sound. The four stereo records themselves are superior, silent-surfaced European pressings, manufactured to highest quality control standards, and come in a handsome hinged case with illustrated brochure and *exclusive* Committee report.

Our preview offer of Brahms' *Four Symphonies* is limited! So please mail the coupon on this page today. Listen for ten days. Then either return the album and pay nothing, or keep it for only $6.98 plus small postage/handling charge (and sales tax where required). This is a saving of *$20.94 off* the suggested retail price when the regular edition is finally released.

How you can be first to enjoy the greatest new recordings!

By replying now, you assure yourself of the privilege of receiving, for future ten-day free previews, special *advance editions* of major new albums . . . chosen *exclusively* for the International Preview Society by the Carnegie Hall Selection Committee. Since the Society is *not* a record "club", you will never be required to buy any album, nor will you receive a "record-of-the-month" like clockwork. Only when the Committee discovers a new album of the most extraordinary musical interest, artistic excellence and technical quality will it be sent for your free audition, so that there will often be months when no album at all is offered by the International Preview Society.

In every case, you will preview these records *long before* they reach the stores . . . with the option of purchasing any album you wish, in a superb imported pressing, including exclusive Committee report . . . never at a higher price than the ordinary edition will retail for months later, and frequently at *big savings!* You may cancel this arrangement at any time.

If you truly love great music, please mail the coupon today for your exciting Brahms' *Four Symphonies* preview.

"Claudio Abbado has his own way with Brahms, a beautiful way. He endows all of his readings with an emotional quality tempered by a sense of stylistic elegance."

—Julius Bloom

EXECUTIVE DIRECTOR/CARNEGIE HALL

ALL FOUR RECORDS

(regular list price $27.92)

YOURS FOR THE PRICE OF ONE

$6.98

(plus small postage/handling charge) if you decide to keep them after your free audition!

SEND NO MONEY

SIMPLY MAIL ATTACHED CARD TODAY

If card is missing write to:

THE INTERNATIONAL PREVIEW SOCIETY

Dept. DSA

333 North Michigan Avenue

Chicago, Illinois 60601

Only new members eligible for this offer. Offer limited to U.S. and its possessions.

Established 1845 SCIENTIFIC AMERICAN February 1974 Volume 230 Number 2

Soybeans

Now in a class with wheat and not far behind corn as a major U.S. crop, they have become one of the nation's main exports. They are thus playing an important role in balancing payments for imports

by Folke Dovring

In 1930 soybeans were harvested from a million acres of farmland in the U.S. and the crop amounted to about 14 million bushels. Last year 56 million acres were in soybeans and the harvest is estimated at more than 1.5 billion bushels. Even more remarkable is the expanding role of soybeans in U.S. exports. As recently as 1950 soybean exports were negligible; last year they had a value exceeding $3 billion, or about 5 percent of all income from exports. This year soybeans may account for an even larger part of U.S. exports, thereby becoming a critical factor in the nation's balance of trade as a source of the money needed to pay for such essential imports as petroleum.

The economic value of soybeans lies in their protein and fat (extracted as soybean oil). Even though the fat content of soybeans (18 to 20 percent) is less than that of most other oilseed crops, the output of fat per acre is often higher than that of other oilseeds because soybeans have a higher yield. Oil alone, however, would not have enabled soybeans to compete with other crops for farm acreage in the U.S.; of equal importance to the fat content is the protein content, which is as high as from 40 to 45 percent of the weight of the beans, an unusually large fraction for any crop.

The soybean is a relative newcomer to America. The plant originated in China, but (unlike many other crops) it did not spread around the globe during the era of geographical discovery. Until the early part of this century nearly all soybeans were grown in eastern Asia. The soybean is a staple food in the mainly vegetable diets of China, Japan and Indonesia but for a long time seemed to have little to contribute to the meat-oriented diets of the West, where it was used chiefly as soy sauce for Chinese cooking and for British ketchups. Attempts to grow soybeans in Europe have met with scant success because of the climate, which is characterized by drought in early summer. Efforts have also been made since late in the 19th century to grow the crop in Russia, but the U.S.S.R. has had a measure of success only in its Far East region.

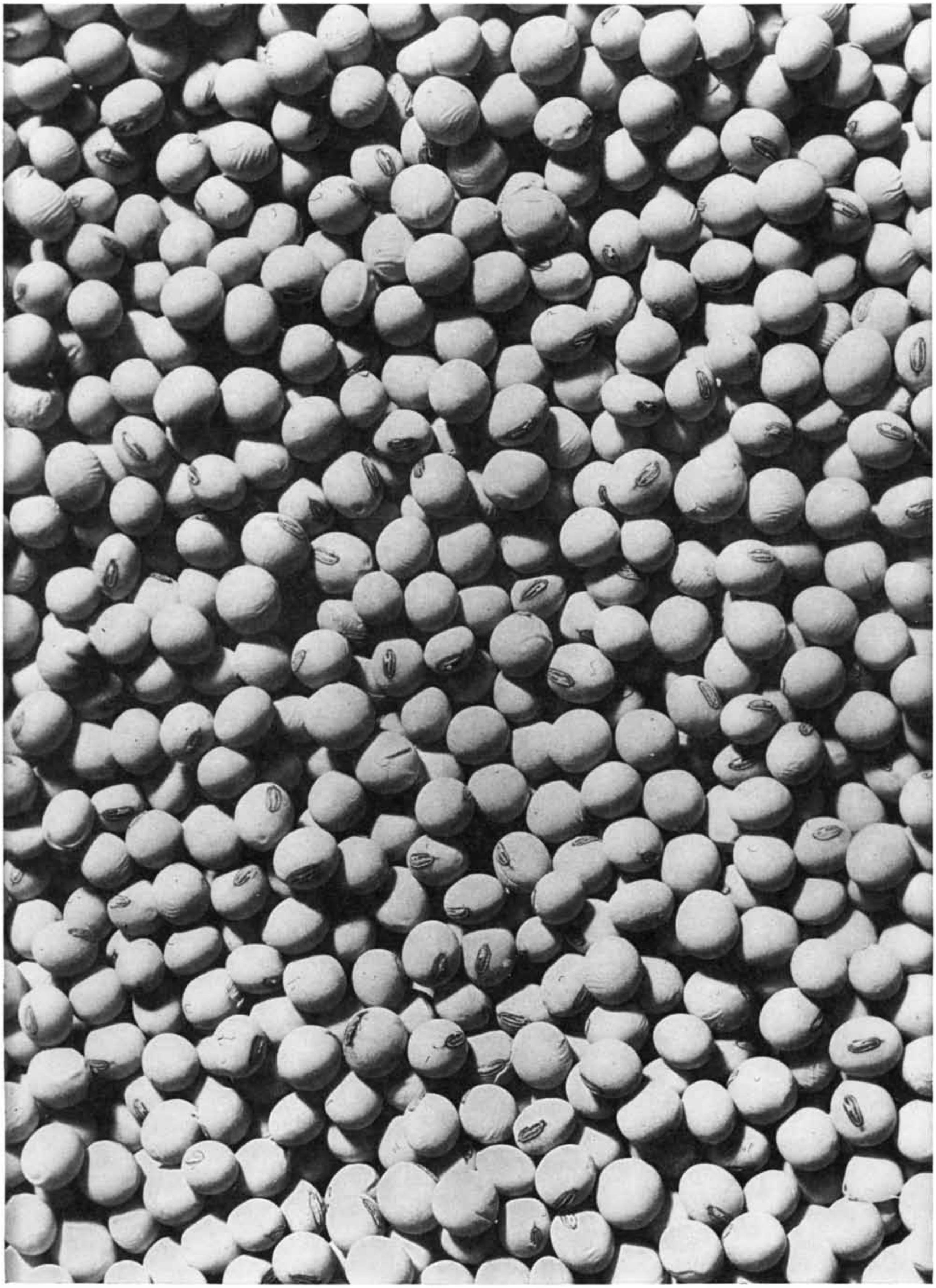
The climatic possibilities for introducing soybeans on a large scale lay elsewhere. Both North and South America have large areas with predominant summer rains, as in eastern Asia. The wetter parts of southern Asia and tropical Africa also qualify as potential soybean country. Soybeans can also be grown under irrigation in any climate with warm summers.

Soybeans remained of small importance in the U.S. until World War II brought shortages of butter. With soybean oil going increasingly into margarine, the harvested soybean acreage rose to 10 million in 1943, 30 million in 1964 and 40 million in the late 1960's. The 56 million acres in soybeans last year put the crop in a class with wheat and not far behind corn.

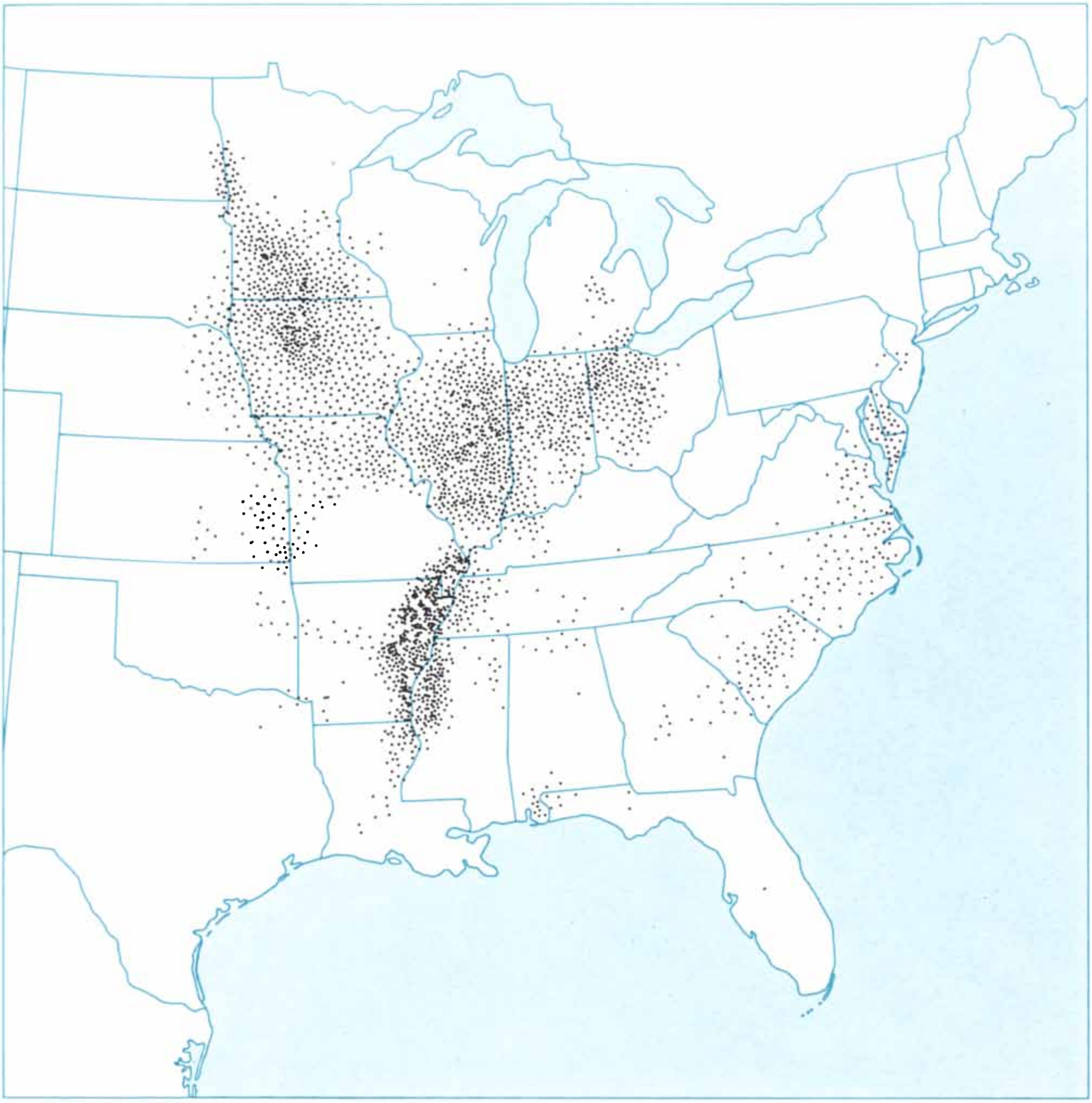
Soybeans first began to spread in the corn belt. In recent years the crop has become increasingly important in the lower Mississippi Valley and the Southeast [*see illustration on page 16*]. Thirty states have a significant acreage in soybeans; the first seven in order of contribution to the 1973 crop are Illinois, Iowa, Indiana, Missouri, Minnesota, Arkansas and Ohio. Soybeans are not much grown in the Western states for the same climatic reason as in Europe.

The soybean is a leguminous plant. Like other members of the family, such as peas and clover, it can harbor nitrogen-fixing bacteria on its roots and thus supply part of its own fertilizer. Although the yield of soybeans is much lower than the yield of corn and wheat in the high-rainfall areas where most soybeans are grown, the price per bushel is higher. As a result soybeans have in recent years surpassed wheat as a cash crop. With the prices that are now being paid, the 1973 crop of soybeans may prove to rival the corn crop in value.

All of this has made the U.S. the world's major producer of soybeans [*see top illustration on page 18*]. About three-fourths of all the soybeans produced are grown in the U.S., and the bulk of the soybeans moving in world trade comes from the U.S. The second-largest producer is still China, but its production has not increased much in recent decades and only a minor part of the crop is exported. The only other exporter of consequence is Brazil, which



HARVESTED SOYBEANS tend to be globular in shape and frequently have distinctive dark markings. They are approximately the size of peas. The harvest of soybeans in the U.S. in 1973 is estimated at between 1.5 and 1.6 billion bushels from 56 million acres.



SOYBEAN-GROWING AREAS of the U.S. are indicated. Each dot represents 10,000 acres planted to soybeans as determined by the 1964 census of agriculture. A map based on later data, which are not yet fully available, would show an increase in soybean-growing in the South. Soybeans are not extensively grown in the western part of the country because the summer droughts inhibit the crop.

in recent years has developed a big soybean enterprise and exports a substantial part of the crop. Other major soybean countries such as the U.S.S.R., Japan and Indonesia each produce less than a million tons per year; Japan has a large net import need of soybeans. Many other countries, notably Mexico and Colombia, have begun to grow soybeans commercially, but their output is not yet significant in world markets.

The expansion of the soybean business in the U.S. has had a profound impact both on the domestic food economy and on foreign trade. A comparison of the trends in butter and margarine illustrates the role that soybeans have come to play in the domestic food economy. In 1940 the ratio of butter to margarine in domestic consumption was 7 : 1. Now margarine is in the lead by a ratio of 2 : 1. Most of the increased output of margarine is made with soybean oil, which now accounts for three-fourths of the country's margarine production.

Even so, margarine takes only about a fifth of the soybean oil produced in the U.S. Other foods such as salad oil and shortening are made with soybean oil. Soybean oil also goes into a number of nonfood products and into the export trade. Soybean oil now accounts for more than half of all the food fat produced in the U.S. In the world as a whole soybeans have become the largest single source of edible fat, ahead of butter and also ahead of any other oilseed. Soybeans deliver some 15 percent of all world supplies of edible fat, and about a fourth of all the edible fat in international trade is soybean oil or the fat in whole soybeans.

Because of the rather low fat content of soybeans, the part of the bean that remains after the fat has been extracted is usually worth somewhat more than the oil, even though the oil brings a higher price per unit of weight. Soybean cake and soybean meal for feeding livestock are the chief products made from what is left after the oil is removed. Leaving aside feed grains such as corn and sorghum, the soybean products now make up nearly half of all commercial feedstuffs sold in the U.S. They have an even larger role in U.S. exports of livestock feed. For the rising supply of all kinds of meat, particularly poultry, the swelling flow of soybean cake and meal is of the greatest importance. U.S. exports are also important in the expansion of animal husbandry and the meat supply in Western Europe and (more recently) Japan.

In foreign trade soybeans over the past five years or so have become the largest single item among U.S. agricultural exports, exceeding in value both wheat and corn. The devaluation of the dollar early last year, combined with a strong demand for farm products, brought the prices of such products to unprecedented levels. The soybean harvest of 1973 proved when finally added up to meet the expectation of from 1.5 to 1.6 billion bushels; if 40 or 50 percent of it is exported at current prices, soybeans could account for 10 percent of the value of U.S. exports in 1974.

The present high level of demand for American soybeans may prove to be a temporary extreme condition. What are the prospects for the longer term? Could soybeans come to occupy an even more strategic position in the national economy? The answer will depend on two trends: the trend toward a higher yield per acre of soybeans and the trend toward the more intensive consumption of soy protein in human foods, either as "artificial meat" or in other forms.

Soybean biology has been and is the subject of considerable work in the U.S. Varieties have been developed that are suitable for cultivation in various parts of the country. Yields have been slowly but steadily increased.

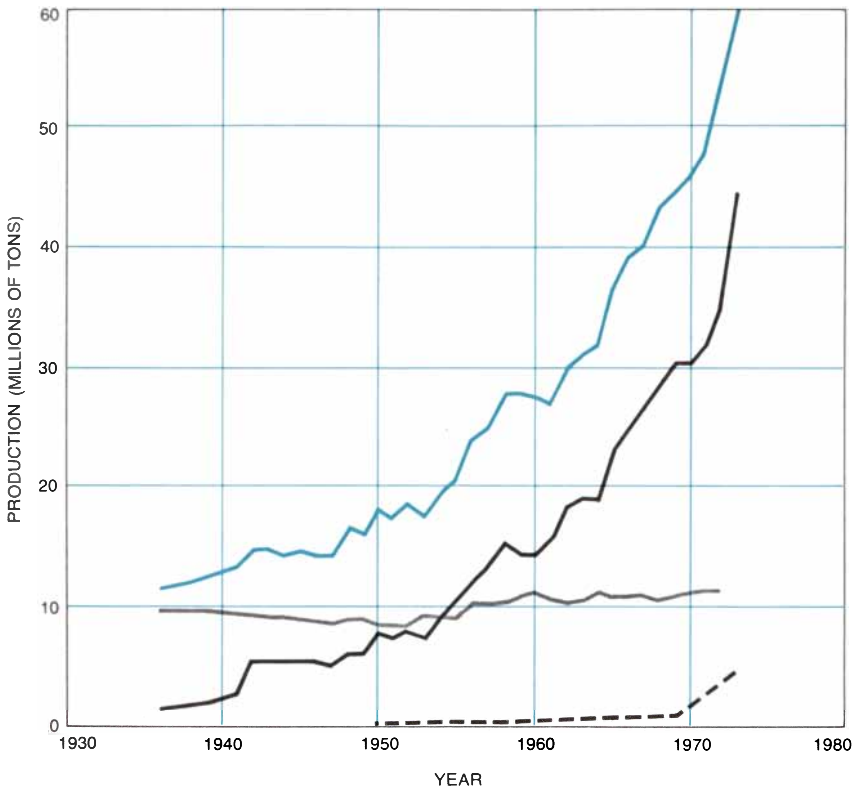
Other countries besides the U.S. have benefited from this work, including Mexico and India, which both recently imported from the U.S. soybean strains that proved to be suitable for local conditions. The results are particularly promising in India, where certain American varieties have turned out to give higher yields than they do in the U.S. The potential for soybean production in India appears to be large.

Even though soybean yields are generally higher than those of other oilseeds or of other dry pulse crops such as peas and beans, the consensus among soybean biologists is that the yield potential of the soybean is far from exhausted. It may well prove possible to increase the yield of soybeans per acre to two or three times the levels now reached in the U.S. If such results are achieved, the scope for growing high-protein food will be correspondingly larger. The perspective is plausible; what remains uncertain is when it may materialize.

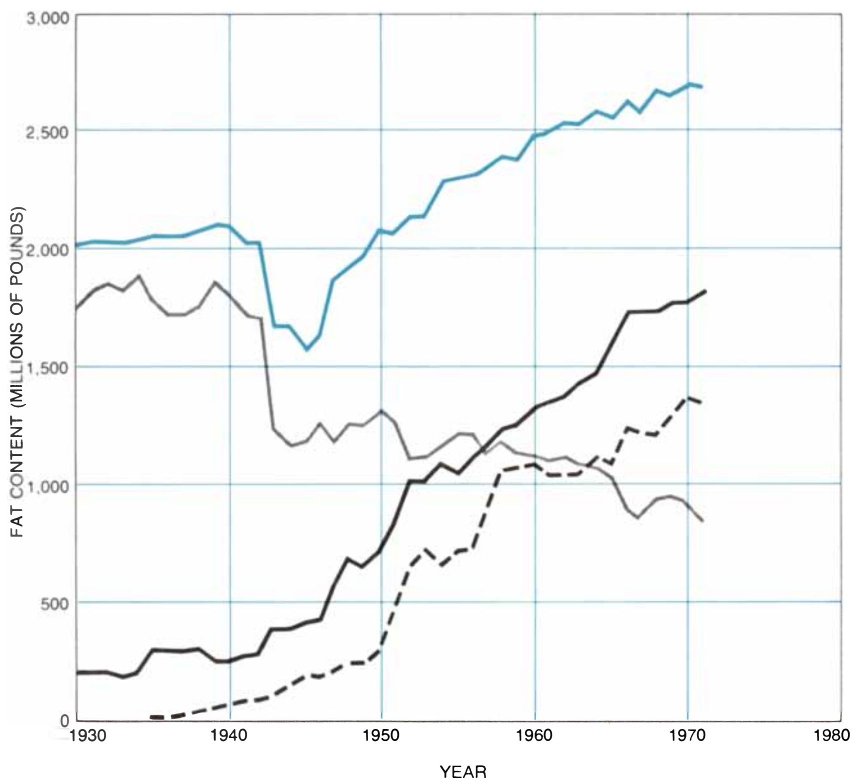
The perspective of producing soy-protein food for direct human consumption is not just a thing of the future. Meat analogues and other soy-protein dishes that many people find tasty are already on the market. Widespread acceptance by consumers may be only a matter of time. The commercial promotion of soybean recipes and a conscious economic policy favoring soybeans as human food will also be important.

The manufacture of imitation meat from soybeans rests on a technology that has been under way for some time. Soy-protein isolates have been available commercially in the U.S. since about 1960; by 1970 the production was some 20,000 metric tons per year. The protein isolate is spun into fibers, fashioned to resemble whatever kind of meat it is intended to

SOYBEAN PLANT is portrayed with a leaf array that has been modified so as to make the pods visible. A mature plant, standing three feet high or more, would have more stems and leaves and the pods would be hard to see. A mature pod with beans is shown at bottom.



PRODUCTION OF SOYBEANS for the past four decades is charted. Curves portray world production (*color*) and yields of the U.S. (*black*), China (*gray*) and Brazil (*broken line*).



BUTTER AND MARGARINE have changed positions in the U.S. since 1930, accounting for a large rise in the consumption of soybean oil for making margarine (*broken curve*). The total consumption of butter and margarine has risen (*color*), but margarine (*black*) now has a larger share of the market than butter (*gray*), a reversal of the earlier relation.

resemble and flavored accordingly. The name for the fibers is "textured vegetable protein," usually shortened to TVP. As defined in a regulation issued by the U.S. Department of Agriculture, the material must have "a structural integrity and identifiable texture such that each unit will withstand hydration and cooking and other procedures used in preparing the food for consumption." In other words, it must not cease to look and taste like meat if it is boiled or fried.

A marketing test of imitation bacon strips conducted in Fort Wayne, Ind., in 1970 showed ready acceptance of the soybean product by the public. The price on an "as served" basis (that is, after cooking, which in natural bacon results in a loss of weight) is about half the price of natural bacon. At present imitation meat is available to the general public in two forms: imitation bacon bits, which are flavored and colored to resemble bacon, and unflavored textured vegetable protein. The bacon bits are recommended as a spice or an additive in dishes where natural bacon would be appropriate. Unflavored textured vegetable protein serves as a meat extender in hamburgers and other dishes where ground meat is a main ingredient. Here too the soy product costs about half as much as real meat.

The manufacturers of these meat substitutes have not seemed to be rushing to capture a large share of the meat market. Part of the explanation for their caution appears to be the fact that certain soy products put on the market a number of years ago without careful development gave soybean foods a bad name. It may also be that the manufacturers want to avoid conflict with the farmers and food processors involved in meat production—some of whom are members of the same companies or conglomerates as the manufacturers of textured vegetable protein.

Whatever the reasons for the cautious attitude of the manufacturers toward meat analogues, the fact is that the American public is already eating soy products to a larger extent than is indicated by the production of textured vegetable protein. Breakfast foods, biscuits, various kinds of bread, sausages and many other foods contain soybean meal in less expensive forms than textured vegetable protein. As of 1970 such foods consumed 10 times as many soybeans as the manufacture of textured vegetable protein. American food scientists have recently perfected a soy beverage that could substitute for milk, and they have also improved the taste of soybeans for consumption as whole beans (in bean salad, for example), as flakes in baby food and in other forms.

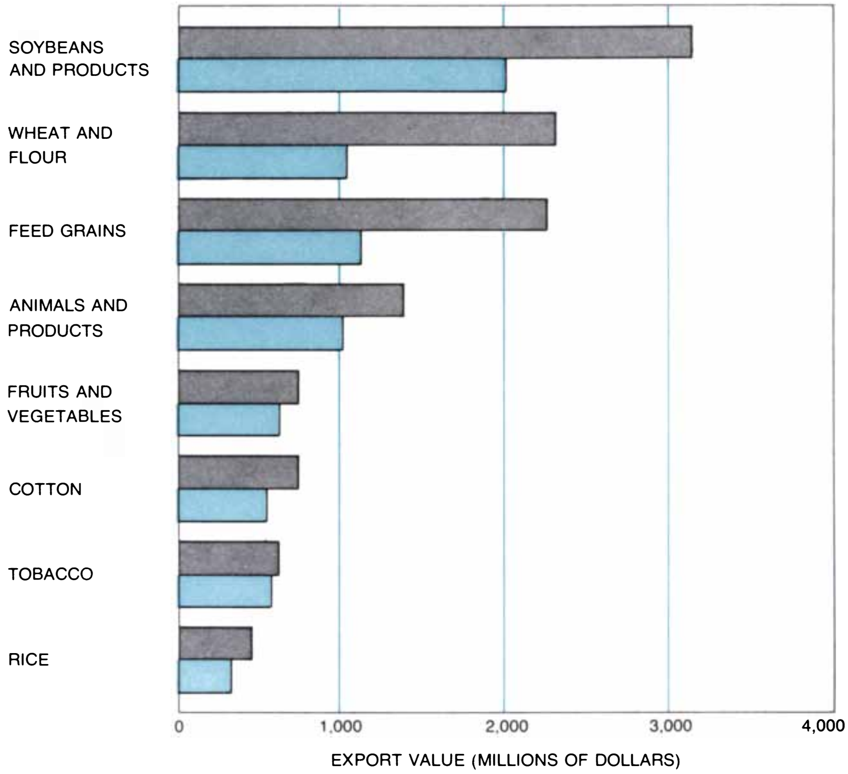
Religious groups and health-food groups with an interest in vegetable-based diets have developed many more soybean products and recipes than are available from standard sources. As a result a large variety of meat analogues can be found in specialty stores. Moreover, anyone who likes to experiment in his own kitchen can discover more uses for textured vegetable protein than the major manufacturers offer. For example, if the unflavored product is mixed with a strongly flavored bacon analogue, the blend serves with good effect in recipes calling for minced meat.

Soy protein is nutritionally somewhat less complete than meat, but the few deficiencies can be made up easily from other vegetables, including corn. A complete diet without animal products is therefore readily attainable. Many more meat imitations than are currently available could be made from soybeans, including analogues of several boneless meat and poultry cuts.

A move toward a more vegetable diet would have several advantages for the public, such as lowering the cost of living and reducing the ingestion of fat. It would lessen the pressure of economic activity on the environment, since the current energy crisis is symptomatic of an impending scarcity of resources that would be greatly relieved if less land were used to feed domestic animals.

The production of natural bacon, for example, requires 10 times as much cropland as the same quantity of bacon analogue from soybeans; beef production calls for from 15 to 20 times as much land as is needed for soy protein. Moreover, since soybeans can obtain their own nitrogen fertilizer, they draw less on the energy-intensive industries that manufacture commercial fertilizers. The ecological pressure from high levels of application of nitrogen fertilizers would also diminish, because less nitrogen would leak into the ground water and the streams.

For countries with a low level of per capita income, foods based on vegetable protein offer even greater advantages. Although at present in the poorest countries soybeans are best consumed directly (or in such forms as flour additives and "soy milk"), in the longer term meat analogues could become highly important in economic development. Experience shows that when per capita income rises, the demand for meat goes up. It rises most rapidly at certain intermediate levels of income (between about $400 and $1,600 per capita per year in current prices). "Inflation feeds on red meat" is an expression that originated in France in the 1950's when the country was on such a middle-income level. The expression has particular point today in such countries as Argentina and Mexico. Sooner or later, if all goes well, the problem will come up in the present low-income countries. It would be far less difficult economically for those countries to plan toward large supplies of vegetable protein than to try to build up meat production on a scale that would satisfy the aspirations of their people, since vegetable protein is so much more sparing of resources. Indeed, the food situation of the entire world would benefit from a greater reliance on vegetable protein.



AGRICULTURAL EXPORTS by the U.S. are led by soybeans and soybean products. Bars represent exports of major commodities in the years ending June 30, 1972 (*color*), and June 30, 1973 (*gray*). Soybeans now account for more than 5 percent of all U.S. exports.

For the U.S. the question of agricultural priorities is closely tied to issues of foreign trade. In recent years the U.S. has become increasingly dependent on imports of petroleum, particularly from the Arab nations of the Middle East and North Africa. Apart from the political embarrassment of being dependent on such distant sources, the nation confronts the impact of petroleum imports on the U.S. balance of payments. The balance-of-payments situation is quite bad and was deteriorating until the early part of 1973. A modest surplus of payments in recent months has redeemed only a tiny fraction of the huge amount of dollars held in foreign countries. At present it appears that agricultural exports, with soybeans foremost among them, may bail out the U.S., at least for a time. The question is how long they can continue to do so. The answer lies, among other things, in the trends of consumption of meat and of energy.

Both trends have been upward, and since the early 1960's they have risen at rates exceeding those of the previous several decades [*see illustration on next page*]. These two indicators are sure signs of rising standards of living. It is clear from the figures that Americans in general have become better off over the past dozen years and that the improvement has been faster than in earlier decades.

These developments would be all to the good if it were not for the problem of resources that underlies the rising figures for meat and for energy consumption. The problem is intensified by the fact that beef accounts for much of the increasing consumption of meat. At present the U.S. is paying for its imports (including meat as well as petroleum) by exporting soybeans, wheat, corn and other raw agricultural products. If the upward trend in meat consumption continues, however, more of the nation's crops will be needed to feed meat animals and less can be exported. If energy consumption continues to rise, petroleum imports can be expected to rise too.

In other words, the trends in meat consumption and energy consumption are on a collision course. If they continue upward, the two together will bring the U.S. into increasing difficulty. To put the matter rather crudely, the American public may have to choose between steaks and pleasure trips. A turn toward mass transit and other means of economizing on the consumption of energy could bring down the need for importing petroleum. In addition, however, a shift toward meat analogues and other high-protein foods based on vegetable sources—soybeans foremost among them—could reduce the requirement for feed crops within the country and allow for higher exports of agricultural products to pay for petroleum and other essential imports while keeping the foreign-trade account in balance.

Agricultural products have historically been a large part of U.S. exports. Even now the U.S. has the largest potential for agricultural production of any nation. (Brazil could become a match for the U.S. in the future, but that remains to be seen.) It is logical for the U.S. to employ its comparative advantage by paying for imports with the proceeds from farm exports.

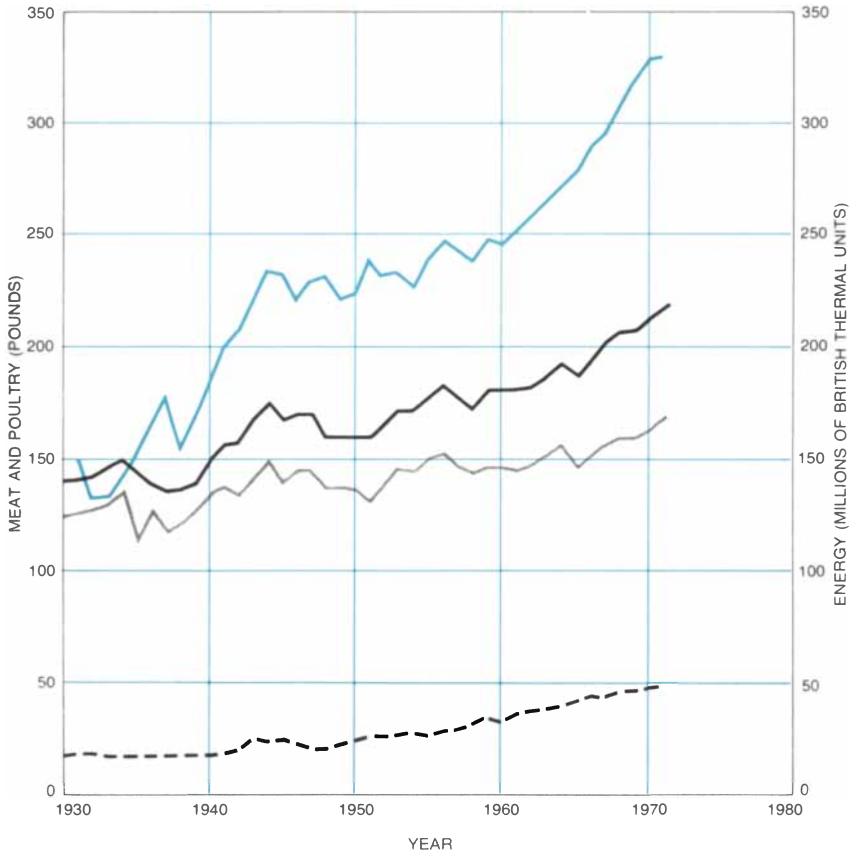
In recent decades, however, the nation's farm exports have become relatively less important. Manufacturing industries have increased their share of the export trade because of their comparative advantage over other countries in technology and management. Their real export has been embodied technology.

Lately the industrial advantage has declined somewhat because much technology has been exported and much has been reproduced abroad. This was probably inevitable; a comparative advantage based on technology tends to be short-lived in an era of rapid technological progress. Comparative advantage based on natural resources is more dependable, which is why it is logical that the U.S. should again rely on its agricultural base to strengthen its foreign-trade position. The logic is even stronger now that worldwide demand for U.S. farm products has risen, partly because of rising standards of living in Europe, Japan and elsewhere and partly because many countries have more difficulty than the U.S. has in increasing agricultural production. The prospect for trade between the U.S. and the U.S.S.R. rests essentially on the superiority of the U.S. in agriculture and the large mineral resources of the U.S.S.R.

A revival of this country's role as food grower for much of the world will be easier to achieve if the U.S. can slow down or halt the trend toward devoting an increasing amount of land to feed domestic animals for an increasing consumption of meat. Recent experience is suggestive: if butter had continued to dominate margarine in the marketplace, less grain would now be available for export. In the future a shift toward meat analogues made with vegetable protein, soy protein being the cheapest and most abundant, could help to release more acreage for export crops.

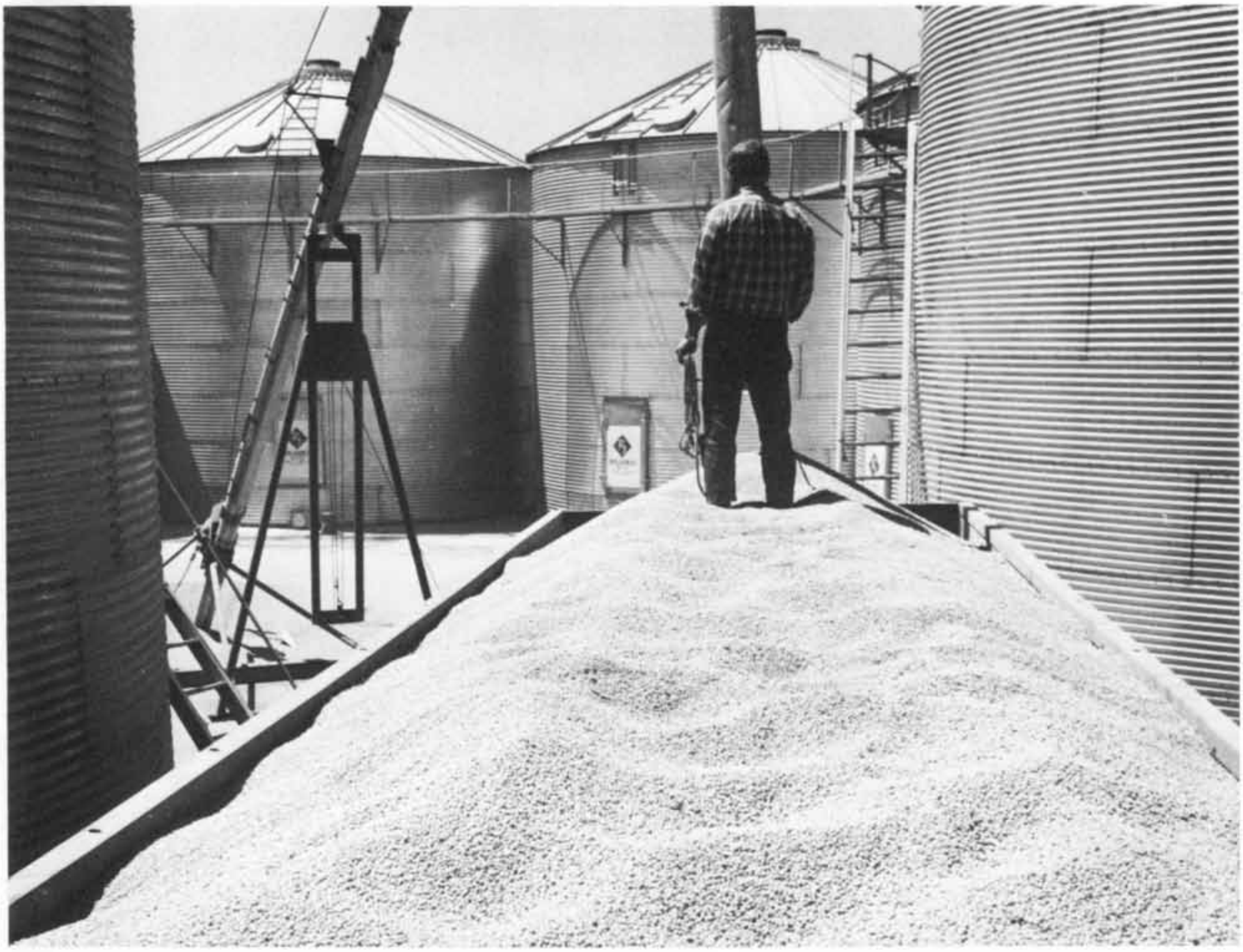
The shifts in production, consumption and trade implicit in the argument set forth here cannot be effected suddenly. What is important is to look ahead at national problems and to be aware of the probable consequences of alternative courses of action. If foresight had been applied to the long-term prospects of the economy in the 1950's, for example, it is likely that neither the current energy crisis nor the adverse balance of foreign trade would have arisen.

The superiority of the soybean both for domestic purposes and for export comes from the wide range of products that can be made from it. An important general characteristic of a modern economy is an increasing versatility of raw materials. This characteristic, together with rising productivity and per capita income, is one of the sources of the widening freedom of choice that makes the socioeconomic environment an "optional society"—a society where the range of options open both to individuals and to society as a whole is wider than ever and still widening. The soybean ranks high in this respect, and so it is an eminently modern commodity. The biologists and food chemists who work with soybeans therefore hold important keys to the nation's economic future.

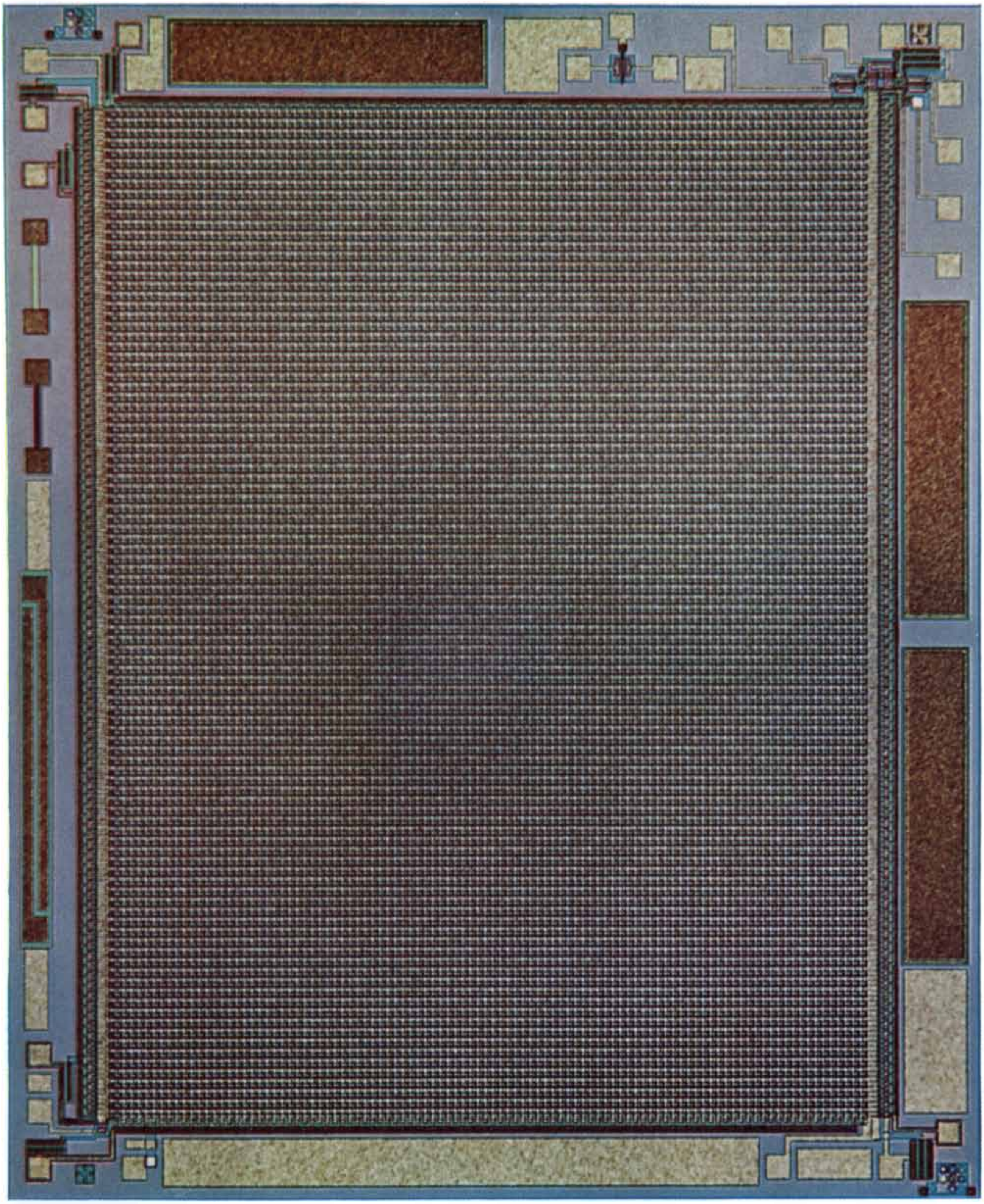


CONSUMPTION OF MEAT AND ENERGY is rising in the U.S., leading simultaneously to an increased dependence on imports of petroleum and meat and to the use of more farmland to raise food for domestic animals. The author argues that more reliance on such crops as soybeans would relieve pressure on the land and also yield more agricultural exports to earn money for fuel imports. Top curve shows energy consumption. The curves below it successively show total meat and poultry consumption, then meat alone and poultry alone.

ACRES OF SOYBEANS grow on a farm in Missouri, which is fourth among the seven states that made the major contribution to the 1973 U.S. soybean crop. Others are Illinois, Iowa, Indiana, Minnesota, Arkansas and Ohio. Crop is important in 23 other states.



STORED SOYBEANS are unloaded from storage bins on a farm in Mississippi. As in this case, part of the crop that is harvested in the fall is frequently stored for sale in subsequent seasons. A large part of the soybeans moving in international trade comes from the U.S.



CHARGE-COUPLED IMAGE SENSOR shown in this color photomicrograph, made in the author's section of the Research and Development Laboratory of the Fairchild Camera and Instrument Corporation, consists of 10,000 photosensor elements arrayed in a rectangular 100-by-100 grid on a silicon chip measuring only .12 by .16 inch. (Large brown and white shapes around border are metal contacts for electrodes.) The device is designed to be used in an experimental television camera based on the concept of charge-coupling, that is, the collective transfer of all the mobile electric charge stored within a semiconductor storage element to a similar, adjacent storage element by the external manipulation of voltages. Although such charge-coupled image sensors are still in a somewhat primitive state, in the author's view they "clearly point the way toward a powerful camera technology." An enlargement of a small portion of a similar charge-coupled photosensor array appears in the scanning electron micrograph on the opposite page.

CHARGE-COUPLED DEVICES

The products of a new concept in semiconductor electronics, they hold considerable promise in applications as diverse as image sensors and information-storage elements for computer memories

by Gilbert F. Amelio

For the past four years there has been a growing excitement among solid-state physicists about a new concept in semiconductor electronics that may someday have an impact on our lives as dramatic as that of the transistor. The new concept is charge-coupling and its practical manifestation is the charge-coupled device.

Like the transistor, the charge-coupled device is a concept of semiconductor electronics; as such it is subject to the same physical laws that govern the transistor's dynamics and fabrication. That, however, is where the similarity ends. Although the charge-coupled device shares much the same technological base with its distinguished predecessor, it is a functional concept that focuses on the manipulation of information rather than an active concept that focuses on the modulation of electric currents. Transistor technology has made possible computer-memory components with thousands of memory elements on a single chip of silicon; charge-coupling is making possible comparably sized memory components with tens of thousands or even hundreds of thousands of memory cells per silicon chip at approximately the same cost.

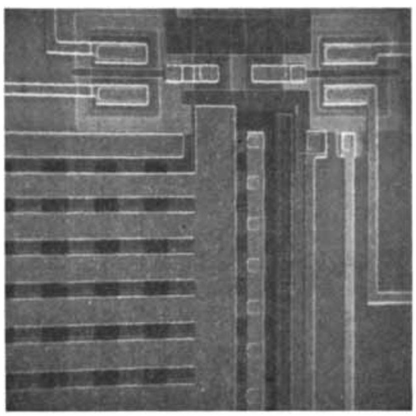
What is charge-coupling? It is the collective transfer of all the mobile electric charge stored within a semiconductor storage element to a similar, adjacent storage element by the external manipulation of voltages. The quantity of the stored charge in this mobile "packet" can vary widely, depending on the applied voltages and on the capacitance of the storage element. The amount of electric charge in each packet can represent information.

Perhaps the easiest way to visualize the operation of a charge-coupled device is through the use of a mechanical analogy. Imagine a machine consisting of a series of three reciprocating pistons with a crankshaft and connecting rods to drive them [*see top illustration on next two pages*]. On top of one or more of the pistons is a fluid. Note that rotating the crankshaft in a clockwise manner causes the fluid to move to the right, whereas rotating the crankshaft in a counterclockwise manner would cause the fluid to move to the left. Since it takes three pistons to repeat the pattern, this arrangement is called a three-phase system. If it is desired to move the fluid in one direction only, a two-phase system can be devised by imposing an asymmetry on the piston design [*see bottom illustration on next two pages*]. Regardless of the direction of rotation, the fluid now advances to the right.

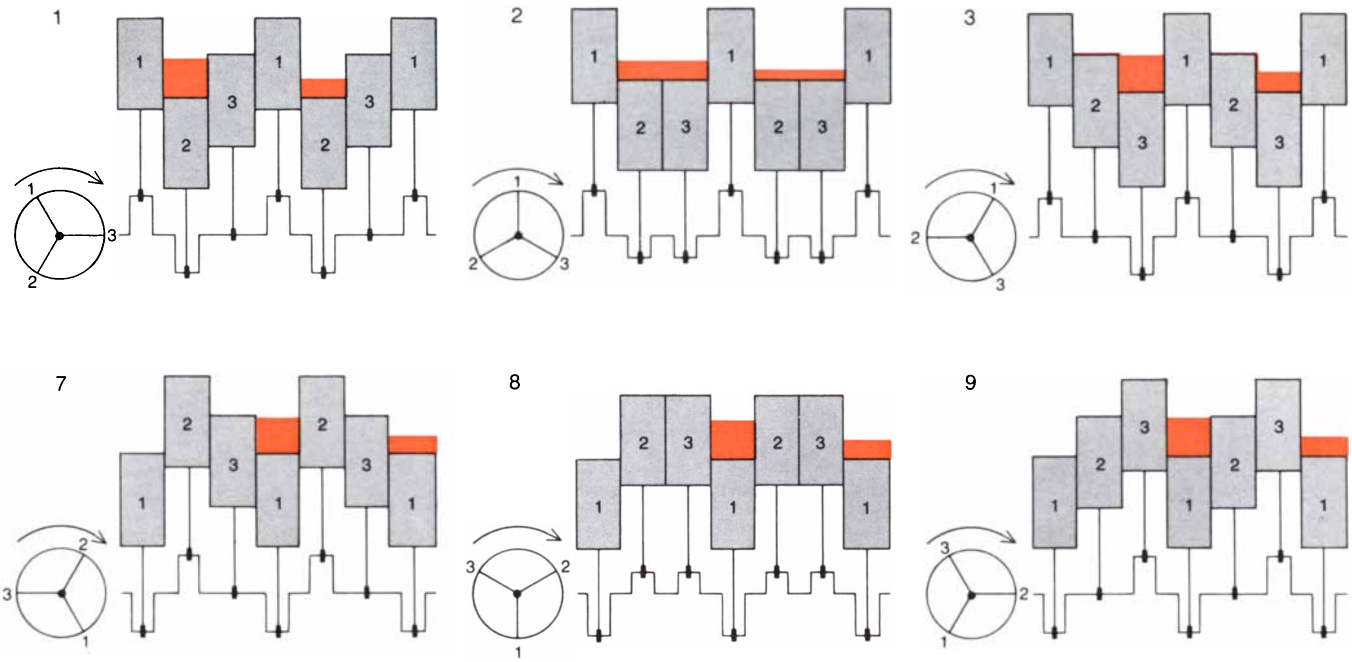
Analogous charge-coupled devices can be fabricated of silicon [*see illustrations on page 26*]. The devices consist of a "*p* type" silicon substrate (in which electrons are normally the signal carriers) with a silicon dioxide insulating layer on its surface. An array of conducting electrodes is deposited in turn on the surface of the insulator. The electrodes can be interconnected to establish either two-phase or three-phase operation. Underlying the insulator and within the bulk of the semiconductor the electrical conductivity of the silicon can be selectively altered to form "*n* type" material (in which not electrons but electron "holes" are normally the signal carriers). The correspondence with the machine in the mechanical analogy is realized by supposing that the fluid represents an accumulation of electrons, that the pistons represent the potential energy associated with the voltages applied to the electrodes and that the crankshaft and connecting rods represent the driving voltages and their relative timing.

When a periodic wave form called a "clock" voltage is applied to the electrodes, some of the electrons in the vicinity of each electrode will form a discrete packet of charge and move one charge-coupled element, or unit cell, to the right for each full clock cycle. The packets of electron charge therefore move to the right as a result of the continuous lateral displacement of the local "potential well" in which they find themselves. They are thus—or so it seems—always falling.

The creation of the necessary potential well in the semiconductor substrate deserves some elaboration because of its central importance to the charge-coupling concept. In this context a potential well is a localized volume in the silicon that is attractive to electrons; in other words, it is the most positive place around and hence is a desirable location from the point of view of the negative electron. Potential wells are formed in a charge-coupled storage element by the interaction of the different conductivity-



CLOSEUP VIEW of a small portion near the output of a charge-coupled photosensor array is provided by this scanning electron micrograph. Each element and its associated readout electrode measure 1.9 square mils.



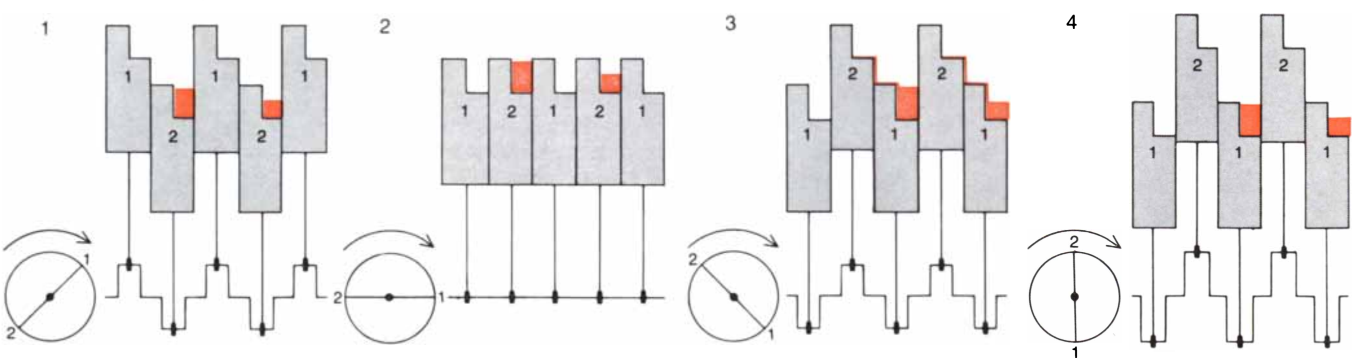
MECHANICAL ANALOGY useful in visualizing the operation of a charge-coupled device is depicted in this sequence of idealized drawings. The machine illustrated consists of a repeating series of three reciprocating pistons with a crankshaft and connecting rods to drive them. On top of one or more of the pistons is a fluid (*color*). Rotating the crankshaft in a clockwise manner, as shown

type regions of the silicon [*see illustration on page 27*]. This interaction forms a well for electrons such that the higher the clock voltage, the deeper the well. Any electrons in the well will move with the clock voltages.

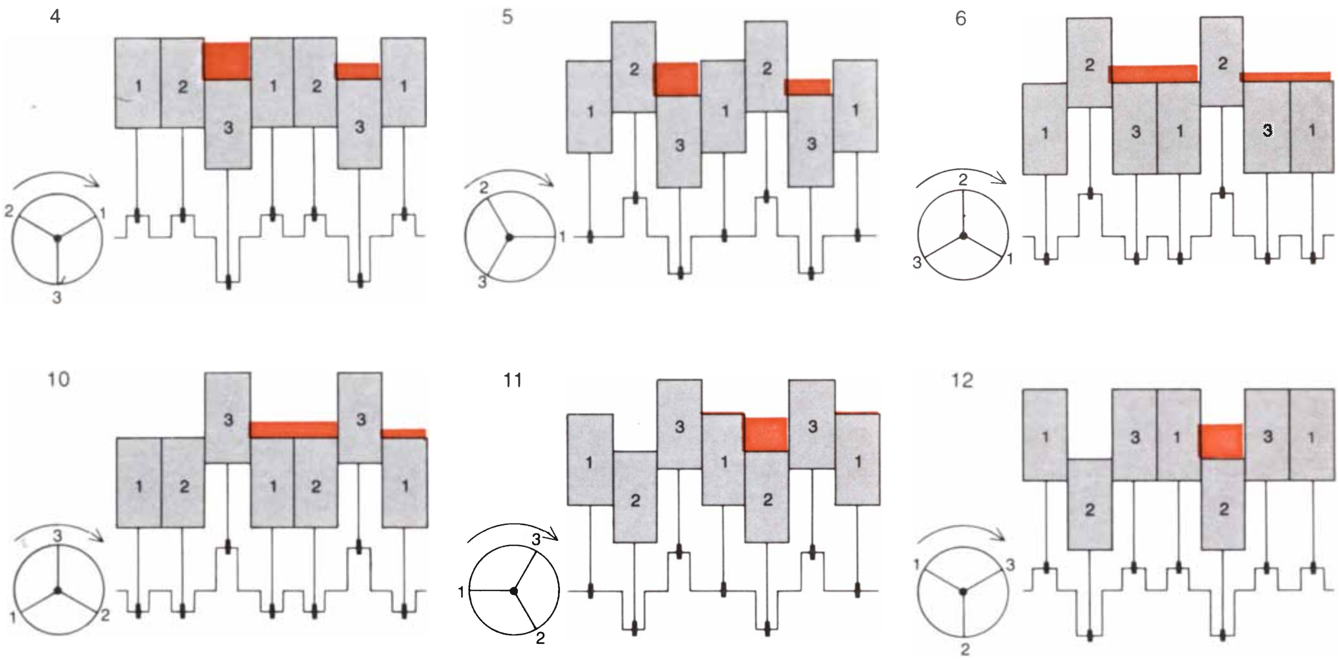
Now, if two or more wells of different depths are placed close to one another, the wells will overlap and charge may be "coupled," or transferred, from one storage element to the next as the depth of the well is altered by the clock voltages. Thus the external clock voltages on the electrodes cause the electrons to move in packets through the semiconductor in a potential-energy trough known as a channel. This mode of electron transfer is the essence of charge-coupling.

The phenomenon of charge-coupling is in itself inadequate for the purpose of constructing a useful device. A practical charge-coupled device must be able to introduce the necessary electrons into the structure and also have a means at some location in the channel for detecting the amount of charge in a packet. Thus for a structure to be classified as a charge-coupled device it must possess at least three attributes: input, charge-coupling and output.

As an example of a simple yet functionally complete charge-coupled device, consider a "shift register" consisting of eight three-phase elements, an input diode and gate and an output diode and gate [*see illustration on page 28*]. This structure is in fact very similar to the first charge-coupled device ever fabricated. The signal that is to be entered into the charge-coupled device is connected to the input diode, which acts as a source of electrons. If the input gate is held at a low voltage, no signal electrons can enter the channel. In order to put a packet of electrons into the device it is necessary to wait until the first-phase electrodes are in the high-voltage condition and then "turn on" the input gate by raising its voltage. Electrons fill the potential well until the energy level for electrons in the well is the same as that for the electrons in the source. The input-gate voltage is now lowered to isolate the source, and the charge packet created is ready for transfer down the channel. In the detection of the signal



ASYMMETRICAL PISTONS are added to the mechanical analogue in order to introduce the operating principle of a two-phase system. Regardless of the direction in which the crankshaft is rotated, the fluid now advances to the right. In the correspondence



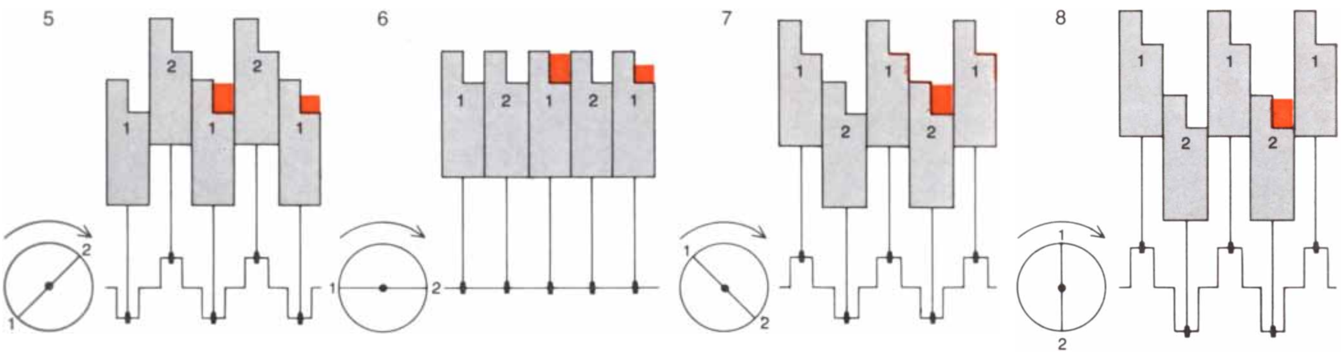
in this instance, causes the fluid to move to the right. If the crankshaft were to be rotated in a counterclockwise manner, on the other hand, the fluid would move to the left. This particular type of arrangement, which requires three pistons to repeat the pattern, is called a three-phase system. An analogous charge-coupled device can be fabricated of silicon (*see top illustration on next page*).

the charge is merely transferred to a "drain," or output diode, where it appears as a current in some external circuit. This simple charge-coupled device fulfills the function of an eight-bit shift register, a device potentially useful in computer architecture.

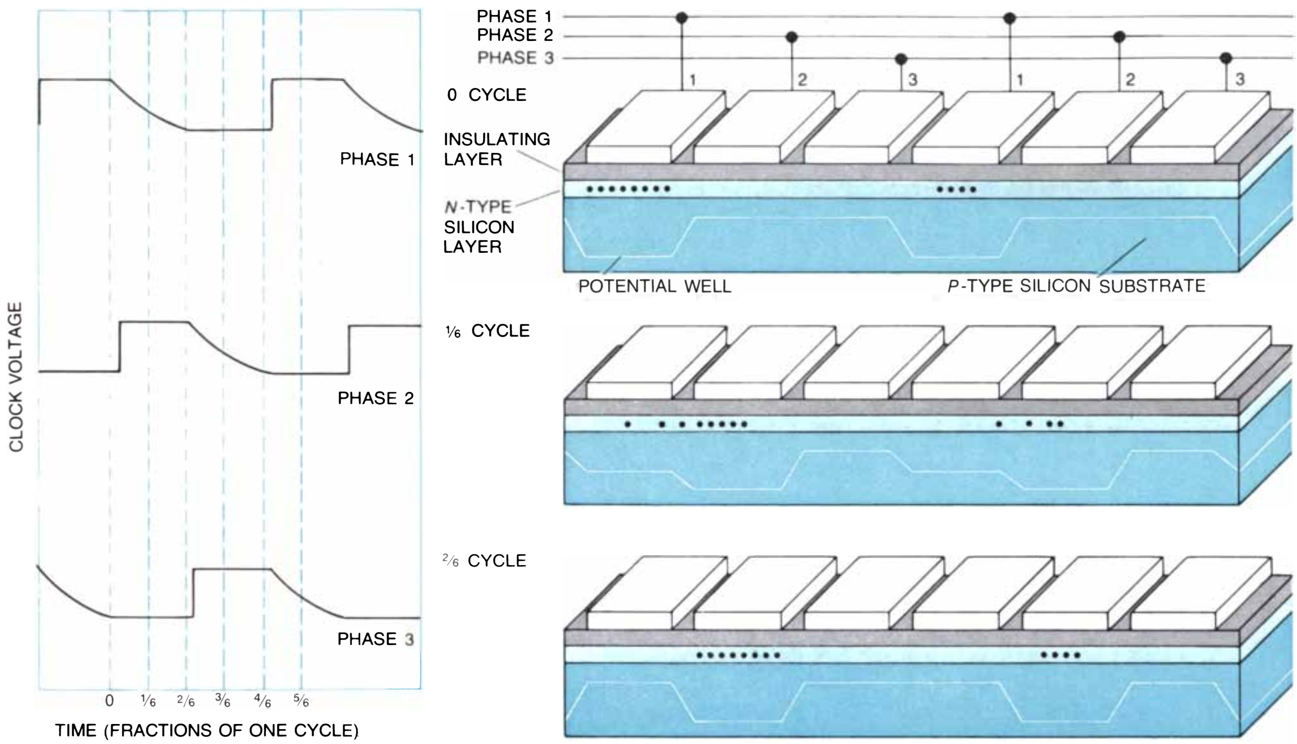
Devices fabricated and operated in this manner verify the predicted performance with one exception. Unfortunately not all the electrons advance with the packet on each transfer, and the residual charge appears in a trailing packet. The magnitude of such "charge-transfer inefficiency" is a function of the design of the device and the frequency of operation. Transfer inefficiency imposes a fundamental limitation on the speed and number of transfers for a practical charge-coupled device because of the resulting attenuation of the charge packet as it is moved through the device from one region to the next.

There are two reasons for charge-transfer inefficiency. First, the electrons may be inhibited from moving because of local regions of lower potential energy (corresponding to dents or ridges in the top of the piston in the mechanical analogy). Second, the frequency of operation may be so high that there is not enough time for all the electrons to follow the moving potential wells. The former problem is one that is influenced predominantly by the design details of the particular charge-coupled device. Researchers working on the development of such devices are continuing to explore this aspect of charge-coupling. Recent advances in technology have significantly reduced the seriousness of the problem. The problem of the speed of the electrons' motion, however, has more basic origins and deserves additional comment.

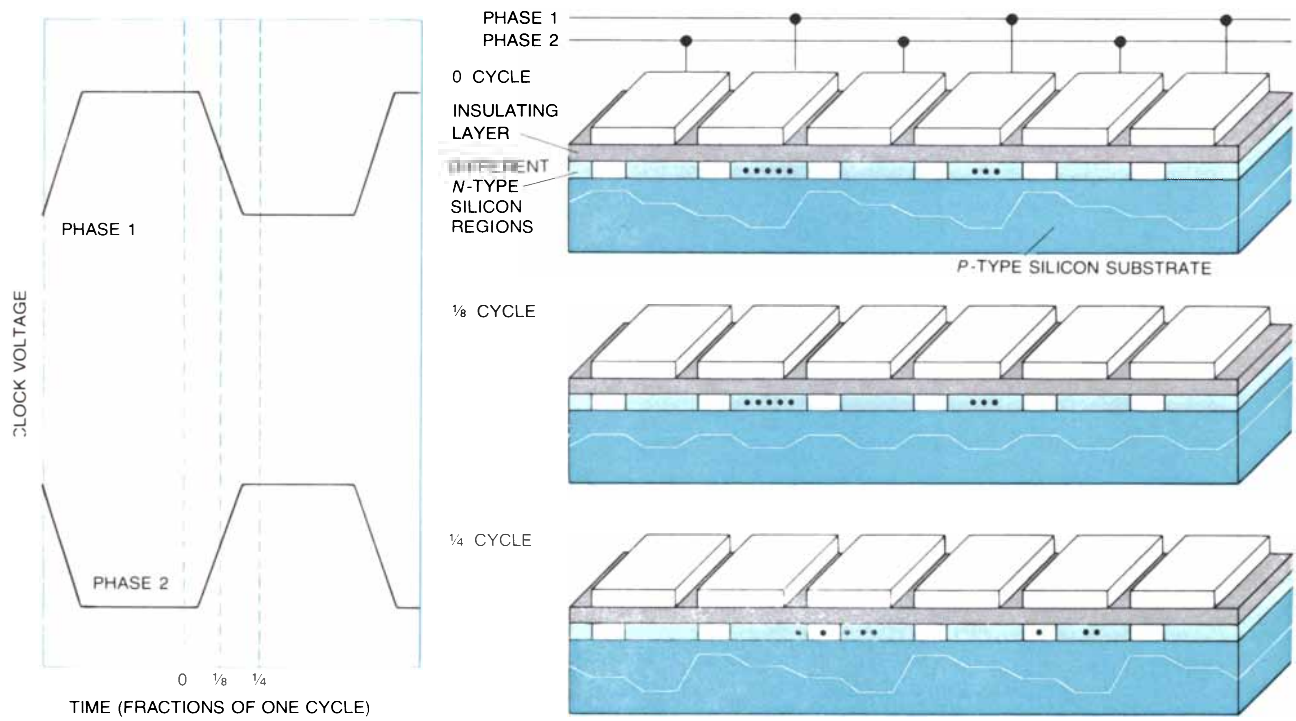
The electrons are induced to move to an adjacent region of lower energy (that is, a deeper potential well) by a combination of three influences: self-induced forces, field-aided forces and thermal forces. Self-induced movement results from the simple fact that a high-density packet of electrons (or any similar particles) tends to spread rapidly if the constraining force is removed, as is the case when the clock voltages change. This type of force is important during the



with an actual charge-coupled device the fluid represents an accumulation of electrons, the pistons represent the potential energy associated with the applied voltages and the crankshaft and the connecting rods represent the driving voltages and their timing.



TWO THREE-PHASE CHARGE-COUPLED ELEMENTS are shown in the cross-sectional diagram at right; the curves at left give the relative timing of the "clock voltage" wave forms for three-phase operation. The device consists of a "*p* type" silicon substrate (in which electrons are normally the signal carriers) with a silicon dioxide insulating layer on its surface. Conducting electrodes are deposited on the surface of the insulator. Underlying the insulator and within the bulk of the semiconductor the electrical conductivity of the silicon can be altered to form an "*n* type" layer (in which electron "holes" are normally the signal carriers). When the clock voltage is applied to the electrodes, some of the electrons in the vicinity of each electrode will form a discrete packet of charge (*black dots*) and move one element to the right for each full clock cycle. In effect the packets of electron charge move to the right as a result of the continuous lateral displacement of the local "potential well" in which they find themselves (*white contours in substrate*).



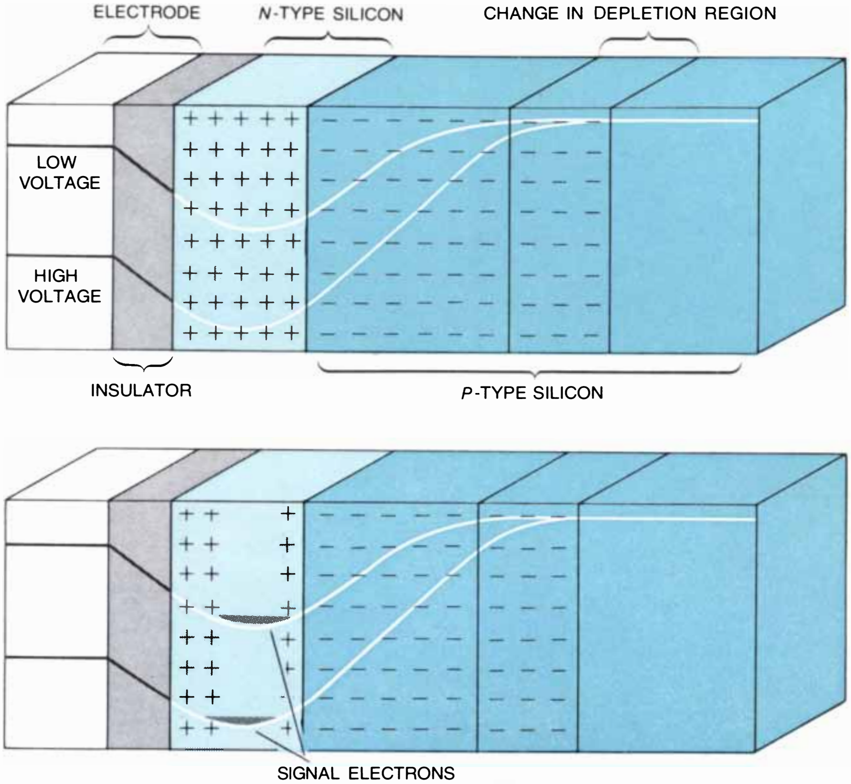
THREE TWO-PHASE CHARGE-COUPLED ELEMENTS are shown in these cross-sectional diagrams; again the curves give the relative timing of the clock voltages, this time for two-phase operation. Here the potential wells are given the required asymmetry by the introduction of different *n*-type conductivity regions just under the insulating layer. As in the illustration at the top, the external clock voltages on the electrodes cause the electrons to move in packets through the *n*-type semiconductor layer toward the right.

early stages of charge transfer. Field-aided movement is important if the structure is designed in such a way that electric fields exist to assist the electrons' motion in the desired direction. This corresponds to adding a slope to the top of the pistons in the mechanical analogy. If such a force is present, it is important only toward the end of the charge-transfer cycle. Thermal forces arise from the fact that the electrons receive thermal energy from the silicon lattice and as a result are free to move about randomly. In their random motion they tend to move to regions of minimum electron energy. This type of force is important at the end of the transfer cycle only if field-aided forces are absent.

The self-induced force lasts for only a brief time at the beginning of the transfer cycle, but it is responsible for moving about 90 percent of a "saturation," or full, charge. If the field-aided force is present, it is responsible for moving most of the remaining charge at a rate directly proportional to the strength of the electric field and inversely proportional to the distance between the electrodes. If the field-aided force is not present, the remaining charge will move under the influence of thermal forces at a rate directly proportional to the temperature and inversely proportional to the square of the distance between the electrodes. This rate is usually lower than that resulting from the field-aided force, although at small dimensions it becomes increasingly significant because of its inverse quadratic dependence on distance.

Although these forces are responsible for moving only a comparatively small fraction of the total charge packet, they are important because very little transfer inefficiency can be tolerated in practical devices. For example, if 1 percent of the charge is left behind at each transfer, most of a charge packet will have dispersed after only 100 transfers. In general the charge-transfer inefficiency must approach one part in 10,000 to be considered acceptable for most practical applications. In spite of this requirement, devices that can be operated at frequencies of up to 100 megahertz (100 million cycles per second) are possible if the structures are made small enough. With modern microelectronic manufacturing techniques it is possible to design and build a charge-coupled unit cell with dimensions of less than a mil (a thousandth of an inch) on a side, although it is not always appropriate to do so.

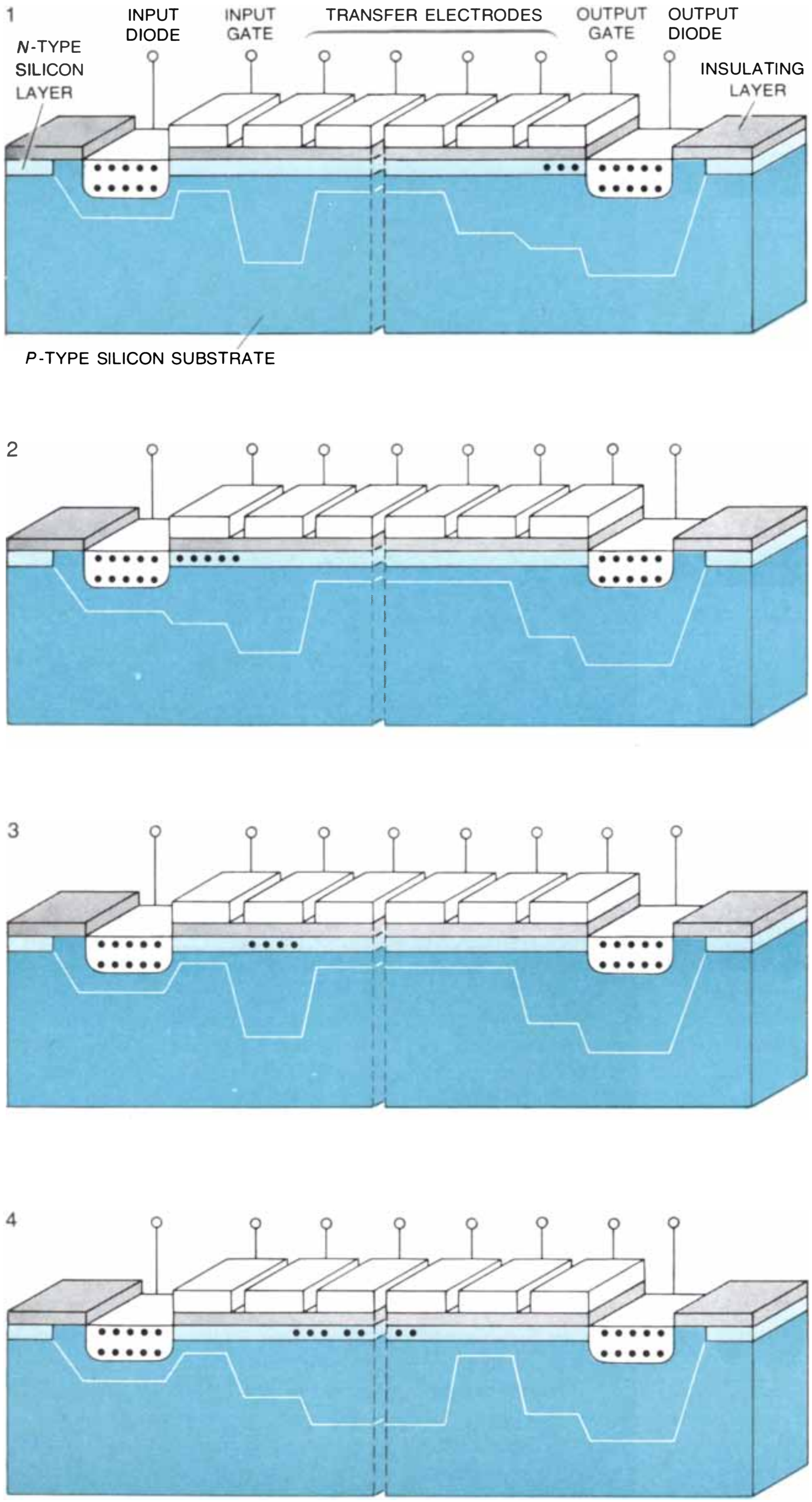
Unit cells of such small dimensions are possible because of the simple nature of the charge-coupled structure, which does not require direct contact with the silicon in the array region. This arrangement is to be contrasted with conventional transistor technology, which in general requires several contacts per functional cell. Contacts consume a significant amount of valuable silicon because of the contact area and the tolerances needed to form a good electrical connection. From the manufacturing viewpoint it is this feature more than any other that makes charge-coupled devices so attractive.



POTENTIAL-ENERGY PROFILES for a typical charge-coupled information-storage element are shown here as a function of distance into the bulk of the semiconductor at right angles to its surface. (In order to show the potential wells clearly, this diagram has been rotated by 90 degrees with respect to the preceding ones.) The charge-distribution patterns are shown for two situations: with no electrons in the well (*top*) and with some electrons in the well (*bottom*). As the curves indicate, the higher the clock voltage, the deeper the well.

The ability to generate, move about and detect many separate packets of electrons in a small piece of semiconductor material suggests that the charge-coupling principle can be applied to fulfill a number of information-processing requirements. In particular the highly ordered manipulation of charge packets characteristic of the operation of charge-coupled devices favors uses such as image sensing, computer-memory operation and sampled-signal processing. In each case the function is achieved by a proper combination of charge-coupled unit cells that operate individually exactly as described above.

Silicon, the semiconductor material of which charge-coupled devices are generally fabricated, is highly sensitive to visible and near-infrared radiation [*see illustration on page 29*]. In other words, when light falls on a silicon substrate, the radiation is absorbed (by means of the Einstein photoelectric effect), which results in the generation of electrons in a quantity proportional to the amount of incident light. If there is present an array of potential wells such as the one formed by charge-coupled devices, these electrons will fill the wells to a level corresponding to the amount of light in their vicinity. This "electro-optic" creation of electrons represents an input to the charge-coupled device that is entirely different from the input method required for the shift register discussed above and makes the charge-coupling concept useful for very different kinds of application. Nonetheless, the packets of electrons generated by the light can be moved, just as in the shift register, to a point of detection and converted to an electrical signal representative of the optical image incident on the device. That signal, after some conditioning, can be displayed on a cathode ray tube. In this way a charge-coupled device can



INPUT AND OUTPUT OPERATIONS of a simple eight-element, three-phase charge-coupled "shift register" are summarized in this series of diagrams. The signal enters the device by way of an input diode, which acts as a source of electrons. If the input gate is held at a low voltage, no signal electrons can enter the potential-energy "channel" (*1*). In order to put electrons into the device one must wait until the first-phase electrodes are in the high-voltage condition and then "turn on" the input gate by raising its voltage (*2*). Electrons fill the potential well until the energy level for the electrons in the well is the same as that for the electrons in the source. The input-gate voltage is now lowered to isolate the source (*3*), and the charge packet created is ready for transfer down the channel (*4*). The signal is detected by transferring the charge packet to an output diode, where it appears as a current.

become the heart of a television camera.

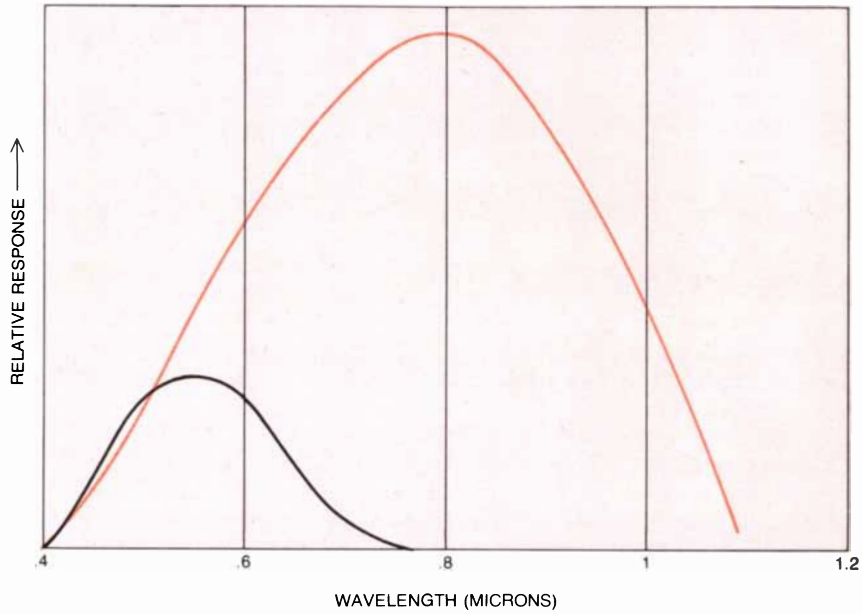
One of the significant advantages of charge-coupled image sensors over vacuum-tube sensors is the precise knowledge of the photosensor locations with respect to one another. In a camera tube the video image is "read" from a photosensitive material by a scanning electron beam. The position of the beam is never precisely known because of the uncertainty in the sweep circuits resulting from random electrical noise. In a charge-coupled sensor the location of the individual photosensor sites is known exactly, since it is determined during the manufacture of the component. Such "metric" accuracy is important for proper alignment in color cameras and in applications requiring data reduction of the acquired image (as in photographic missions in space and photogrammetry).

It is generally convenient for purposes of discussion to separate charge-coupled sensors into two categories: linear sensors and area sensors. A linear image sensor is a simple straight-line array of photosensors with the associated readout and sensing circuitry. An area image sensor is a two-dimensional mosaic of photosensors, again with the associated readout and sensing circuitry.

Linear image sensors are used for a host of applications, including air-to-ground and space-to-ground imaging, facsimile recording and slow-scan television. The image to be viewed is obtained by providing relative motion between the sensor and the scene with the axis of the array perpendicular to the direction of the motion. A resolution of 500 or more photosensor elements is usually required. A primitive linear imaging device can consist of nothing more than a charge-coupled shift register and an output diode. In this structure the image is acquired when one holds the potential wells stationary by stopping the voltage clocks for some period of time (the "integration time") and then rapidly reads out the information by starting the clocks. Such a simple charge-coupled device should be practical only in special applications that allow very long integration times. The reason for this limitation is the "smearing" of the image that results when the shift register is clocked at the same time that it is illuminated.

A really practical charge-coupled linear image sensor is more complex. It consists of a photosensor array for accumulating the photocharge pattern plus an associated charge-coupled shift register with one charge-coupled element for each photosensor element in order to move the resulting charge packets to an output point. The elements of the photosensor array are individual charge-coupled storage elements with a common electrode called a photogate. They are electrically separated from one another by a highly concentrated *p*-type region called a channel stop. The photosensor array is separated from the charge-coupled shift register by a region over which there is an electrode called the transfer gate.

In operation the photogate voltage is held high and the charge generated by the incident radiation (the photocharge) is collected by the individual photosensor elements. At the end of the integration time the transfer-gate voltage is raised from its normally low voltage condition. The charge-coupled shift-register electrodes adjacent to the photosensor elements are also brought to a high-voltage state. The photogate voltage is then lowered and the accumulated photocharge transfers to the shift register. After that is accomplished the transfer-gate voltage is lowered and the photogate voltage is brought back to its normally high state for another integration period. Meanwhile the charge-coupled shift register is clocked for the purpose of reading out the charge pattern.



RELATIVE SPECTRAL RESPONSES of a charge-coupled silicon photosensor element (*colored curve*) and the human eye (*black curve*) are compared. The semiconductor material absorbs not only visible light (.4 to .7 micron) but also near-infrared radiation (.7 micron to 1.1 microns). The absorption of such radiation by a silicon substrate results in the generation of electrons in a quantity proportional to the amount of incident radiation. It is this "electro-optic" property that enables charge-coupled devices to be used as image sensors.

A high-density image sensor is more economically designed with one shift register on each side of the photosensor array. Since there must be one charge-coupled element for each photosensor element, the distance between photosensor elements is equal to the distance between the shift-register electrodes for a two-phase charge-coupled shift register and is equal to 1.5 times the distance between shift-register electrodes for a three-phase charge-coupled shift register. In this example the signal charge from the two three-phase shift registers is transported to a three-phase, two-element register for delivery to the on-chip preamplifier. If two-phase technology is used, however, it is possible to shift the charge directly into an output diode, which is in turn the input to the on-chip preamplifier. Note that in either case the information-output rate of the device is twice the rate of either of the long shift registers. It is clear from this example that a two-phase charge-coupled structure not only is easier to clock but also is more economical to lay out for a practical device. Even though it is somewhat more difficult to manufacture because of the required asymmetry, it is likely to dominate future designs of charge-coupled devices when fully developed.

A linear image sensor can be made to produce conventional two-dimensional images [*see illustration on next page*]. The image to be sensed is placed on a rotating drum, which provides the necessary motion of the image with respect to the device. The speed of rotation is synchronized with the vertical scan of the monitor. The charge-coupled linear image sensor provides each horizontal video line for the monitor by a complete sensing and readout operation repeated rapidly to supply all the horizontal lines for a full frame. In many applications the device is the moving element in the system, as in aerial reconnaissance, where the device is located in an airplane or a satellite.

The quality of image reproduction achievable with a linear charge-coupled sensor is excellent, reflecting the large dynamic range of the image sensor [*see illustration on page 31*]. The dynamic range is the ratio of the maximum to the minimum detectable image intensity. The quality of the reproduction demonstrates the very high transfer efficiencies and low electrical noise levels that can be achieved in existing charge-coupled devices.

Area image sensors are useful primarily for television-type camera applications. The image is obtained by conventional line-by-line scanning of the array mosaic and reproduction of the resulting video signal on a standard raster-scanned cathode-ray-tube monitor. A charge-coupled area image sensor designed for such a readout mode can be designed in a manner analogous to the linear image sensor. As in standard broadcast television, the image is read out in two separate fields by first reading all the even-numbered photosensor elements in each column and then all the odd-numbered photosensor elements in each column rather than by reading the odd and even elements in parallel, as in the case of the linear image sensor.

The area image sensor operates as follows. Light falling continuously on the photosensor sites produces electrons, which accumulate as charge packets in the potential wells created by the photogate voltage. After an interval of a thirtieth of a second the charge packets collected in the photosensors adjacent to all the phase-1 electrodes are transferred to the region under the phase-1 electrodes by raising the phase-1 voltage and lowering the photogate voltage. The charge packets in photosensor sites adjacent to the phase-2 electrodes do not transfer because the phase-2 voltage remains low. After the phase-1 transfer takes place the photogate voltage again goes to its normally high state and more electrons begin to accumulate in the depleted photosensor sites. The charge packets in the opaqued shift register are now transferred to the horizontal shift register at the top of the array. Each vertical transfer fills the horizontal register, which is then read out completely, producing a line of video information at the output. After all these lines are read out (a procedure that takes only a sixtieth of a second) the photosensors adjacent to all the phase-2 electrodes are read out, and in a similar manner this second field is delivered as a video signal at the output. Finally, the entire operation begins again and is completed at regular intervals of a thirtieth of a second.

A typical image sensor designed to operate in this fashion consists of a rectangular 100-by-100 photosensor grid [*see illustration on page 22*]. Each photosensor element and associated readout electrode occupies only 1.9 square mils. All 10,000 elements fit on a chip that measures .12 by .16 inch. An image taken with a camera system using such a device can be displayed on a television monitor.

This image-sensing device and others made by charge-coupled techniques are still somewhat primitive, but they clearly point the way toward a powerful camera technology. The combination of solid-state reliability, low-voltage operation, low power dissipation, large dynamic range, metric reproducibility and visible and near-infrared response offers to the potential user a compelling advantage over vacuum-tube image sensors and other solid-state image sensors.

The charge-coupling concept is basically one of semiconductor electronics rather than one of electro-optics. Because of the electro-optic characteristics of silicon, however, the light-sensing properties of charge-coupled arrays have tended to dominate this new technology. Nonetheless, the data-handling properties of such arrays may be of equal or even greater significance.

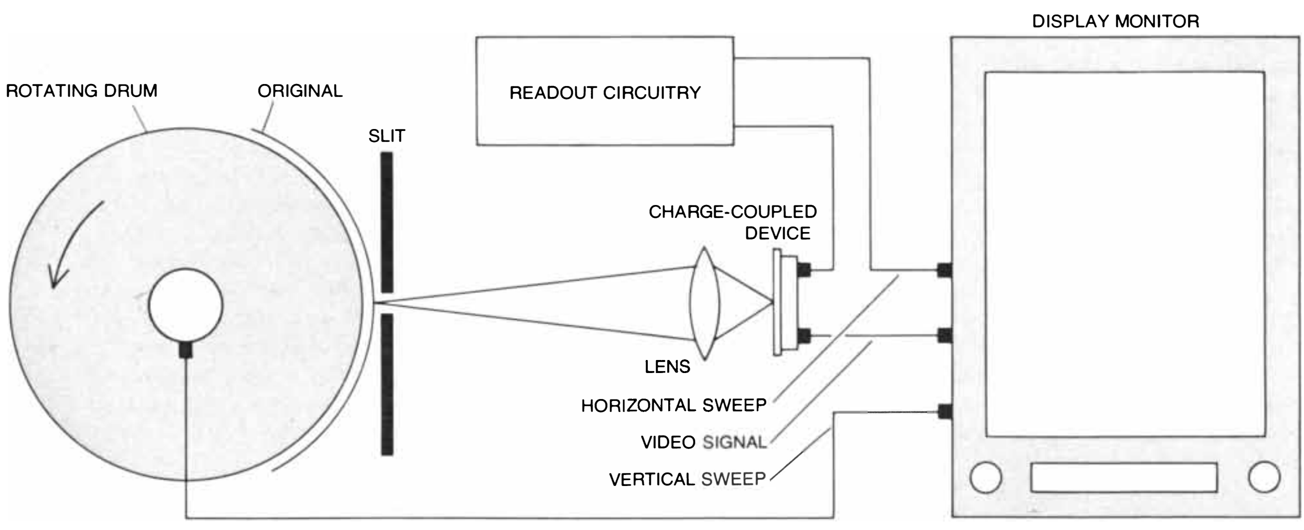
A charge-coupled semiconductor array is virtually ideal as a time-sampled analogue shift register. From the viewpoint of the electrical engineer this means a delay line where the delay is proportional to the readin/readout rate; if the array is long enough to contain the complete message, the readin and readout rates can be different and the maximum delay available is limited only by the thermal generation of random electrons. At low temperatures several minutes of delay are possible.

As a memory or digital-storage device, charge-coupled arrays can perform the functions of sequential access or hybrid tasks such as drum or disk storage. The use of solid-state charge-coupled arrays to eliminate all mechanical motion and parts is a strong advantage of a memory consisting of charge-coupled devices.

The intrinsic analogue nature of the charge packet in a charge-coupled device suggests broad potential for application to sampled-signal processing. In a fundamental sense the use of charge-coupled devices as image sensors is merely a special application of the device as an analogue shift register. If one restricts the definition of sampled-signal devices to those with an electrical (rather than an optical) input, then the predominant members of this class are variable delay lines and filters.

A delay line is a circuit that reproduces as accurately as possible an input signal delayed by a finite period of time. A delay line is "variable" if the time delay can be altered electrically. The charge-coupled device acts as a natural delay line since any signal placed on its input diode will appear at its output in sampled form after the interval required for the charge packets to be shifted through all the elements of the structure. The charge-coupled device can be used as a delay line in several ways. First, in the simple continuous mode the delay is equal to the number of unit cells divided by the frequency at which the device is clocked. Alternatively, whenever data appear in bursts, the charge-coupled shift register can be loaded with these data during the burst and the data retained for the desired interval and then read out. In this way the charge-coupled device is said to perform a "buffer" function.

A charge-coupled delay line offers major advantages over the more conventional glass delay line and even significant advantages over the more exotic acoustic-surface-wave devices [see "Acoustic Surface Waves," by Gordon S. Kino and John Shaw; SCIENTIFIC AMERICAN, October, 1972]. Among these are wide dynamic range (better than 60 decibels after 30 milliseconds at room temperature) and separate electronic control of propagation velocity and delay time. Delay lines with such flexibility will be of great value in communications and television applications and will simplify existing methods of producing controlled signal delays. One special application of significant interest is a "scan-rate converter" often required in video communications. Here the charge-coupled device operates in the buffer mode described above to accept and then read



TWO-DIMENSIONAL IMAGES can be reproduced with the aid of a linear charge-coupled image sensor in a variety of ways, one of which is outlined in this schematic diagram. The image to be sensed is placed on a rotating drum (*left*) whose speed of rotation is synchronized with the vertical scan of a conventional television display monitor (*right*). The charge-coupled device and the associated readout circuitry produce horizontal video lines at a rate rapid enough to build up a full-frame image on the screen of the monitor.

EXCELLENT REPRODUCTION obtained with a 500-element linear charge-coupled image sensor under widely varying light conditions is evident in these photographs. An apparatus similar to the one in the illustration on the opposite page was employed to scan the image. The photograph at left shows the original image to be scanned. The photograph at center shows the video display obtained from the charge-coupled system under optimum lighting conditions (30 foot-candles of illumination). The photograph at right shows the video display obtained from the same system but with the light level reduced 1,000 times; to produce this picture the charge-coupled device had to move packets of approximately 400 electrons each through a centimeter of silicon without dispersion.

out video frames at different rates so as to match practical transmission-system bandwidths with standard television-display requirements.

Extension of the simple delay-line concept leads to other sampled-signal processing devices. If a delay line is fabricated with interim taps at which the signal can be sensed and fed back to earlier stages in such a way as to affect the transmission of the data, then this structure can be used as a filter. Such a structure can be conveniently configured as a band-pass filter where the resonant frequency of the circuit is a direct function of the clock frequency. An improvement in the signal-to-noise ratio to within a decibel of the theoretical maximum has already been achieved.

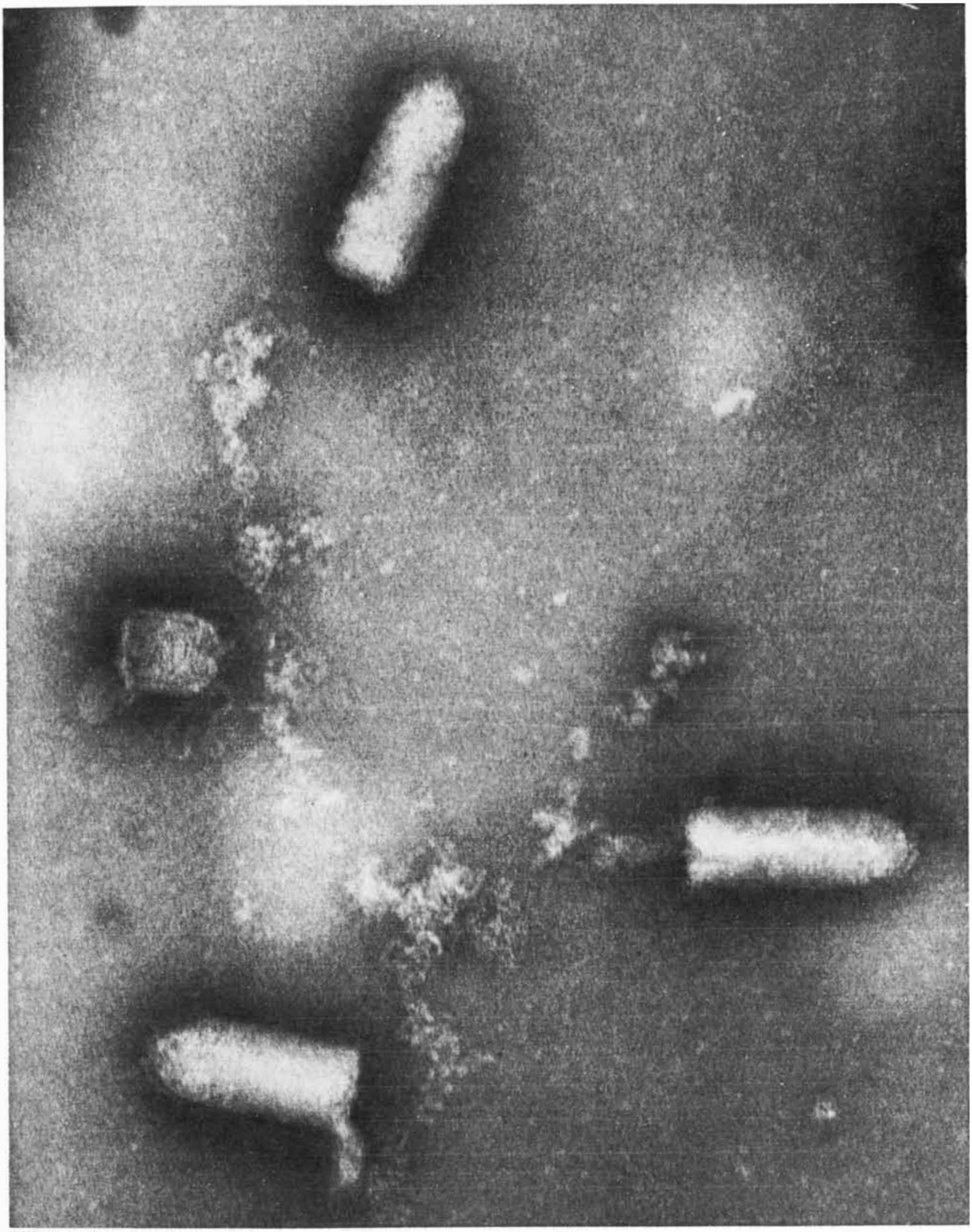
Matched filters find application in wide-spectrum communications and in radar to detect weak signals in high noise backgrounds. In such applications charge-coupled devices will complement acoustic-surface-wave devices, which generally are useful only for delays of less than 100 microseconds.

As mentioned above, a charge-coupled storage element is capable of storing a packet of electrons with a varying amount of charge, depending on the design and operating conditions of the charge-coupled unit cell. Nonetheless, there is no reason one cannot conceptually quantize the charge-handling ability of the cell and view the device as a binary digital element. For example, one can arbitrarily say that if a storage element contains a charge less than half the saturation charge, it contains a "zero," whereas if it possesses a charge greater than half the saturation charge, it contains a "one." In this way the storage element becomes a memory "bit" and a charge-coupled delay line can be made to serve the function of a digital shift register or serially accessible memory. Since this function can be performed by other technologies also, one must ask what charge-coupling has to offer. The answer is cost-effectiveness. A charge-coupled memory not only has all the advantages of a conventional semiconductor component (compatibility with other electronic circuit elements, no mechanical motion, low power and voltage, variable clocking rates and other similar features) but also offers a potentially low cost-per-bit ratio approaching that of a magnetic memory. This is a result of the inherent structural simplicity of the charge-coupled device. By virtue of this simplicity, memory arrays as large as a quarter of a million bits per component on a piece of silicon less than half an inch on a side can be envisioned.

In addition, the power necessary to sustain a charge-coupled memory device is very low since the storage element is not active. The power required to move the charge stored on one charge-coupled element to an adjacent element in a microsecond is approximately a microwatt. Moreover, in a properly organized memory it is not necessary to have all bits moving simultaneously. Thus a one-megahertz, one-megabit charge-coupled memory device would require a power of somewhere between a milliwatt and a watt to sustain it, excluding logic and other functions. The volume required for such a memory is less than that of a pack of cigarettes.

Another advantage lies in the fact that the charge-coupled device is basically analogue in nature. It is thus possible to store more than one data bit in each memory cell. This can be done by storing any one of a number of discrete levels of charge in each cell, thereby greatly increasing the information-packing density. For example, a 100,000-cell device capable of handling eight levels of charge is comparable to a 300,000-bit conventional memory. Such a memory chip would be of great value in digital-to-analogue and analogue-to-digital converters and other applications where multiple levels are achieved only by the addition of vast amounts of memory.

In view of these important prospective features of charge-coupled memory devices it appears that we are at the dawn of a revolution that will ultimately bring today's powerful digital computers directly into our everyday way of life. The charge-coupling concept, in short, is a major new innovation in semiconductor electronics. By virtue of its simplicity in design and fabrication, its high performance in terms of dynamic range and low power, and its high packing density and potentially low cost, the technology of charge-coupling will create major and unique new applications for semiconductors that will have a direct impact on our lives.



BULLET-SHAPED PARTICLES of vesicular stomatitis virus (VSV), which causes mild disease in man but severe disease in some animals, are shown enlarged about 250,000 diameters in this electron micrograph. The roundish particle to the left of center of the photograph is a defective-interfering virus. It has only a portion of its ribonucleic acid (RNA) genome, or set of genes, and is dependent on the complete virus to support its replication. The coiled strands between the virus particles are nucleoproteins that were released from the virus particles during their preparation for electron microscopy. These nucleoprotein coils contain the ribonucleic acid genome of the virus. The photomicrograph was taken by Ronald Leavitt of the University of California at San Diego.

SLOW, INAPPARENT AND RECURRENT VIRUSES

Certain viruses can persist in their host without giving rise to the usual signs of infection. Evidence is accumulating that such viruses can cause or trigger chronic degenerative disease in man

by John J. Holland

In recent years it has become apparent that some slow, persistent diseases that do not superficially appear to be infectious diseases can be caused or triggered by unusual "slow" viruses. Strong evidence has accumulated that four severe neurological diseases are caused by slow viruses: kuru and Creutzfeldt-Jakob disease in man, scrapie in sheep and transmissible encephalopathy in mink. There is also preliminary evidence, as yet inconclusive, that a number of common human degenerative diseases, for example multiple sclerosis, rheumatoid arthritis, leukemia and diabetes, may actually be the result of slow and inapparent virus infection. If the preliminary findings of virus involvement in degenerative diseases are substantiated, then virtually all the chronic and slowly progressive diseases in man for which no cause is known will have to be screened to determine if any are of viral origin.

In slow-virus disease the slowness is not necessarily a feature of the virus itself but rather of the course of the disease. In the typical acute virus infection the infected cell generally dies between 10 and 48 hours after it is infected. In slow-virus infection the virus can persist in cells and replicate without causing the death of the host cell. The presence of the virus does not go unnoticed by the body's immune system; indeed, it is the cumulative reaction of the immune system to a persistent virus infection that may be responsible for much of the damage found in degenerative diseases.

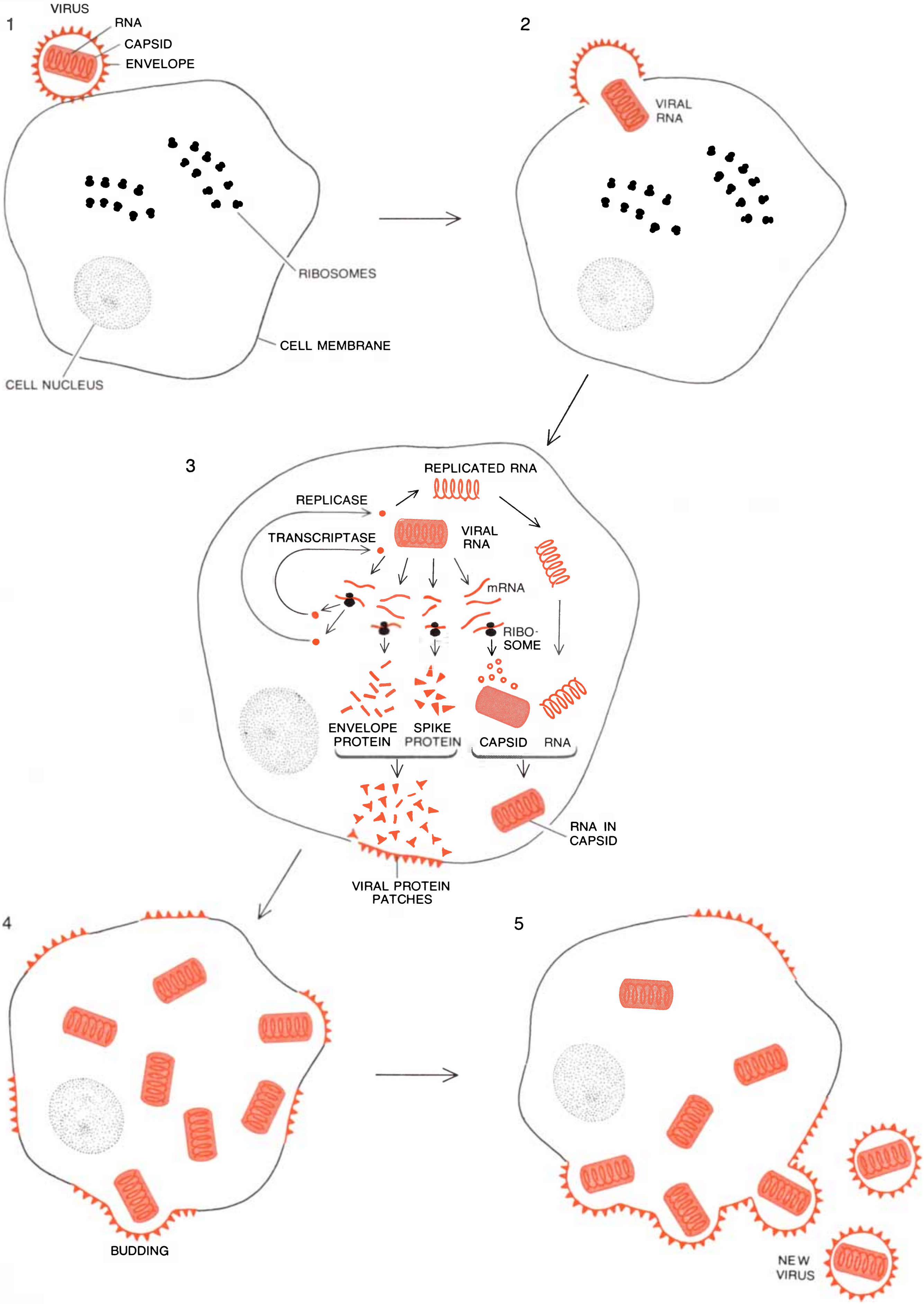
In an acute virus infection such as influenza or poliomyelitis the invading virus takes over the cell's machinery and begins to synthesize new virus components. Eventually a crop of new viruses bursts from the cell, which then dies, and the new virus particles go on to infect other cells. Not all virus infections, however, result in the death of the host cell. It has long been known that the herpes simplex virus that causes cold sores or fever blisters can remain dormant or silent in the host for months, years or even decades. Jack G. Stevens of the University of California at Los Angeles School of Medicine has shown that the herpes virus lies latent within the nerve cells of the ganglia, producing no disease symptoms until stresses (for example sunburn, fever and exposure to extreme cold) trigger the virus and small skin eruptions result. Yet between outbreaks the herpes virus apparently does not destroy the nerve cells in which it continues to exist.

How do some viruses manage to persist in cells and replicate without causing the death of the host cell? Part of the answer lies in the fact that such viruses do not inhibit the host cell's synthesis of ribonucleic acid (RNA) and protein, and this enables the infected cell to continue functioning. On the other hand, slow viruses almost always exhibit an ability to alter the protein composition of the surface of the infected cell's outer membrane. Although many acute viruses are able to do the same thing, from the host's point of view an altered cell surface in surviving cells is much more dangerous than it is in dying cells. When an altered cell remains alive and multiplies, it and its progeny cells are likely to be attacked as "foreign" objects by the immune system. The result may be a disease: perhaps one of the severe neurologic diseases, cancer or one of the autoimmune diseases.

Myxoviruses (influenza viruses) and paramyxoviruses (the viruses of measles, mumps and related diseases) have a set of genes that encodes the manufacture of several viral proteins. Some of these proteins have a strong affinity for the outer membrane of the infected cell. Other viral genes direct the synthesis of viral RNA. The newly synthesized RNA is neatly packaged along with virus-coded protein in small cylindrical tubes in the cell's cytoplasm. These ribonucleoprotein (RNP) tubules apparently provide a protective framework for the viral RNA. As the concentration of the RNP tubules builds up, portions of the cell's cytoplasm can become literally packed with them. Some RNP tubules move to areas of the cell membrane that contain viral proteins. When these virus-coded membrane proteins bud, that is, begin to break away from the cell, the RNP is incorporated into the interior of the budding portion. The bud enclosing the RNP contains the viral genome: the complete set of viral genes. Therefore when the bud breaks free, it can go on to infect another cell [*see illustration on next page*].

It might be expected that since the new virus buds from the cell membrane, the outer coat of the virus would contain cell proteins as well as virus proteins. My colleague E. D. Kiehn and I, working with a radioactive-labeling technique at the University of California at San Diego, found that little or no host-cell protein is incorporated into a budding influenza virus particle but host-cell fats and carbohydrates are taken up. This suggests that the virus-encoded protein completely alters patches of the cell membrane. It is therefore not surprising that in acute infection by influenza virus or measles virus the cells cease functioning as their cytoplasm fills up with viral RNP tubules and their external membrane is infiltrated with viral protein.

In a pioneering series of investigations begun more than a decade ago Robert


1
VIRUS
RNA
CAPSID
ENVELOPE
RIBOSOMES
CELL MEMBRANE
CELL NUCLEUS
2
VIRAL
RNA
3
REPLICATED RNA
REPLICASE
TRANSCRIPTASE
VIRAL
RNA
mRNA
RIBO-
SOME
ENVELOPE
PROTEIN
SPIKE
PROTEIN
CAPSID
RNA
RNA IN
CAPSID
VIRAL PROTEIN
PATCHES
4
BUDDING
5
NEW
VIRUS

Rustigian of the Tufts University School of Medicine has demonstrated that a normally lethal measles-virus infection of cells grown in laboratory tissue culture could become nonlethal under certain conditions. He was able to establish cell cultures that were persistently infected with measles virus. He did so by cultivating measles-infected cells in the presence of measles antibody for a time. After a period of time the majority of the cells became persistently infected. They continued to divide through hundreds of generations and continually transmitted the measles virus at cell division to their daughter cells. Although the cytoplasm of the cells contained typical viral inclusions and large amounts of viral protein, only tiny amounts of infectious virus could be isolated from the cultures. Rustigian concluded that incomplete virus was being produced in the cells.

The medical significance of his findings became apparent in 1969, when measles virus was isolated from brain cells of patients suffering from the brain inflammation called subacute sclerosing panencephalitis (SSPE). This rare and ultimately fatal disease leads to a progressive loss of mental and motor functions. It has been known for decades that SSPE generally occurs months or years after a severe case of measles and that SSPE patients tend to have a greatly elevated level of measles antibody. Attempts to isolate measles virus from the brain tissue of SSPE patients failed, however, until these cells were grown together with measles-susceptible tissue-culture cells. It appears that the measles-virus genome in SSPE brain cells is in a latent form and that it is activated by fusion with susceptible cells. Luiz Horta-Barbosa, John L. Sever, David A. Fuccillo and their colleagues at the National Institute of Neurological Diseases and Stroke found that latent measles virus is also harbored in the lymph-node cells of SSPE patients. Their finding suggests that SSPE is a generalized persistent virus infection involving more of the body than just brain cells.

It is not understood why most infected people recover from ordinary measles whereas a rare few individuals develop a slowly progressing, fatal brain disease. It is clear that, since SSPE patients show an elevated level of measles antibody, they are not deficient in their ability to manufacture the antibody. Antibody that can neutralize mature viruses outside cells, however, may be powerless to combat the spread of the viral genome from cell to cell through cytoplasmic bridges or by cell division. Some investigators suspect that SSPE patients have a more subtle immunological defect: an inability to destroy virus-infected cells that have had their outer membrane altered by the virus. Of course, the virus isolated from SSPE patients by cell fusion is not truly representative of the state of the virus in the cells of the patients. The virus in the patients is very likely in the form of an incomplete RNP particle that is noninfectious. The incomplete virus probably replicates in the same unbalanced way that Rustigian observed in his cultures of cells infected with measles virus.

Multiple sclerosis patients as a group also tend to have a higher measles-antibody level than patients with other neurological diseases. In multiple sclerosis there is a progressive sclerosis, or hardening of patches in the brain and spinal cord, resulting in loss of muscular coordination. John M. Adams and David T. Imagawa of the University of California at Los Angeles School of Medicine have observed inclusion bodies of the measles type in the brain lesions of patients suffering from multiple sclerosis. Measles virus has not, however, been isolated from multiple sclerosis patients by cell fusion or any other technique. If measles virus is a cause of multiple sclerosis, the virus must be replicating in an even more defective way than it does in SSPE.

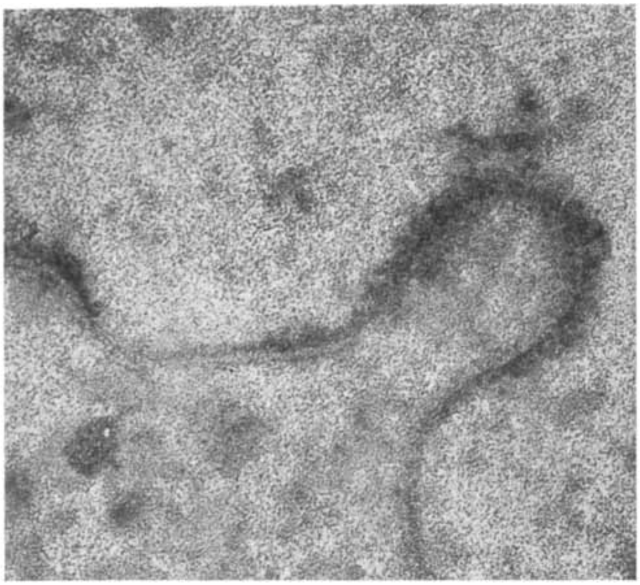
Recently V. ter Meulen of the University of Göttingen and Hilary Koprowski and his colleagues at the Wistar Institute of Anatomy and Biology, using a cell-fusion technique, have isolated parainfluenza virus Type 1 from the brain cells of patients with multiple sclerosis. The parainfluenza virus is quite similar to the measles virus, both being classified as paramyxoviruses.

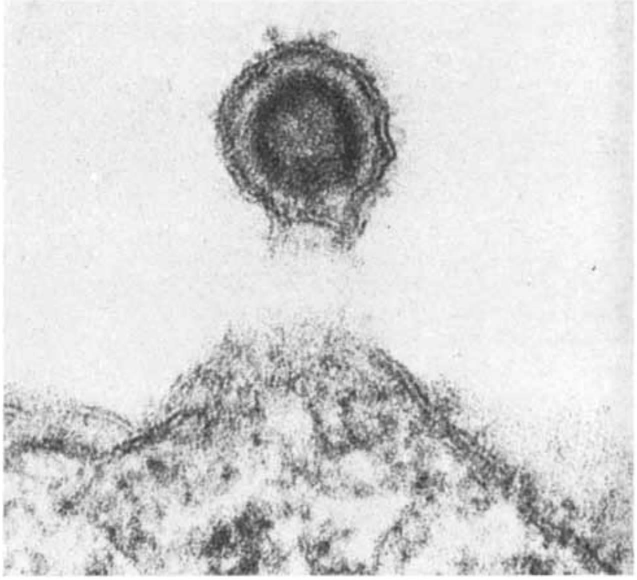
Additional evidence for the involvement of paramyxovirus in multiple sclerosis comes from the discovery of densely packed viruslike tubules in electron micrographs of brain cells from victims of the disease. The viruslike material was found only in tiny early lesions and was not present in larger lesions. The discovery was made by John Prineas of the Department of Medicine of the University of Sydney, who suggests that multiple sclerosis brain lesions may be initiated by periodic seeding of virus-bearing lymphocytes at selective sites in the white matter of the brain.

In 1970 Koprowski and his colleagues reported finding a measles virus and a small DNA virus in the brain cells of some SSPE patients. The DNA virus was similar to viruses known to cause tumors: the polyoma virus (a virus of rodents) and the SV 40 virus (a virus of monkeys). Billie L. Padgett, Duard L. Walker, Gabriele M. Zu Rhein and their colleagues at the University of Wisconsin Medical School had previously observed that there were a large number of SV-40-like virus particles in the brain cells of people suffering from another slowly progressing and fatal brain disease: progressive multifocal leukoencephalopathy (PML). This is a disease of the white substance of the brain; its symptoms include drowsiness, dimmed vision and unsteady gait. Most PML victims have a deficient immune response. Significantly, many of them had previously undergone treatment to suppress their immune responses, either to prevent the rejection of transplanted tissue or for other purposes. In 1971 and 1972 viruses of the SV 40 type were isolated from the brain tissue of PML victims by the Wisconsin workers. This was the first evidence that the SV 40 virus may be involved in human disease.

There are a number of other chronic or slowly progressive diseases of the central nervous system in man that probably have a viral origin. Perhaps the best-understood of these diseases are kuru and Creutzfeldt-Jakob disease. Both are slowly progressing presenile dementias: they result in the premature development of mental deficiencies associated with old age. In both diseases there are numerous vacuoles, or voids, in the brain

INFECTION OF A CELL by an RNA virus can lead to the production of new virus particles by "budding" without killing the host cell. An infectious virus particle becomes attached to a cell (*1*) and empties its genome (RNA in a nucleoprotein capsid, or coat) into the cytoplasm of the cell (*2*). A transcriptase enzyme present in the viral genome gives rise to several types of messenger RNA (mRNA) that attach to ribosomes in the cytoplasm. The mRNA produces envelope protein, spike protein, capsid nucleoprotein and the enzymes replicase and transcriptase (*3*). Replicase combines with the original viral RNA to produce new viral RNA, which becomes incorporated into a nucleoprotein capsid and migrates to the cell's outer membrane. Meanwhile the envelope and spike proteins have also migrated to the outer membrane, where they form discrete patches (*4*). The viral RNA–protein genome aligns itself under the larger patches. When the patches bud, they take up the viral genome (*5*). The particles that break away are complete viruses and can infect other cells.





VIRUS BUDDING FROM CELL MEMBRANE is visible in these electron micrographs. The micrograph at left, made by Garth Nicholson of the Salk Institute for Biological Studies, shows a new influenza virus particle being formed by budding. Dense areas in the cell membrane are patches of viral protein. Micrograph at right, supplied by George J. Todaro and his colleagues at the National Cancer Institute, shows a Type C virus particle budding from membrane of a Chinese hamster cell that had been grown in culture.

cells and a proliferation of glia: the "nurse" cells associated with nerve cells. The gray matter of the brain degenerates and becomes spongy in appearance.

Kuru is found only among the Fore people of eastern New Guinea. The Fore formerly practiced a ritual cannibalism that involved eating the brain and entrails of dead relatives. When cannibalism was outlawed and the practice ceased, the incidence of kuru dropped. Creutzfeldt-Jakob disease is also rare but its distribution is worldwide. D. Carleton Gajdusek, Clarence J. Gibbs, Jr., and their colleagues at the National Institute of Neurological Diseases and Stroke have successfully transmitted kuru and Creutzfeldt-Jakob disease to chimpanzees from the homogenized brain tissue of patients who have died of these diseases. The incubation period for chimpanzees, during which there were no symptoms of infection or disease, was between 10 and 24 months. The duration of the disease (which was invariably fatal) was between one and 15 months. Monkeys at first appeared to be resistant to the disease, but on longer observation it was found that many monkey species were susceptible. In one rhesus monkey kuru appeared 8½ years after inoculation. Gajdusek and Gibbs have also been able to pass the disease serially from chimpanzee to chimpanzee. Even though the transmission of these diseases provides strong evidence that they are caused by viruses, the viruses themselves have not yet been isolated or grown in cell cultures.

Scrapie in sheep and encephalopathy in mink are slow-virus diseases similar to kuru and Creutzfeldt-Jakob disease [see "The Kinship of Animal and Human Diseases," by Robert W. Leaders; SCIENTIFIC AMERICAN, January, 1967]. In all four diseases the transmissible "virus" shows a surprising resistance to heating and to incapacitation by radiation or chemicals. Here again no one has succeeded in isolating the agent of either scrapie or mink encephalopathy, although many efforts have been made.

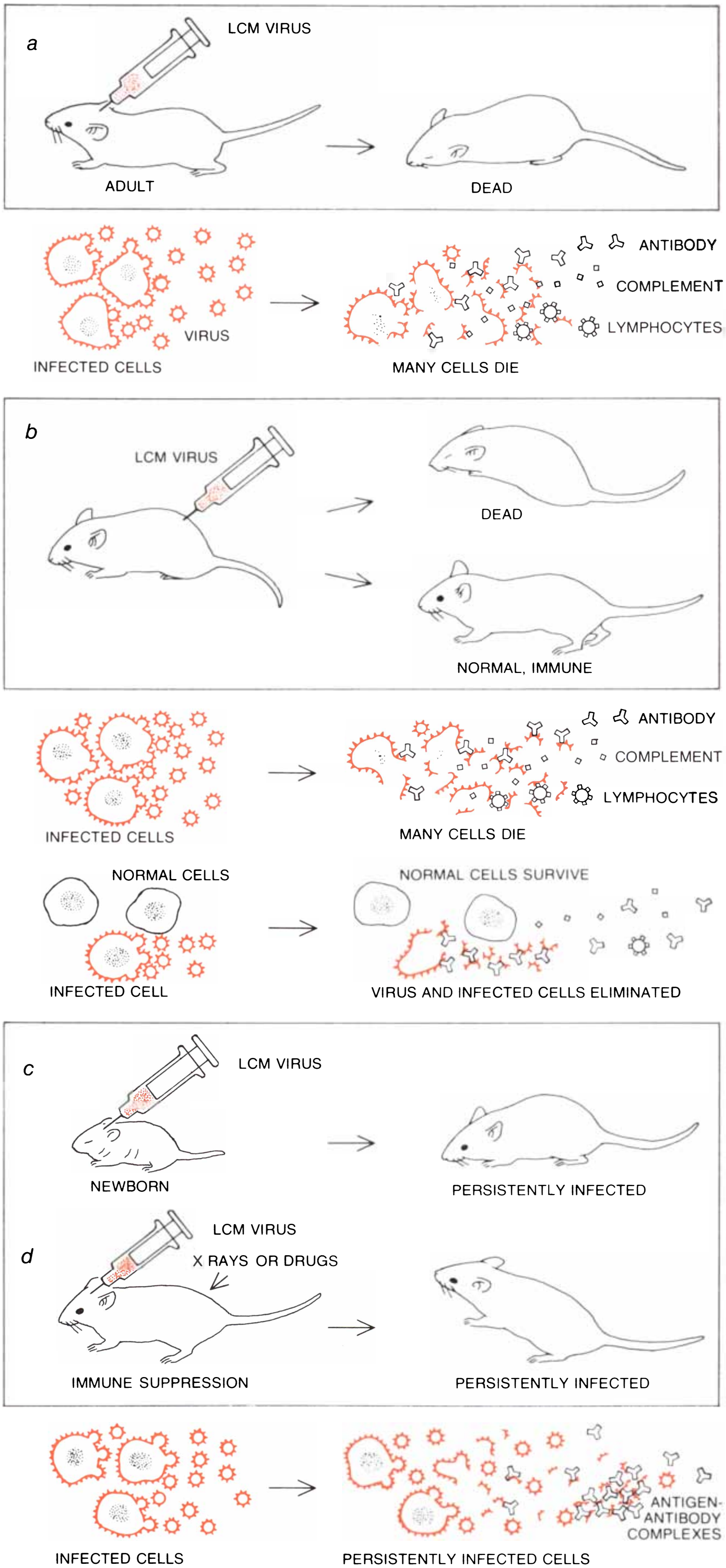
Because of the unusual properties of the transmissible agents implicated in kuru, Creutzfeldt-Jakob disease, scrapie and mink encephalopathy, many investigators believe the causative factor may not be viruses of the usual kind but instead some new type of infectious agent. On the other hand, everything known about the four diseases is quite consistent with the explanation that they are caused by typical viruses that happen to be extraordinarily resistant to the immune defenses of the body. Even the elusive scrapie agent has at long last probably been seen. H. K. Narang and his colleagues at the Newcastle-upon-Tyne General Hospital in England recently reported finding clusters of dense virus-like particles in electron micrographs of the nerve cells of rats infected with scrapie.

Some of the basic principles of slow-virus infection are well illustrated by a disease that is found in wild mice in certain areas of the world and that occasionally infects humans. The disease is lymphocytic choriomeningitis (LCM); it causes inflammation of the membranes surrounding the brain. Some important early investigations of this virus were conducted by E. Traub in Germany, who found that infected mice either rapidly developed a fatal disease or experienced a milder disease followed by a strong immunity and the disappearance of the virus. In a few mice, however, the infectious virus persisted without any apparent sign of infection.

John E. Hotchin of the New York State Department of Health later demonstrated that the development of lifelong infection by LCM virus depends on a suppressed immune response to the virus. Newborn mice inoculated with the virus survive because they lack a fully developed immune response. When the virus is injected into the brain of adult mice, the mice invariably die; when the virus is injected into a vein, the mice either die or recover with full immunity, depending on how many brain cells become infected with the virus and are then killed by the immune response. Hotchin also showed that in adult mice inoculated with the LCM virus, treatment with X rays or drugs to suppress the immune response will give rise to a lifelong persistent virus infection instead of a fatal disease [*see illustration on opposite page*].

It seems that part of the fatal disease process is due to an attack by the body's immune system on the infecting virus and, probably to an even greater degree, to an attack on infected cells that have viral proteins on their surface. Since LCM virus by itself does not kill the cells it infects, it is not surprising that most of the cells in a mouse inoculated with LCM virus become infected and continue to function until they are destroyed by the animal's immune response. This leads to a paradoxical situation in which the immune system causes a "virus disease" even though the virus infection is relatively innocuous.

Hotchin noted that persistently infected mice do not have a normal life span. The mice are usually healthy during the first year after infection, but gradually they develop a disease involving kidney damage similar to what is found in the human disease systemic lupus erythematosus (SLE). Michael B. A. Oldstone and Frank J. Dixon of the Scripps Clinic and Research Foundation in La Jolla, Calif., showed that mice with a persistent LCM-virus infection suffered from an immune-complex disease. The mice produced antibody that combined with the virus to form virus-antibody complexes. These complexes become trapped in the capillaries of the kidneys and produce the inflammatory kidney disease glomerulonephritis. Similar damage to the kidneys of humans with lupus erythematosus has led to a search for viruses that may be involved in the disease. Oldstone and Dixon also



LCM VIRUS, which causes lymphocytic choriomeningitis, an inflammation of the membranes of the brain, kills adult mice when it is injected directly into the brain (*a*). When LCM virus is injected into a vein, adult mice either die rapidly or develop a milder disease followed by immunity to the virus (*b*). In the fatal disease process infected cells manufacture a large number of new virus particles that infect other cells. The response of the immune system is to attack both the virus particles and the infected cells. Death results when a great many infected cells die. If the infection is stopped before a large number of cells are involved, the animal survives and maintains immunity to the virus. When the immune response is weak, as it is in newborn mice (*c*), injection of LCM virus into the brain results in a persistent infection. Treating adult mice with X rays or certain drugs suppresses the immune response, and when they are injected with LCM virus, they also develop a persistent infection instead of a fatal one (*d*). Their immune response is not adequate to kill all the infected cells, so that the animal remains infected for life. The infected cells continue making virus and virus antigens. The antigen-antibody complexes that are formed are carried by the blood to the kidneys, where they can cause severe, progressive damage.

found that a New Zealand strain of mice appears to be genetically predisposed to develop immune-complex disease when infected with certain viruses.

Mink encephalopathy, also known as Aleutian mink disease, likewise affects genetically predisposed animals. The affected animals belong to a homozygous blue-gray strain, that is, a strain in which the gene for this trait is inherited from both parents. The disease triggers a massive immune response that gives rise to circulating virus-antibody complexes, which in turn cause massive tissue damage.

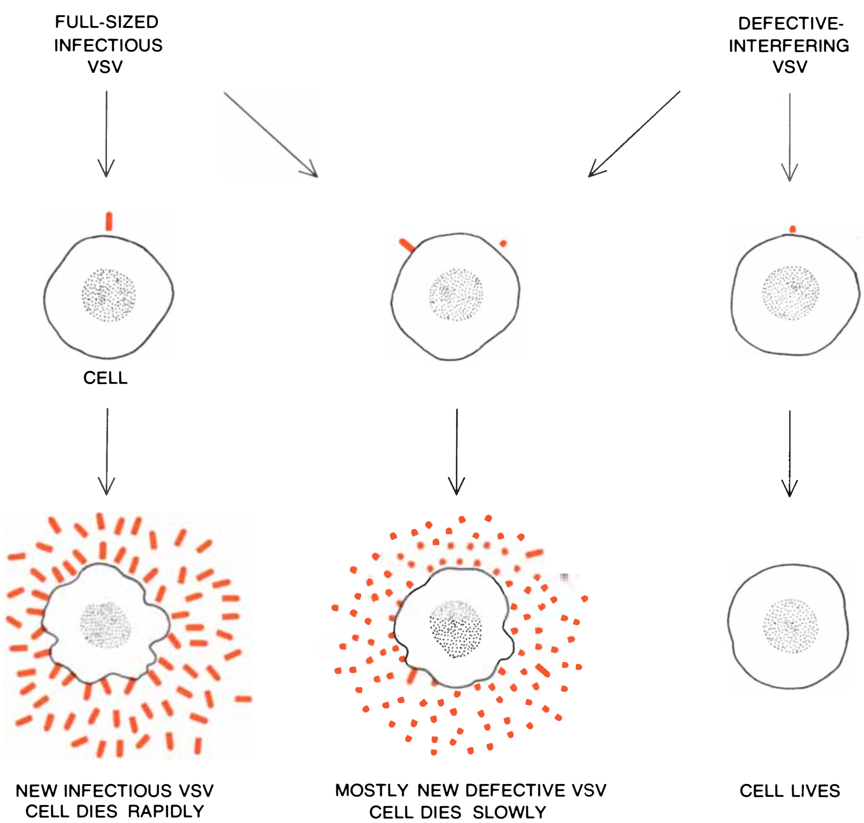
With the discovery that certain animals are genetically predisposed to react too strongly, too weakly or in an unbalanced manner to persistent virus infections, investigators began to wonder if some of the autoimmune diseases associated with aging in man might not also involve a persistent virus infection. Some workers have reported preliminary evidence that implicates viruses as causative agents of rheumatoid arthritis and other degenerative diseases, but definite proof is still lacking.

One of the most interesting aspects of virus infection is the ability of some viruses to depress all or part of the immune system in certain circumstances and to enhance the immune response in other circumstances. Part of the reason may be that some viruses grow in cells, for example lymphocytes, that are involved in the immune response. A wide variety of viruses (LCM, RNA-tumor, measles, influenza and poliomyelitis viruses) can depress the production of antibody against other antigens. Thus in persistent virus infection the virus itself may initiate and maintain persistent aberrations of the immune system, which in turn enable the virus to persist.

Viruses of the herpes group appear to be particularly capable of establishing persistent or recurrent infections in man. Herpes simplex virus Type 2 causes recurrent genital lesions just as Type 1 causes recurrent fever blisters, and evidence is accumulating that recurrent herpes virus genital infection may be the cause of carcinoma of the cervix. It is not clear why most people recover completely from the virus that causes chicken pox (herpes zoster) whereas a few people harbor the virus for life and suffer recurrent painful attacks of shingles (localized chicken pox lesions along sensory nerve paths).

Another herpes virus, the Epstein-Barr virus, has been strongly implicated in the causation of Burkitt's lymphoma, a tumor of the connective tissue of the jaw, and in the causation of nasopharyngeal carcinoma, a cancer of the nose or throat [see "Herpes Viruses and Cancer," by Keen A. Rafferty, Jr.; Scientific American, October, 1973]. Again, patients with infectious mononucleosis manufacture antibodies against the Epstein-Barr virus, yet not everybody infected with the virus develops mononucleosis. Presumably the Epstein-Barr virus can lie dormant in the body for years. It does not regularly cause the death of the cells in which it replicates, but it may transform the cells in such a way that in some individuals cancer eventually develops.

The circumstances that change a virus such as measles, which usually kills the cells it infects, into a persistent nonlethal virus are not well understood. Alice S. Huang of the Harvard Medical School and David Baltimore of the Massachusetts Institute of Technology have suggested that defective-interfering viruses may play a role in establishing persistent virus infection. Defective-interfering viruses are faulty virus particles that lack large portions of their RNA or DNA genome and thus are unable to replicate by themselves; they can, however, replicate in cells that are concurrently infected with the intact virus genome. In our laboratory Michael Doyle, Luis Villareal and I have shown that a defective-interfering virus can convert a virus infection that normally leads rapidly to death into a more slowly progressing disease. When vesicular stomatitis virus (VSV) is injected into the brain of mice, a fatal encephalitis rapidly develops. It usually takes only a few VSV particles to give rise to the fatal disease. Mice given a very large dose of defective VSV along with a small dose of intact VSV, however, develop an immunity and do not develop the disease when they are later injected with a large, lethal dose of intact VSV. Finally, when

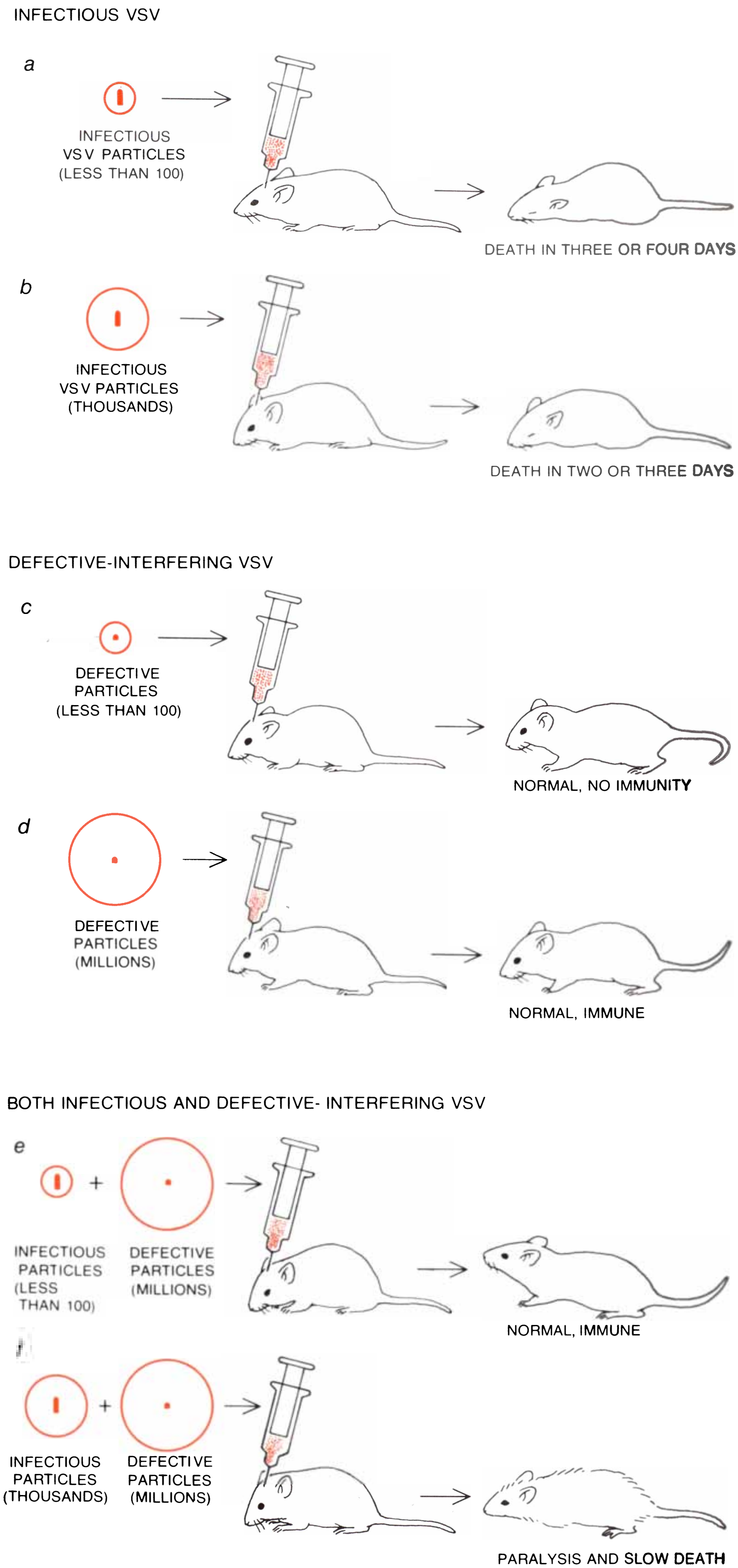


CULTURED CELLS INFECTED BY INTACT VSV PARTICLES manufacture thousands of new intact, infectious virus particles by budding. Normally all the infected cells die (*left*). Addition of defective-interfering VSV to cultured cells does not lead to the production of new virus particles because the defective virus does not have the genetic code for making all the necessary viral proteins (*right*). Cells do not die unless millions of defective particles are added for each cell. When both intact and defective virus infect the same cell, many new defective particles along with a few intact, infectious VSV particles are made (*middle*). The defective virus utilizes proteins made by the intact virus and thus interferes with the rate of production of the intact virus particles. The degree of interference depends on the number of virus particles added and on the ratio of the defective-interfering virus to the intact virus. Cell death ensues from double infection, but it comes relatively slowly.

a very large dose of VSV is injected into mice along with defective virus, the animals do not develop the rapidly fatal encephalitis but rather show a slowly progressing paralytic and wasting disease that leads to death after a relatively long period (two weeks instead of two to three days).

We have also found that in cell cultures the defective virus can alter the pattern of infection by intact VSV. Normally all VSV-infected cells will be destroyed, but with the addition of the defective virus a chronically infected carrier culture can be produced. The carrier cells continue to grow while producing both the intact VSV and the defective VSV particles [*see illustration on opposite page*]. Mutations occur in the virus during its growth in the carrier cultures, and these changes affect the ability of the virus to cause disease in mice. Further study of defective-interfering viruses may yield considerable insight into the mechanisms of slow, persistent virus infections in general.

It is believed the defective-interfering virus competes with the intact virus in the replicative process, in effect borrowing some of the proteins made by the in-



DEFECTIVE-INTERFERING VIRUS does not by itself cause disease but can affect the course of a normally fatal virus disease. Mice injected with a small number (fewer than 100) of complete, infectious VSV particles die in three or four days (*a*). Injection of a larger number of infectious particles speeds up the fatal process (*b*). Mice injected with a small number of defective VSV particles show no noticeable response (*c*). When millions of defective particles are injected, there is also no infection, but an immunity to the virus develops (*d*). The large dose of defective virus acts like a vaccine: it contains an amount of antigen sufficient to stimulate the production of antibodies. When a small but normally fatal dose of the infectious virus is injected with a large protective dose of defective virus, all the injected mice survive and develop an immunity (*e*). The replication of the infectious virus is greatly slowed by the defective virus. This enables the immune system to arrest the infection. Finally, a model for slowly progressing virus disease is found in mice injected with a large dose of infectious virus and a large dose of the defective-interfering virus (*f*). The mice become persistently infected and develop a progressive paralysis of the hind limbs. They rapidly waste away, and most of them die within two weeks. Further study of how the defective virus interferes with normal virus reproduction may provide additional clues to the mechanisms of other slow, persistent virus infections.

tact virus because the defective virus does not contain the genetic information for producing all the necessary viral proteins. This competition interferes with the rate of production of intact viruses. With fewer infectious viruses being produced, the infectious process is greatly slowed.

Many human and animal viruses are known to generate defective-interfering virus particles. These defective viruses probably play an important role in virus evolution. It may be that the defective-interfering virus serves to keep the intact virus from being too virulent; otherwise the virus might decimate the entire population of host organisms and thus eliminate itself. The occasional accidental triggering of slow-virus diseases in host organisms with certain inadequacies in their immune system may be only an untoward side effect of defective virus activity.

Defective-interfering virus has also been found in cell cultures infected with LCM virus. This indicates that defective-interfering virus may be an important factor in LCM disease of mice and in autoimmune and immune-complex diseases.

Further research will probably uncover other molecular mechanisms by which normally virulent viruses can establish persistent, nonlethal infection of cells. It is already clear, however, that it may be misleading to name a virus after the commonest acute infection it causes. This tends to direct attention away from the equally important degenerative processes that can result from chronic virus infection.

DISEASES IN WHICH THERE IS STRONG EVIDENCE OF VIRUS INVOLVEMENT	SOME TYPICAL SIGNS	EVIDENCE
Creutzfeldt-Jakob	Impaired mental function, presenile dementia, usually leads to death	Transmissible agent demonstrated but not identified
Kuru	Impaired mental and motor function, presenile dementia, usually leads to death	Transmissible agent demonstrated but not identified
DISEASES IN WHICH THERE IS PRELIMINARY EVIDENCE OF VIRUS INVOLVEMENT	SOME TYPICAL SIGNS	EVIDENCE
Burkitt's lymphoma Nasopharyngeal carcinoma	Malignant tumors	Epstein-Barr virus implicated
Carcinoma of the cervix	Malignant tumors	Herpes simplex virus implicated
Leukemias	Blood cancer	Tumor-causing RNA virus found in animals, preliminary evidence in man
Lupus erythematosus	Inflammation of the skin	Paramyxoviruses implicated
Multiple sclerosis	Loss of muscular coordination	Measles and parainfluenza viruses implicated
Progressive multifocal leucoencephalopathy	Drowsiness, unsteady gait, dimmed vision, usually leads to death	Virus of SV 40 type isolated
Subacute sclerosing panencephalitis	Impaired mental and motor function, usually leads to death	Measles virus and in some cases viruses of the SV 40 type implicated
SOME OTHER DISEASES WITH POSSIBLE VIRUS INVOLVEMENT	SOME TYPICAL SIGNS	EVIDENCE
Amyotrophic lateral sclerosis	Muscular atrophy, usually leads to death	Scattered preliminary information; no firm evidence as yet
Diabetes	Loss of pancreas islet cells, lack of insulin, excessive sugar in blood	
Glomerulonephritis	Inflammation of kidneys	
Hemolytic anemias	Red blood cells destroyed	
Muscular dystrophy	Progressive muscular atrophy	
Optic neuritis	Optic-nerve damage leading to blindness	
Parkinson's disease	Tremors	
Psoriasis	Scaling of skin due to rapid cell growth	
Rheumatoid arthritis	Inflammation of joints	
Ulcerative colitis	Chronic and recurring ulceration of colon	

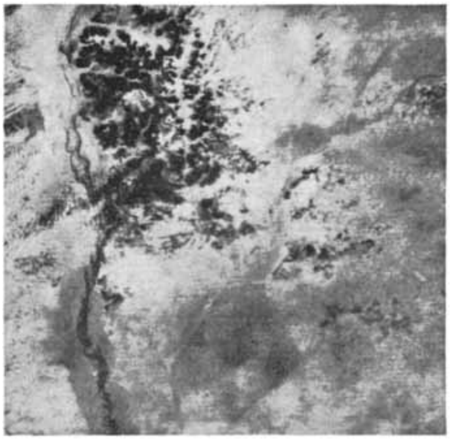
DEGENERATIVE DISEASES listed here are some of diseases in man that deserve thorough study before ruling out slow or inapparent virus infection as a causative factor. In any of these diseases, however, isolation of a virus from patients does not necessarily prove that the disease is caused by the virus, which may be present fortuitously or because of a failure in the patient's immune system.

Know the land and the people*

Land

Imagery from satellites and aircraft utilizes our products. Let the benefits fall where they may.

The Sahara is expanding. Millions are suffering in West Africa. Where the sand blows, the land was once green, just as all the Sahara was in prehistory, before man turned from hunting to herding.



On April 20, 1973, NASA's ERTS caught this image from 567 miles high over the Mali-Niger border. Both are large countries drying up. Note the faint hexagon at bottom center. Its great significance was not understood. Turned out to have been drawn in five years with barbed wire that kept ruminants out long enough for grass seed dropped by the wind to do its work. Now the Niger government may convince its herdsmen citizens that barbed wire and range management threaten their way of life far less than would acceptance of employment as haulers of concrete for damming rivers.

In the U.S.A., where beautifully marbled beef is much prized, the feedlots where that beef is finished for market carry up to 200 animals per acre. Herding reaches its pinnacle. The herdsman must now worry about what comes out of the beasts as well as what to put into them. A stationary 100,000 cattle on 500 arid acres limits ground cover to salt-tolerant vegetation. In non-arid country run-off may pollute water supplies far away. Vegetation patterns, drawn by nutrients and salt and photographed from the air, can guide the feedlot operator on expanding or moving the operation.

Or so it would appear from a paper entitled "Remote Sensing for Detecting Feedlot Runoff." The American Society of Photogrammetry chose it for an award as the best photogrammetry paper by a college undergraduate. For a reprint write Kodak, Dept. 55W, Rochester, N.Y. 14650.

People

How to give them a better-than-offhand talk:

1. Think about:

What do they have in common?
Why are they coming?
How much do they know already?
How much more do they want to know?
What attitudes will they bring with them?
What do I want to change or accomplish with my message?

2. Get a pack of 4″ x 6″ index cards.

3. Draw a large box in the upper left-hand corner of a card. In that box draw a crude sketch of what comes into your mind when you concentrate on one of the principal points you want to make. It may be a chart, clipping, symbol, diagram, or a photo of a person, place, or thing. Underneath state the point in as few words as needed to cue yourself to the thought.

4. Do a similar card about the thought that leads into the thought you have just expressed. Now do one about the thought that follows the first one. Keep going like that.

5. When you run out of ideas to tack on ahead or behind, think of the important points that haven't fallen into sequence yet. Make cards for them. **Always work up the sketch before the words.** (If lively words flow out of you too easily to work that way, probably anything you'd say on any subject would fascinate any audience. In that case you hardly need any of this advice.)

6. Arrange the cards on a table in an order that makes sense.

7. Now get critical. Is the development of the ideas too plodding? Would some other scheme of arranging the cards liven up the beginning and the end? Which cards should be tossed out? Where are you skipping too fast? Where are you trying to pack too much into a single card? Make out the additional cards you need.

8. Now get critical about your stack of cards from the standpoint of practicality. Some of your sketches would take too much time and art talent to turn into presentable slides. Substitute images easier to obtain from internal sources or the public domain. If you are on your own for this, a KODAK EKTAGRAPHIC Visualmaker kit can much simplify your slide-making problems, both in copying extant material and in snapping originals.

9. Run through the talk. Make believe your sketches are already slides on the screen. Speak from the cues you've written under the sketches. (When you give the actual talk, the slides themselves may suffice as cues. Then you wouldn't be reading at all. Why assemble an audience just to hear you read?)

10. Decide whether you have too much or too little material. Discard or add cards accordingly.

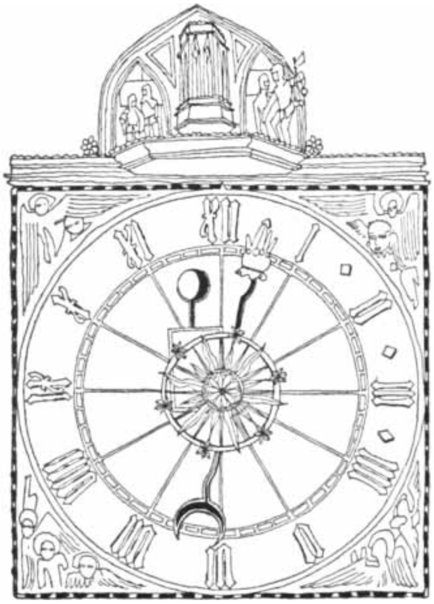
11. Now you are ready to prepare your slides. Worst of the pitfalls is type or other detail that the important but shy people in the back row can't quite make out on the screen. To avoid this and other mistakes, send for the free Kodak publication "Slides with a Purpose." Address request to Dept. 55Z, Kodak, Rochester, N.Y. 14650. We'll throw in literature about the KODAK Visualmaker and other handy products.

12. You're on!

13. You're great!

Kodak

*A service mark of Eastman Kodak Company. It will be our theme for the nation's bicentennial. It summarizes much of our reason for existence.



SCIENCE AND THE CITIZEN

Grow, No-Grow

Two extreme positions have been established in the current debate on economic growth, and Ronald G. Ridker of Resources for the Future suggests that both of them "are wrong—indeed, that they border on the irresponsible." The pro-growth school, he writes in *Science,* holds implicitly that material economic growth is the primary social goal, taking precedence over equity and measures to cope with the social and environmental costs of growth. The no-growth school, on the other hand, seems to hold that social problems will disappear if growth ends. "The relevant question," he says, "is not whether to grow or not to grow, but how to channel and redirect economic output . . . in ways that will make it better serve humanity's needs." Ridker grants that the earth is finite, that the law of entropy cannot be violated and that growth must therefore someday cease. But when? It makes a difference if the limitation must come now or in 100, 1,000 or 100,000 years.

Ridker criticizes the well-known study *The Limits to Growth,* which suggests that the limitation must come quickly, on three grounds. Its model, he says, does not include important adjustment mechanisms: prices, which redirect exploration, research and consumption; government, which can do similar things; simple learning from experience, which leads people to change their behavior as they learn its potential consequences. Then too, the model is based on extreme aggregation, so that it deals with a composite "output," one "resource," one "pollutant" and the world as a whole. This compounds the lack of adjustment mechanisms by eliminating the possibility of substitutions by producers, consumers and society and thereby reducing the choice to one between growth and limitation of growth. Finally, Ridker maintains, *The Limits to Growth* makes pessimistic assumptions about technological progress, reserves, pollution control and population growth. Changing these assumptions makes a significant difference in estimates of when growth must stop even in the absence of the many adjustment mechanisms that would come into play.

For most problems associated with growth, Ridker writes, direct attacks on the problems are probably better than indirect attacks on economic or population growth. He cites a Resources for the Future study's conclusions with regard to various forms of pollution. In each case the amounts of pollutants generated and actually emitted would increase substantially by the year 2000, even assuming low economic and population growth. In the presence of an active abatement policy, however (defined as the application by 2000 of standards that have been recommended by the Environmental Protection Agency for 1976), emissions would be cut well below present levels even in high-growth projections—and at a cost of less than 2 percent of the G.N.P. in the year 2000. In other words, a direct attack through pollution abatement is more effective than a generalized "meat ax" restriction of growth.

There are large areas of ignorance, Ridker concedes, and he asks: "Should we not, in effect, stop the ship, or at least slow it down, until we know more about what lies ahead in the fog?" No, since we are ignorant not only about the negative developments that may accompany growth but also about positive developments, which may improve the well-being of future generations in addition to avoiding disaster. It is therefore "not obvious that the prudent course is to save resources for future generations, at least not obvious to any but the most affluent on this earth."

If society fails to attack such growth-associated problems as pollution directly, then they will indeed build up to the point that they do call for reductions in growth, Ridker writes. Advocates of no-growth may believe that to be the situation; Ridker would hope to prove they are wrong. To do so, however, requires bypassing the conventional debate on limiting growth. "Both schools, it seems to me, are copouts. What we must do is get on with the solution to the problems that obviously and directly face us."

Outward Bound

Having been accelerated from 38,000 to 83,500 kilometers per hour as a result of its close encounter with Jupiter in December, *Pioneer 10* is racing outward through the solar system, leaving experimenters on the earth with reels of magnetic tape for future study. Technically the most notable finding was that a carefully designed spacecraft can penetrate the intense radiation fields of Jupiter and survive—barely. *Pioneer 10* will now continue to monitor interplanetary conditions, reaching the orbit of Saturn in 1976 and three years later the orbit of Uranus, where its radio signals will finally become too weak for detection. By 1983 it will cross the orbit of Pluto, headed for outer space at an "interstellar cruising speed" of 46,000 kilometers per hour. Barely a quarter of a century after the launching of *Sputnik I,* the first artificial satellite, an object made by man will be on its way to the stars.

Pioneer 10 revealed that the charged particles trapped in Jupiter's magnetic field produce radiation belts 10,000 to a million times stronger than the earth's radiation belts. The spacecraft's survival was a narrow thing. Designed to withstand between 10^4 and 10^5 rad units (on a scale in which 500 rad units are lethal to man), *Pioneer 10* was subjected to 5×10^5 rad units in the form of energetic electrons and 5×10^4 rad units in the form of protons. The electron flux was at the high end of the expected range, based on Jupiter's radio emission. There had been no way, however, to guess the proton flux. The maximum measured flux of protons with energies above 30 megavolts was 4×10^6 protons per square centimeter per second, a value considered impressively high. The maximum flux was recorded 250,000 kilometers from the planet.

Jupiter's radiation belts form a flat disk some four million miles in diameter. If such a disk were visible from the earth, it would equal the apparent diameter of the moon. The disk appears to be tilted, indicating that the planet's magnetic pole is located about 15 degrees from its pole of rotation. As a result the disk of trapped particles wobbles up and down about 30 degrees with every 10-hour rotation of the planet itself. Jupiter's magnetic field is opposite in polarity to the earth's and about 10 times as intense.

Pioneer 10's ultraviolet instrument has revealed for the first time that Jupiter's atmosphere contains helium. When all the data are reduced, it should be possible to determine the planet's hydrogen-to-helium ratio, which in turn should shed light on the origin of the solar system. More detailed knowledge of the planet's atmosphere awaits analysis of photopolarimetry measurements and of the radio signals that passed through the atmosphere when the spacecraft swung briefly behind the planet.

The photo-imaging device and other instruments show that the bands in Jupiter's atmosphere range from brownish-orange to blue. The brighter blue bands are at greater altitude and some 15 degrees Fahrenheit colder than the darker brown bands. The coldest clouds at the top of the atmosphere are −230 degrees F. The famous red spot shows up clearly in many of the images produced by *Pioneer 10*. There is now some evidence that the spot is a free-floating vortex that may extend several kilometers above the surrounding clouds.

The celestial-mechanics experiment based on precise measurements of *Pioneer 10*'s trajectory has made it possible for the first time to discriminate clearly between the mass of Jupiter itself and the mass of its four largest satellites: Io, Europa, Ganymede and Callisto. The entire Jupiter system (planet plus moons) is heavier than was previously thought by .004 percent, an amount roughly equal to twice the weight of the earth's moon. Io has long been known to have the highest surface reflectance of any body in the solar system. It is also brightest when it first emerges from Jupiter's shadow. *Pioneer 10* has now discovered that Io has a thin atmosphere, which suggests that when the moon passes behind Jupiter and cools off, some kind of "snow" (perhaps methane) precipitates out on the surface but soon vaporizes when the moon emerges into sunlight.

Federal Energy

As the Federal Government moves to make the U.S. self-sufficient in energy, particularly by encouraging the development of alternatives to petroleum, it appears to hold much of the leverage in its own hands because it owns a substantial part of the resources that might be developed. Federal holdings include about 80 percent of the land with oil shale, 65 percent of the land with geothermal energy and a substantial part of the land with low-sulfur coal. Moreover, through its control of offshore areas the Government has a major role in the development of additional sources of petroleum and natural gas.

Until recently the Government has moved slowly and with apparent reluctance to make these sources of energy available. Over the past few weeks, however, signs of a change in policy have appeared. The Department of the Interior announced that it would lease to the highest qualified bidders some 50,000 acres of known geothermal areas in California for the production of geothermal energy and would accept applications for leases allowing the exploration of potential geothermal areas. In another move the department opened 7.7 million acres of California's ocean shelf to exploration for oil and gas.

With a view to improving the generation of electricity from geothermal energy the U.S. Atomic Energy Commission has provided money for work aimed at producing electricity from hot water, hot brine and hot rock. The only commercial geothermal field in the U.S. (at The Geysers in California) generates electricity from naturally occurring steam, which is clean and easy to work with. Much geothermal energy, however, is in the form of hot water or hot rock rather than steam and so presents technological problems such as corrosion and precipitation. Three laboratories under contract with the AEC are attempting to solve the technological problems. The Lawrence Livermore Laboratory will try to demonstrate by 1980 that the unprocessed and highly salty brines of the Salton Sea Trough can be put through a turbine with high efficiency and minimal corrosion. The Lawrence Berkeley Laboratory is investigating the production of energy from hot brines with low salt content. The Los Alamos Scientific Laboratory is studying the technology of extracting energy from hot rock.

Palomar in Space

The National Aeronautics and Space Administration, in collaboration with some 40 leading astronomers and engineers, is planning a 120-inch reflecting telescope to be launched into orbit around the earth in the early 1980's. The instrument, known as the Large Space Telescope, will probably be the first large payload put into orbit by the "space shuttle." The construction, capabilities and astronomical usefulness of the Large Space Telescope were discussed at the 12th Aerospace Sciences Meeting of the American Institute of Aeronautics and Astronautics in Washington at the end of January.

The telescope will have a primary mirror of either Cer-Vit or ULE (ultra-low expansion) fused silica; both materials are resistant to the expansion and contraction that would result from temperature extremes in space. The Martin Marietta Corporation reported that it is developing a system for pointing the telescope and guiding it with so little vibration that theoretically it could make measurements to within .0005 second of arc. Such accuracy would actually be 10 times better than is required, because of the natural limit imposed on the telescope's optical system by the diffraction of the light entering it.

Since the observatory will be above the earth's atmosphere, star images will appear to have angular diameters of the order of .05 second of arc or even smaller, and the images will be free of distortion due to atmospheric turbulence. One program will be the observation of the positions and motions of stars, particularly double stars, with a precision unattainable by telescopes on the ground. The telescope will also make observations at wavelengths that are filtered out by the earth's atmosphere. The infrared wavelengths will be open to study. It will also be possible to observe interstellar gas, very young stars and very old stars at ultraviolet wavelengths.

Within the solar system the Large Space Telescope will be suited for monitoring Martian weather and the atmospheric dynamics of Jupiter. It will also be effective for gathering information about the size, shape, surface details, rotation and other properties of asteroids, comets and planetary satellites. Cosmology will benefit greatly from the telescope. It will be possible to determine the precise distance of galaxies that are within 100 million parsecs. With these distances in hand, astronomers will be able to map out the recessional velocities of galaxies to see if there are any irregularities in the Hubble expansion of the universe that could be interpreted as evidence for a deceleration of expansion.

The Large Space Telescope will be utilized to study quasars and their spectra in the ultraviolet wavelengths as well as the visible wavelengths. Such observations may clear up current doubts about how quasars are distributed in space and how their red shift is produced. Moreover, because of its ability to probe the universe to magnitudes 100 times fainter than can be reached from the ground, the telescope should reveal many new objects in other galaxies.

Lost Microscope

One of the lesser known casualties of recent cutbacks in the U.S. Government's support of basic research has been a promising new form of microscopy based on the harnessing of the X-ray portion of the powerful synchrotron radiation emitted by a high-energy electron accelerator. The prototype synchrotron X-ray microscope—constructed more than a year ago by Paul Horowitz, John A. Howell and their associates at Harvard University following a suggestion from E. M. Purcell—is now regarded by its inventors as "a finished chapter," owing to the permanent closing in June of the Cambridge Electron Accelerator. The Harvard physicists had sought funds (estimated at between $500,000 and $700,000 per year) to convert the accelerator to the full-time production of synchrotron radiation, but their proposal was rejected by the National Science Foundation for budgetary reasons.

The synchrotron radiation from an electron accelerator contains some three to six orders of magnitude more "soft" X radiation than what is available from conventional X-ray tubes. The radiation is emitted tangentially to the circular orbit of the electrons in the form of a fan-shaped beam that can be focused (by means of an ellipsoidal quartz mirror) and collimated (by means of a pinhole) to construct a scanning X-ray microscope "of extremely simple design." The microscope, the operating details of which were presented by Horowitz and Howell in *Science*, is described as "capable of making stereoscopic element-discriminating pictures of relatively thick specimens in an atmospheric environment."

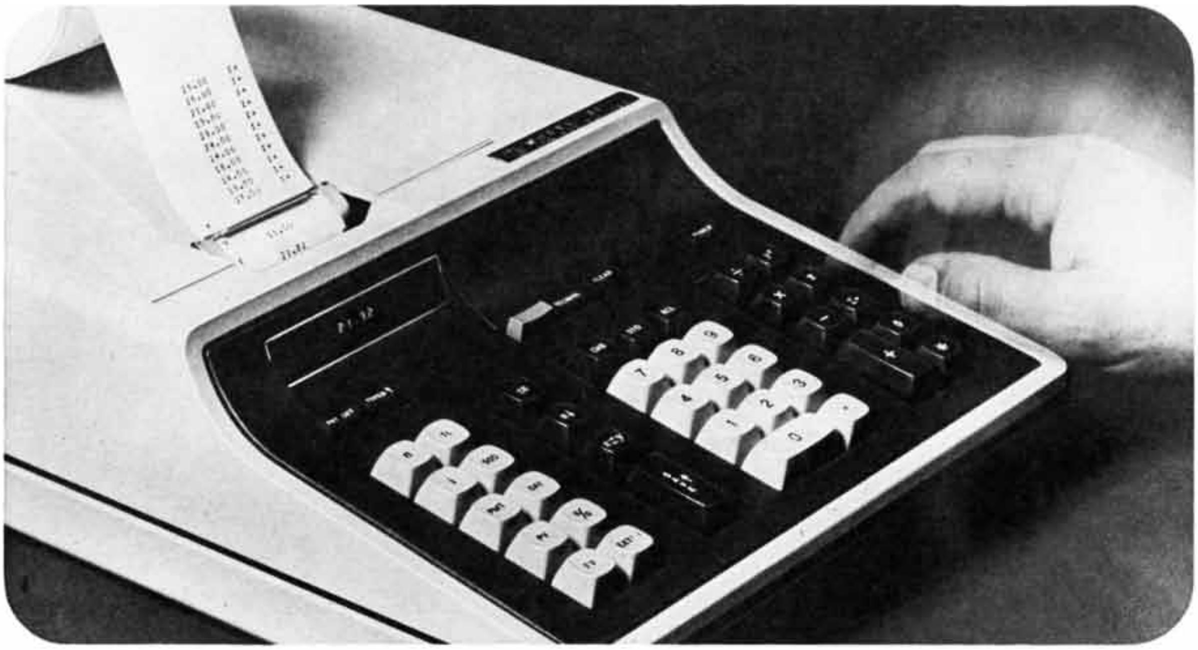
The original *Science* article by Horowitz and Howell also summarized the work they did during the first several weeks of the microscope's operation. "The idea," Horowitz reports, "was to encourage people in diverse fields who had problems for which the microscope might be useful to contact us. Since that time we have looked at a number of specimens people sent us—microscopic sea shells (ostracods), diatoms, slices and pieces of human tissue, integrated circuits and so forth—which will probably form the basis of a second article."

In spite of the closing of the Cambridge Electron Accelerator and the abandonment of the conversion scheme that would have enabled the accelerator to supply enough synchrotron radiation for dozens of simultaneous experiments Horowitz is not entirely discouraged. "It is quite possible," he says, "that the idea might inspire someone at one of the other synchrotron-radiation machines to carry on, and I think this would be good, since the microscope has great promise, combining as it does chemical-element discrimination, enormous depth-of-field, minimal radiation damage, operation in an atmospheric environment, moderate resolution and good penetrating power."

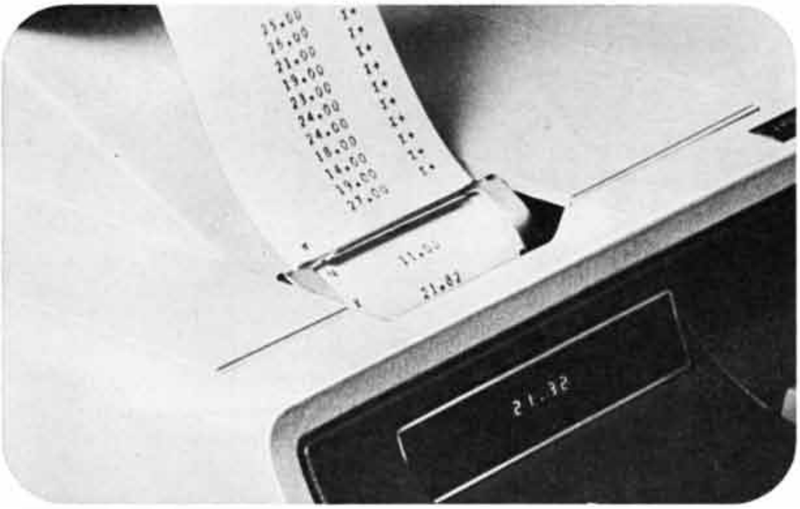
Socialized Medicine

The fact that public funds already pay for a substantial portion of medical care in the U.S. is often overlooked in the continuing controversy over the equitable delivery of health services. The government role has long been particularly large in New York City, and its current dimensions are revealed by a study conducted by Nora Piore and Purlaine Lieberman of the Columbia University School of Public Health and the Center for Community Health Systems. In 1971, they found, public expenditures accounted for 72 percent of the care rendered in all the city's hospitals and nursing homes. The figure was 50.5 percent in 1960 [see "Metropolitan Medical Economics," by Nora K. Piore; SCIENTIFIC AMERICAN, January, 1965].

Between fiscal 1961 and 1971 appropriations of public funds (at all levels of government) for personal health care of all kinds increased from $529 million to $2.45 billion. There was a significant shift in the source of those public moneys, due primarily to the advent during the 1960's of Medicare and Medicaid. Whereas New York City tax dollars accounted for 44 percent of the appropriations in 1961, municipal taxes paid only 27 percent of the bill in 1971; the share paid by state taxes also decreased; the difference was made up by the Federal Government, whose contribution increased from 18 percent in 1961 to 43 percent in 1971. The new Federal programs also produced a shift in the institutions providing services. Medicare, in particular, made it possible for more older patients to be treated in voluntary and proprietary hospitals. As a result the number of patient-days increased by 20 percent in the voluntary hospitals and 19 percent in the proprietary ones between 1960 and 1971, whereas the number of patient-days in municipal hospitals decreased by 15 percent.



The HP-81: The new and powerful computer calculator that's pre-programmed to perform 63 functions specifically related to business and finance, printing solutions in words, numbers and symbols.



The HP-81's clearly labeled tape is a permanent record of your calculations. This printout of a standard deviation problem illustrates how the HP-81 automatically annotates each entry with easy-to-read symbols identifying each operation.

The HP-80: The pocket-sized computer calculator that started the revolution in time and money calculations. For professionals in business and finance. $395*

The HP-81 helps you make business a little more of a science.

The plethora of calculators now on the market can baffle potential buyers with a confusing array of choices. Fortunately, HP continues to provide clear choices when it comes to calculators for science and engineering, and for the world of business and finance.

For the latter, the HP-80 is still the only pocket-sized computer calculator with 40 pre-programmed business functions. And now there's the new HP-81, which applies even more computation power to the business of evaluating the complex relationships between time and money.

The HP-81 is considerably more than a desk-top HP-80 with a printer. Its added capabilities include 63 pre-programmed functions for performing – in a few seconds, and in a few keystrokes – complex business problems that used to take up to 20 minutes to calculate manually. For example, with the HP-81 you can compute and print a 30-year mortgage table in 83 seconds, get a coupon equivalent yield in 21 seconds, or figure a discounted cash flow in 14 seconds. All without tables, calendars or formulas. They're all built in.

The HP-81 contains two 200-year internal calendars (360 and 365 days), 20 storage registers for saving dates or numbers for repetitive calculations (and for automatic multiplication or division by any constant you choose), and 10 special modes for bond and note calculations. Best of all, the HP-81's flexibility helps you explore alternate approaches to a problem. You can check variables using a different date or number, without re-entering the entire problem, and get fast assurance that you have come up with the best decision.

The time and money saved by one or two profitable decisions could easily repay the $925* purchase price. An LED display and high-speed buffered keyboard are optional. For more details on HP's newest time and money machine, send in the coupon.

Finally, a professional quality scope at a practical price.

In oscilloscopes laboratory quality means high performance. Almost every designer has one on his bench. But service and production troubleshooters don't have it as good. They too could often use a high performance scope, but can't risk this expensive and delicate instrument in their more demanding environment.

Now comes the HP 1220A to plug this price/performance gap. This entirely new oscilloscope combines industrial-quality ruggedness with professional performance at a practical price.

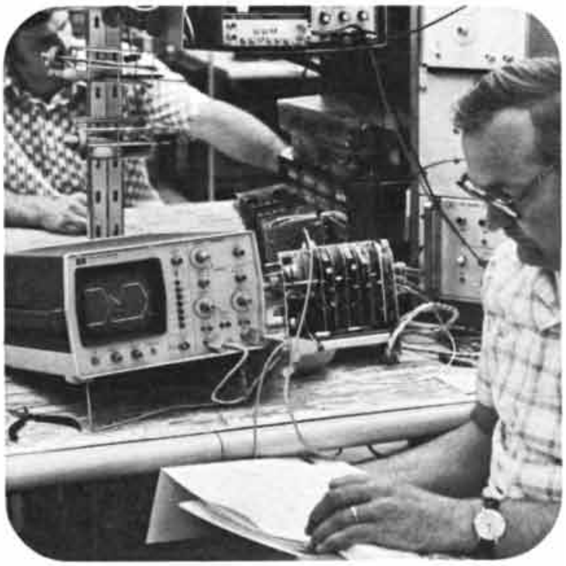
The new scope's large (8 x 10 cm) parallax-free display makes all troubleshooting easier. Because the display is calibrated horizontally as well as vertically, it's easy to locate timing circuit problems and accurately measure ac and dc voltages of waveforms.

TV servicemen don't have to fuss with frame or line triggering because the scope does it automatically. And when they want to pull out vertical interval test signals, they just turn the times-ten magnifier control.

Tough enough for the production line or service truck, the scope's rugged one-piece frame protects it against shock, and its unvented cover shuts out dust and moisture. So stable that it has only one external screwdriver adjustment, it rarely requires recalibration.

And in the classroom, the scope's fast set-up gives the instructor more time to teach. The automatic trigger always displays a baseline and immediately locks on any signal of the correct polarity and amplitude. When the student loses the signal, he can bring it back on screen just by pushing the beam finder button.

The dual-channel Model 1220A has a 15 MHz bandwidth and 2 mV/cm sensitivity, more than enough for audio, video, logic and control circuit measurements even at low levels. The 1220A costs $625*.



For more information on any of the above write to us. Hewlett-Packard, 1505 Page Mill Road, Palo Alto, CA 94304.

*Domestic USA prices only. 00405

Mail to: Hewlett-Packard, 1505 Page Mill Road, Palo Alto, CA 94304
Please send me information on:
( ) HP 80 Calculator ( ) HP 81 Calculator
( ) HP 1220A Oscilloscope

Name________ Title________
Company________
Address________
City______ State______ Zip______

HEWLETT hp PACKARD

Sales and service from 172 offices in 65 countries.
Palo Alto California 94304

THE NATURE OF COMETS

Although Comet Kohoutek was a disappointment to visual observers, it provided a good opportunity for the study of one of the objects that may be relics of the cloud from which the sun and the planets formed

by Fred L. Whipple

Comet Kohoutek, like other comets, is a celestial fountain spouting from a large dirty snowball floating through space. The fountain is activated and illuminated by the sun. It is greatly enhanced because it is spouting in a vacuum and essentially in the absence of gravity. We see the fountain as the comet's head and tail. The tail can extend for tens of millions of miles, but we never see the snowball, whose diameter is only a few miles.

The word "comet" comes from the Greek *aster kometes,* meaning long-haired star. The tail of the comet is of course the hair; the head, or coma, of the comet could be considered the star. Within the coma is the snowball: an icy nucleus that moves in a huge orbit under the gravitational control of the sun. The nucleus spends almost all its lifetime at great distances from the sun, hibernating in the deep freeze of space. When its orbit swings it in toward the sun, its surface begins to sublime, or evaporate, and the sublimated gas flows into space. Pushing against the weak gravity of the relatively small nucleus, the outflowing molecules and atoms carry with them solid particles. Thus does the nucleus give rise to the gaseous and dusty cloud of the coma.

The sun floodlights the dust and gas of the coma, making the comet visible. Some comets are very dusty. Most of their observed light is simply sunlight scattered by the dust and is slightly reddish. Other comets contain little dust. Since molecules and atoms in a gas scatter light feebly, such gaseous comets become bright only through a double process. First the ultraviolet radiation from the sun tears the molecules apart; water, for example, is dissociated into hydrogen (H) and the hydroxyl radical (OH). Then the atom or the broken molecule can fluoresce, that is, absorb solar light at one wavelength and reradiate it at the same wavelength or (more usually) at a series of longer wavelengths.

Almost all the light from gaseous comets comes from such bands of wavelengths, which are mostly emitted by broken molecules of carbon, nitrogen, oxygen and hydrogen such as CH, NH, NH_2, CN and OH, and also C_2 and C_3. What are the parent molecules that split up to produce these unstable radicals? Ammonia (NH_3) and methane (CH_4) are prime suspects, but the suspicion has not yet been confirmed. There is much doubt about the parent molecule for CN. Could it be cyanogen gas (C_2N_2)? Or hydrogen cyanide (HCN)? Or possibly some even more exotic molecule?

Dust Tails and Ion Tails

Regardless of the answers to such questions, it is now understood that the coma of a comet shines with sunlight scattered by dust or with sunlight reradiated by fluorescent gas, usually with both. The tail of a comet is created by another action of the sun. Comet tails, like comet heads, have a gaseous component and a dusty one. For dust tails the action of the sun is uncomplicated: the radiation pressure of sunlight pushes the dust particles out of the coma. Following the laws of motion for orbiting bodies, the dust particles lag behind the coma as they stream away from it; therefore they form a curved tail that can be rich in detail.

Most comets, particularly the brightest, display a huge tail that is only slightly curved. Like the gas in the coma, these tails shine by fluorescence. The molecules responsible for the radiation, however, are ionized, that is, electrons have been removed to leave molecules with a positive electric charge. In such ion tails we find ionized carbon monoxide (CO^+), carbon dioxide (CO_2^+), nitrogen (N_2^+) and the radicals OH^+ and CH^+, but no un-ionized molecules or radicals. Sunlight can ionize some of the molecules, but what pressure can be responsible for pushing them back into space with forces sometimes greater than 1,000 times the gravity of the sun?

The question of how the ion tails are made was long a mystery and has been solved only in the era of space exploration. Space probes have sent back data showing that the sun continuously ejects

COMET KOHOUTEK was photographed on January 11 with the 42-centimeter (16½-inch) Schmidt telescope of the Catalina Observatory on Mount Lemmon in Arizona. The photograph was provided through the courtesy of R. B. Minton of the Lunar and Planetary Laboratory of the University of Arizona, who made numerous exposures of Comet Kohoutek with the instrument. This exposure was made by Steven Kutoroff; it lasted for 10 minutes between 01 : 59 and 02 : 09 Universal Time. The diagonal streak running across the photograph at lower left was made by an artificial satellite; the waxing and waning brightness of the streak is due to the fact that the satellite is irregular in shape and reflects more or less sunlight as it tumbles in its orbit. The photograph was made after perihelion, the comet's closest approach to the sun, so that here the comet is moving in roughly the same direction as that in which the comet's tail is streaming away from its head. Observations of the comet at radio wavelengths show that it contains the "exotic" molecules methyl cyanide (CH_3CN) and hydrogen cyanide (HCN). Since molecules of this kind form in interstellar clouds, their presence supports the hypothesis that comets originate in such an environment.

a million tons of gas per second moving at a radial speed of 250 miles per second. This solar wind, which has a temperature of a million degrees, drags with it chaotic magnetic fields. The fields are carried by currents of electrons in the gas, which is almost completely ionized. Nearly a decade before the first space probe Ludwig F. Biermann of the Max Planck Institute for Physics in Göttingen demonstrated that something like the solar wind was needed to account for the ion tails of comets [see "The Tails of Comets," by Ludwig F. Biermann and Rhea Lüst; SCIENTIFIC AMERICAN, October, 1958]. Although the solar-wind theory of ion tails is not yet very precise, it indicates that two processes couple the solar wind to the cometary gas.

First, the high-energy electrons in the solar wind ionize the molecules in the coma (along with the solar radiation). Second, the solar wind gives rise to a bow wave around the coma. The chaotic magnetic fields now act as a magnetic rake that selectively carries the ions away from the coma, leaving the un-ionized molecules and atoms unaffected. The force of the solar wind on the ions can accelerate them to velocities of several tens of miles per second, so that changes in an ion tail can be seen at distances of many millions of miles on a time scale as short as half an hour.

John C. Brandt of the Goddard Space Flight Center of the National Aeronautics and Space Administration has explained the beautiful curvature of these great tails. It results from the transverse motion of the comet at some tens of miles per second across the movement of the solar wind blowing radially from the sun. The ion tails interact with the high-velocity solar wind in the same way that the smoke rising from a smokestack interacts with moving air to produce a graceful billowy arch on the earth.

Cometary Debris

Comets strew debris behind them in interplanetary space. Some of it is seen from the earth as the zodiacal light, which is visible as a glow in the eastern sky before sunrise and in the western sky after sunset. (It is brightest in the Tropics.) Much of the zodiacal light near the plane of the earth's orbit is sunlight scattered by fine dust left behind by comets. Under ideal observing conditions cometary dust also appears as the Gegenschein, or counterglow: a faint luminous patch in the night sky in a direction opposite that of the sun. Comets need to contribute about 10 tons of dust per second to the inner solar system in order to maintain this level of illumination. Over a period of several thousand years the particles are gradually broken down by collisions with other particles, or are blown away by solar radiation.

In addition to such fine material, comets distribute larger solid particles along their orbit. We see these meteoroids as meteor showers when the earth passes near the orbit of a regular comet. The meteoroids enter the upper atmosphere at speeds as high as 40 miles per second, and atmospheric friction releases the kinetic energy of the object in a short-lived flash of light. The energy released per gram of the meteoroid's weight far exceeds the energy efficiency of the most powerful man-made explosives. Thus an object the size of a pea can create a substantial meteor trail.

No meteoroid associated with a com-

HALLEY'S COMET was photographed in May, 1910, from the southern observing station of the Lick Observatory on Cerro San Cristóbal in Santiago, Chile. Since it has an orbital period of 76 years, it is due for its next return in 1986. Tail of this comet has negligible amount of dust.

COMET AREND-ROLAND was one of two bright comets seen in 1957; the other comet was Comet Mrkos (*see illustration on opposite page*). Photograph was made with the 48-inch Schmidt telescope on Palomar Mountain on April 30.

etary orbit, however, has ever been known to reach the ground, so that it is a very real question whether or not comets contribute any of the meteorites we collect in our museums. In fact, studies of meteors with cometary orbits demonstrate that the cometary debris is extremely fragile. Even though the spectra of such objects indicate that their main constituents are iron, magnesium, silicon and other earthly elements, the material is loosely packed. The meteoroids have an average density of less than half the density of water and can easily be crushed between the fingers. On entering the atmosphere at high speed the material is ripped apart and vaporized. It seldom, if ever, lands in pieces larger than fine dust. Some of the dust, however, can be collected as it floats slowly down through the atmosphere.

Gravel Bank v. Dirty Snowball

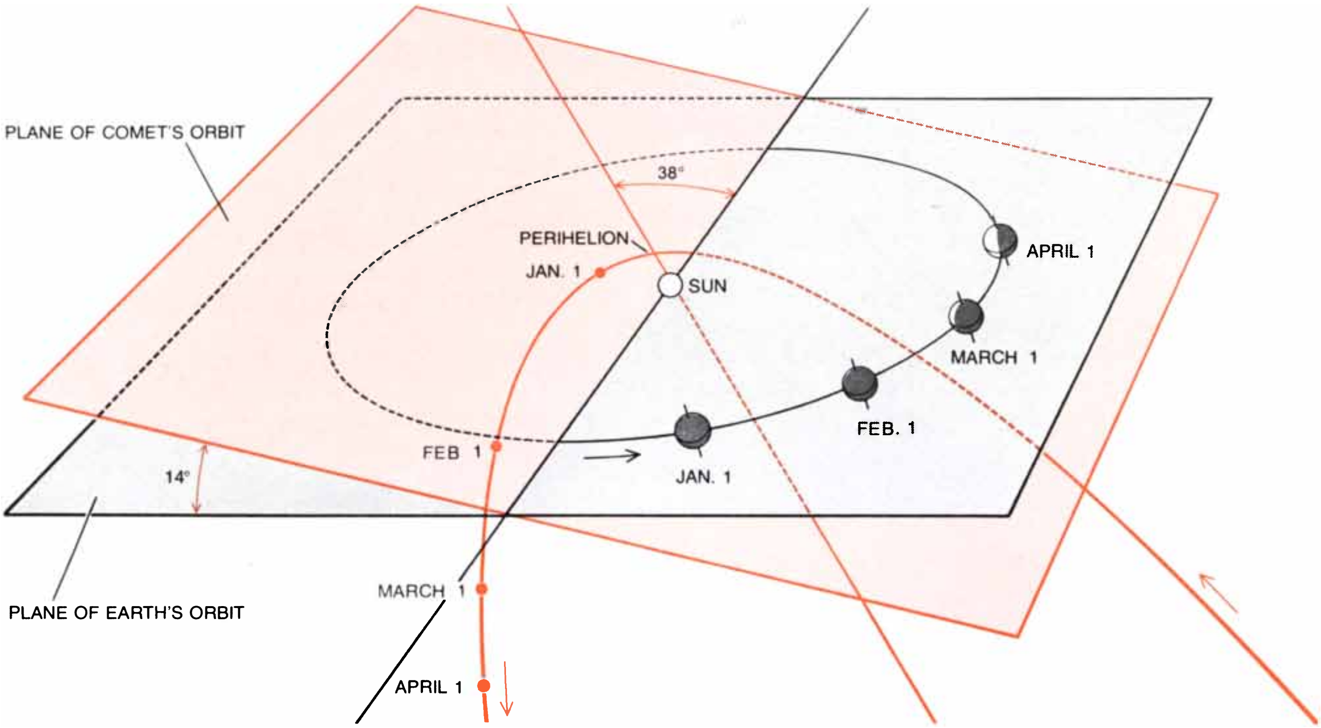
How can we be sure that the invisible nucleus of a comet is actually a dirty snowball? Until 1950, when I proposed the icy-nucleus model, the generally accepted conception was that the cometary nucleus was some kind of gravel bank flying through space. It was assumed that the cometary gases had been absorbed in the solid particles and were released by solar energy as the comet approached the sun. Since it was known that comets contribute to interplanetary space solid particles up to at least the size of a grape, it seemed reasonable to assume that such particles were the primary constituent of a comet and that the gases were of secondary importance.

The gravel-bank theory of the cometary nucleus, however, completely fails to pass three critical observational tests. Each of these tests is in itself fatal to the gravel-bank theory but is satisfied by the icy-nucleus hypothesis. The three tests tell us so much about the nature of comets that they merit some elaboration.

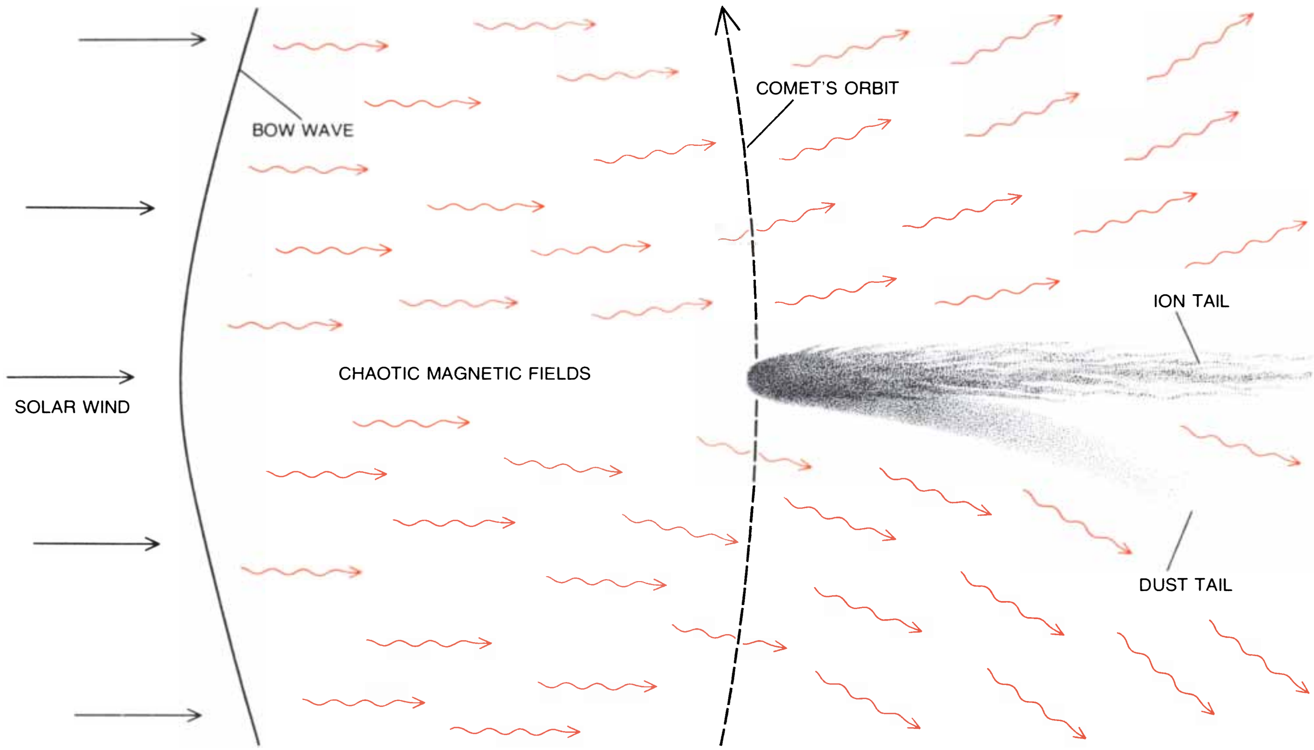
The first test involves the fact that comets that come remarkably close to the sun are not destroyed. In 1965 we had an opportunity to observe Comet Ikeya-Seki, which passed the sun at a distance only a third of the sun's diameter away from the sun's surface. At that distance the solar heat is so intense that it would melt and vaporize practically every common substance. There was plenty of time for earthly objects a foot or more in diameter to be boiled away as the comet passed through this solar inferno. A loose gravel bank would have been completely vaporized as the tidal forces due to the sun's gravity were tearing it apart. The intense solar wind would then have removed the resulting gas and completed the destruction of the comet. Comet Ikeya-Seki withstood the holocaust practically unscathed. Indeed, it is one of a group of eight such comets. They apparently had a common ancestor that broke up into separate comets, presumably because of the tidal forces due to the sun's gravity.

Sun-grazing comets pass so close to the sun as to be within the well-known limiting distance established by E. Roche of France in the middle of the 19th century. Roche showed that when a small fluid body passes a large body within 2½ times the radius of the large body, the gravitational force on the near side of the small body is much greater than it is on the far side. The force is enough to pull the body apart against its own simple gravitational cohesion. Only a body

COMET MRKOS, visible in August, 1957, clearly shows that a comet has two kinds of tail and also that the form of each tail changes with the passage of time. The straight tail to the upper left in both photographs that has the appearance of cigarette smoke caught in a breeze is composed of ionized gases. Fainter, smooth tail curving off below it is composed of dust. The photograph at left was taken on August 22, the one at right five days later. Both were made with the 48-inch Schmidt telescope on Palomar Mountain.



ORBIT OF COMET KOHOUTEK (*dark color*) is shown with respect to the orbit of the earth (*black*). Kohoutek's orbit is an extremely elongated ellipse whereas the earth's orbit is nearly a perfect circle. The plane of the comet's orbit (*light color*) is inclined at an angle of only 14 degrees to the plane of the earth's orbit (*gray*). At perihelion Comet Kohoutek was some 13 million miles from the sun. Dates along the two orbits indicate the position of the comet and the earth with respect to each other at various times.



ION AND DUST TAILS of a comet are created by two different processes. The ion tail is formed by a double process. First high-energy electrons in the solar wind (gases ejected from the sun at high speed) ionize the molecules in the coma of the comet, stripping them of electrons and leaving them positively charged. Second, the solar wind gives rise to a bow wave around the coma. Chaotic magnetic fields within the solar wind act as a magnetic rake that selectively carries the ionized molecules away from the coma at high speeds. Dust tail is formed by pressure of sunlight pushing the dust particles out of the coma. Tail is curved because the dust particles follow the laws of motion of orbiting bodies and lag behind the coma as they stream away at relatively low speeds.

COMA OF COMET BENNETT, a bright comet visible during the spring of 1970, shows a detailed spiral pattern in this photograph made by Minton and Stephen M. Larson on March 28, 1970. They made 10 photographs ranging in exposure time from four to eight seconds, using the 61-inch reflecting telescope of the Catalina Observatory. This picture is a composite of the 10 images. Minton and Larson interpret the spiral pattern as evidence that the icy nucleus is rotating on its axis with a period of about a day and a half as it orbits the sun.

with more than gravitational cohesion can withstand the tidal effects within Roche's limit. Even a body with such cohesion might split up along cracks or in other areas of weakness. We can therefore conclude that the nuclei of comets, at least comets of the sun-grazing family, have some internal strength but not always enough to protect them from the tidal disruption of a particularly close encounter with the sun. These facts alone eliminate the gravel-bank model of the cometary nucleus, unless one wishes to postulate that various comets differ completely in their basic structure.

The Second Observational Test

The second observational test fatal to the gravel-bank model is the persistence of some of the periodic comets. About 100 comets of the 600 so far observed move in orbits with periods of less than 200 years. All of those with periods of less than 30 years move around the sun in the same direction as the planets, and the planes of their orbits are inclined an average of 12 degrees to the plane of the earth's orbit. Some of these comets have completed many revolutions around the sun during recorded history. The painstaking records of the Chinese show that Halley's comet has appeared near the sun 29 times at intervals of between 76 and 77 years. The first well-observed passage was in 239 B.C. and the most recent was in 1910. In 1986 we shall see Halley's comet again.

From the observations made in space by the Orbiting Astronomical Observatory satellites (OAO) we know that bright comets lose tons of gas per second over periods of months as they come close to the sun. On the gravel-bank theory one would expect a single passage to remove a major fraction of the gas that the particles in the gravel bank have absorbed. The comet should then drastically diminish in brightness on subsequent returns. Halley's comet has probably lost a cubic kilometer of ice during its known history, perhaps even more. Interplanetary space contains so little matter that the gravel bank could not absorb enough gas to replenish its supply during the part of its lifetime spent at great distances from the sun. Thus a comet must have a huge reservoir of ice and solids to fuel its spectacular performance each time it passes close to the sun. For Halley's comet an icy nucleus some five miles or more in diameter is clearly indicated. The comet need lose only a fraction of 1 percent of its total substance on each return.

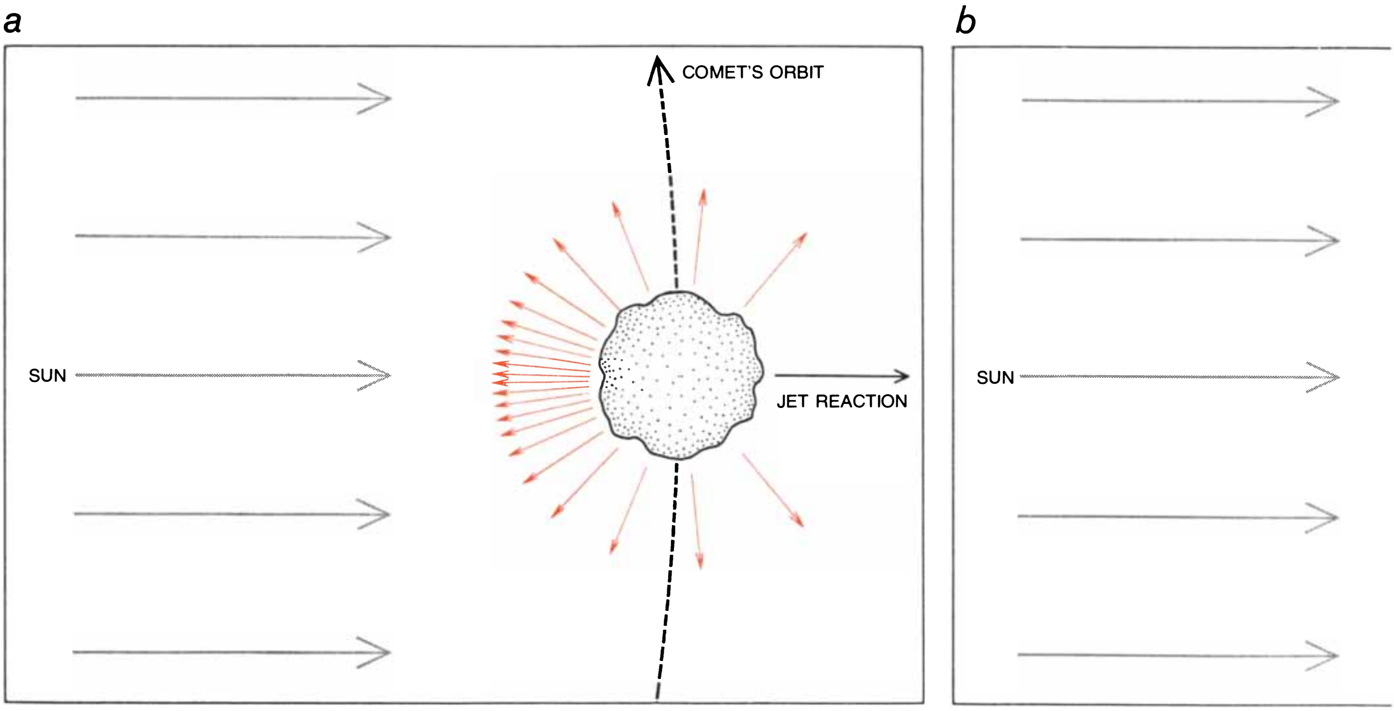
Now we come to the final test. A mysterious idiosyncrasy of comets is that they defy Newton's law of gravity. In 1819 J. F. Encke of Germany studied the motion of a short-period comet discovered in 1786. To his surprise he found that the comet, which has the shortest cometary period known (3.3 years), persisted in returning at each revolution about 2½ hours too soon. Only 2½ hours in 3.3 years may seem like a trivial deviation, but the accuracy with which astronomical positions can be measured places it well outside the possible errors of measurement and calculation.

Nongravitational Motion

Comet Encke still persists in returning sooner than predicted on each revolution. Brian G. Marsden and Zdenek Sekanina of the Smithsonian Astrophysical Observatory find, however, that the effect has slowly diminished to about 10 percent of the deviation it exhibited 150 years ago. Even Halley's comet shows maverick tendencies in its motion. T. Kiang of the Dunsink Observatory in Dublin finds that during its past 11 apparitions the comet persists in arriving late by an average of 4.1 days. Marsden and Sekanina show that of 20 short-period comets all but two show deviations from Newtonian motion, half being accelerated and half retarded.

The gravel-bank model completely fails to account for this nongravitational motion among comets. Until space probes had proved that interplanetary space was a nearly perfect vacuum, one might have assumed that there was a drag caused by some resisting medium. We now know that there is not enough matter in space to produce the required retarding effect. Moreover, even if there were such a drag, there would be no way for it to push certain comets forward in their orbits.

The icy nucleus provides a simple explanation for the wayward motions of comets. In the glare of the solar radiation the ices evaporate and send out molecules at speeds of several hundred feet per second. The molecules are ejected on the sunny side of the nucleus and



COMETS DEFY LAW OF GRAVITY by speeding up and slowing down in their orbits. The radiation from the sun causes the ices to sublime (evaporate) and send out molecules at speeds of several hundred feet per second. The molecules (*color*) are ejected on the sunny side of the nucleus and generate a jet reaction. This force pushes the comet away from the sun (*a*). If the nucleus is rotating, the jet will be displaced along the direction of the comet's orbital motion. If the nucleus is rotating in the same direction as its motion

generate a jet reaction, a real force pushing the comet away from the sun. If such a force is to change the period of a comet, however, it must be directed along the path of a comet, either fore or aft. Now, suppose the comet nucleus is rotating. (And we have yet to observe a celestial object that does not rotate to some extent.) The jet will be displaced by the rotation, that is, there will be a delay in the ejection of gas from any point on the surface as it rotates past the point directly facing the sun.

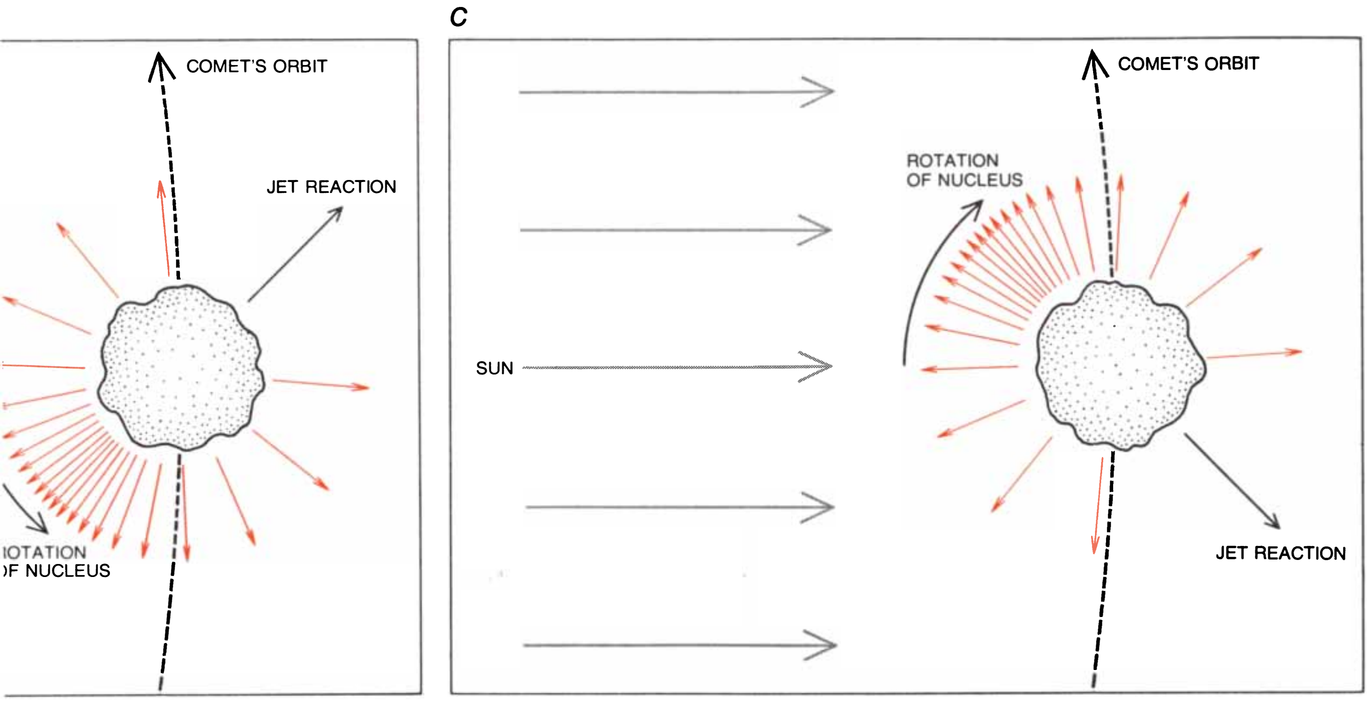
The delay is similar to the lag of the seasons on the earth: the summer temperature in the Northern Hemisphere is highest late in July, not at the time of maximum solar heating on June 21. If a comet is rotating in the same direction as its motion around the sun, the delayed jet action will have a forward component. The comet will drift outward from the sun in its orbit, and its orbital period will be increased. As a result it will show up later than predicted.

On the other hand, if a comet rotates in a direction opposite the direction of its motion around the sun, the jet action will have a backward component. The comet will "feel" a drag force, and it will drift slowly toward the sun in its orbit. Its period will be reduced, and it will show up earlier than predicted. Calculations I made in 1950 demonstrate that for a comet nucleus that has a diameter of the order of a very few miles, and that utilizes solar energy fairly efficiently in subliming its ices, the calculated force is adequate to produce the changes observed in the arrival times of the periodic comets. If the comets' axes of rotation are randomly distributed, we should expect about half of them to be accelerated and half retarded. That, of course, is the case.

The Water Component

I later found that, if the icy-nucleus model is to account for the changes in the periods of comets, it would have to eject gas at a rate that is considerably higher than what can be deduced from the intensity of the lines in the spectrum of a comet at the wavelengths of visible light. Otherwise the jet force would not be strong enough to exert an adequate force on the larger nuclei. Then Biermann showed that if ordinary frozen water is the major component of icy nuclei, bright comets should be surrounded by huge clouds of hydrogen and the hydroxyl radical arising from dissociated

COMETS COME FROM A HUGE CLOUD of isolated icy nuclei, according to the theory put forward by J. H. Oort. The Oort cloud is gravitationally a part of the solar system but extends out some 50,000 astronomical units in radius. (One astronomical unit is the distance from the earth to the sun.) The open circle at the right represents the diameter of the solar system. The small dots repre-



around the sun (*b*), the delayed jet reaction will push the comet forward in its orbit. The nucleus will drift outward from the sun, its orbital period will increase and it will show up later than predicted. If the nucleus rotates in a direction opposite to its orbital motion around the sun (*c*), jet reaction will have a component of motion opposite to comet's direction of travel. Comet will experience a drag force and will drift inward toward the sun. Its orbital period will be reduced and it will show up earlier than predicted.

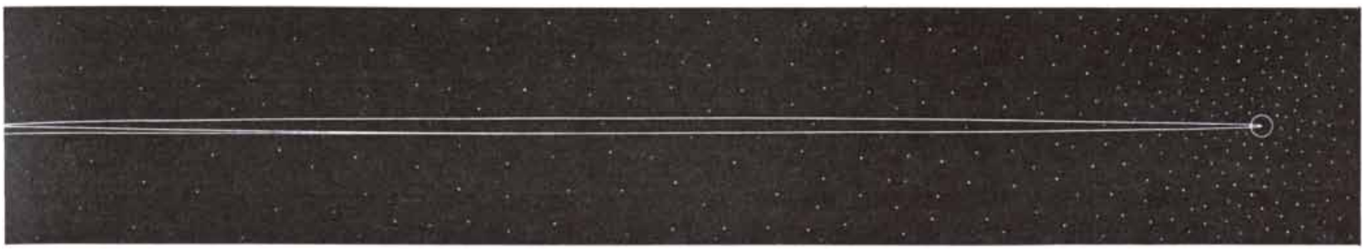
water molecules. A hydrogen atom isolated in deep space sends out almost all its fluorescent radiation not in the visible region of the spectrum but in the very far ultraviolet at the spectral line designated Lyman alpha.

In 1970 Arthur D. Code and Charles F. Lillie turned the first Orbiting Astronomical Observatory to observe Comet Bennett. They discovered a large cloud of hydrogen radiating at the wavelength of the Lyman-alpha line. They also found that the hydroxyl radiation in the far violet, which is barely detectable on the earth, is much stronger than had been suspected. The gas that was radiating at these wavelengths was what was missing from my calculations. Moreover, there seem to be approximately equal numbers of hydrogen atoms and hydroxyl radicals, confirming the suspicion that frozen water is a major constituent of a comet nucleus. Thus the observations verified Biermann's hypothesis and the basic conclusions drawn from the icy-nucleus model. The observations were confirmed by J. Blamont and J. Bertaux of the National Center of Space Studies in Paris; since then large quantities of hydrogen have been detected in three other comets, including Comet Kohoutek.

The Origin of Comets

Now that we know the nature of a comet nucleus, what might its detailed chemical and physical structure be? Could that structure disclose the origin, evolution, lifetime and final disposal of comets? It is easy to make comets, at least sitting at one's desk. All that is needed is to collapse and cool a cloud of interstellar gas and dust. There are many collapsed clouds in the central plane of our galaxy, where new stars and probably new solar systems are now evolving. One of these stellar incubators is the region of the Great Nebula of Orion.

When the temperature in a highly collapsed cloud falls low enough, say to a few tens of degrees above absolute zero, all earthly substances such as water, ammonia and methane will freeze into solid particles. The molecules observed in comets are simple combinations of carbon, nitrogen, oxygen and hydrogen. If comets originate in a typical interstellar cloud of our galaxy that has abundances of atoms much like the abundances in the sun, these four elements are by far the most prevalent ones capable of form-



sent the comet nuclei in this section of the Oort cloud. A star passing through the outer regions of the cloud (approximately 10 feet to the left off the page on the scale of the drawing) will disturb the motions of some of the nuclei just enough for them to enter the solar system. Ellipse shown is the correct scale for orbit calculated for Comet Kohoutek: 44 astronomical units wide and 3,600 long.

ing compounds. Although a quarter of the sun's mass is helium, helium and the other noble gases will not combine chemically with other elements, nor will they freeze out at the expected temperatures and densities.

Are the comets we observe actually created by this process? The solar system's comets form a huge cloud, gravitationally a part of the solar system but extending out several thousand times farther than the outermost planet (Pluto). As J. H. Oort of the University of Leiden showed in 1950, a star passing near the outer reaches of this cometary cloud will disturb the motions of some comets just enough for them to enter the inner part of the solar system where we can observe them. Otherwise they remain deep-frozen, perhaps since the sun and planets were formed 4.6 billion years ago. Comet Kohoutek appears to be one of the "new" comets entering the planetary region for the first time; thus it may be a sample of the primordial material of the solar system.

Interstellar Cloud v. Solar Cloud

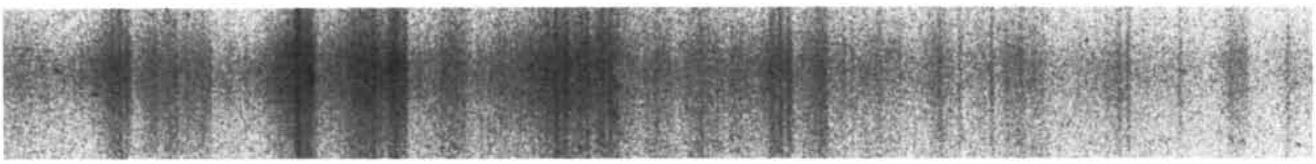
Where did the comets in the great Oort cloud originate? A. G. W. Cameron of the Harvard College Observatory has recently suggested that they formed from fragmentary interstellar clouds at the time when a larger cloud spun down into a disk and formed the solar system. The comets would then be created moving in huge orbits like those of new comets such as Comet Kohoutek before they are disturbed by a passing star. It is not yet certain whether or not the ices and solids needed to form sizable comets could aggregate in a fragmentary interstellar cloud. On the other hand, we are certain that comets must have been formed in huge quantities at the outer edges of the solar system.

The planets Uranus and Neptune have a mean density about twice the density of water, just what one would expect if they were aggregates of literally hundreds of billions of comets. The terrestrial planets (the earth, Mars, Venus and Mercury) must have been formed from large numbers of planetesimals, or small earthlike aggregates. The difference between Uranus and Neptune and the terrestrial planets appears to depend simply on the temperature in the two regions of space at the time they were formed. One would expect that in a great collapsed interstellar cloud rotating as a disk the temperature would remain higher near the burgeoning sun than in the outskirts of the disk. Apparently water could not freeze inside the orbit of Jupiter but could freeze outside the orbit of Saturn. (Jupiter and Saturn themselves, like the sun, are made mostly of hydrogen and helium, compressed by the great mass of the planets to roughly the density of water.) Therefore we can be fairly sure that comets formed beyond the orbit of Saturn. Icy grains and grains of earthlike material formed in the cooling disk of the solar cloud, accumulated into comets and finally aggregated into Uranus and Neptune. After Uranus and Neptune became massive the remaining comets were disturbed by the gravitational attraction both of the new planets and of Jupiter and Saturn, and they were swung into the huge orbits of the Oort cloud.

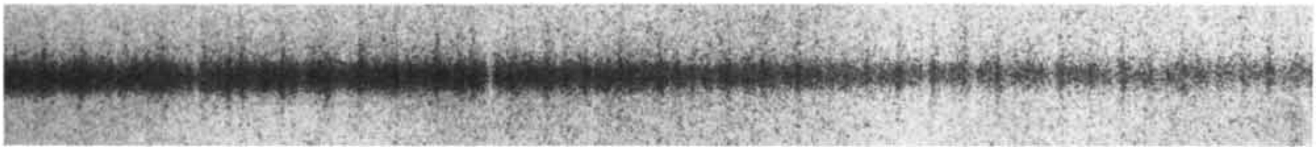
The process would have been extremely wasteful. Many comets would have been captured by Jupiter and Saturn. Many would have decayed into smaller orbits in the inner solar system as they do today. Many would have been lost into the depths of the galaxy. No one has yet shown, however, that comets formed within the orbits of the present outer planets could actually move outward into the great Oort cloud, nor, as we have seen, has it been proved that fragmentary interstellar clouds could condense and aggregate into comets.

Resolving the Dilemma

To determine whether comets formed in the cloud that gave rise to the solar system, in fragmentary interstellar clouds or in some other kind of environment, clues are needed both from new theories of comets and from new observations. Suppose comets did form in fragmentary interstellar clouds. Then each interstellar dust grain would be a nucleus for the condensation of the various ices. The ices might consist largely of the "exotic" molecules now being discovered by radio astronomers in clouds of interstellar gas and dust. These molecules include formaldimine (CH_2NH), methyl alcohol (CH_3OH), methyl cyanide (CH_3CN), hydrogen cyanide (HCN) and more than 20 others, mostly composed of elements that are abundant



SPECTRUM OF A COMET HAVING MUCH GAS in relation to dust shows many distinct spectral lines composing one molecular band. The spectrum, which is of Comet Ikeya (visible in 1963), spans the blue wavelengths from 3,900 angstroms (*left edge*) to 4,100 angstroms (*right edge*). Since this spectrum is a negative print, the dark lines are actually bright emission lines. Almost all the features are from the carbon molecule C_3. The featureless horizontal streak in the center is weak spectrum of reflected sunlight.



SPECTRUM OF A COMET HAVING MUCH DUST in relation to ionized gas shows few and weak distinct spectral bands and lines. The horizontal black streak of the continuous spectrum of reflected sunlight, however, is very strong. The spectrum, which is of Comet Mrkos, is also crossed by absorption lines that show up light on this negative print. These indicate the wavelengths at which the dust absorbs light from the solar spectrum. Spectrum runs from approximately 4,880 angstroms (*left edge*) to 5,100 angstroms (*right edge*). Both this spectrum and the one at the top are from Jesse L. Greenstein of the California Institute of Technology.

in comets: carbon, nitrogen, oxygen and hydrogen.

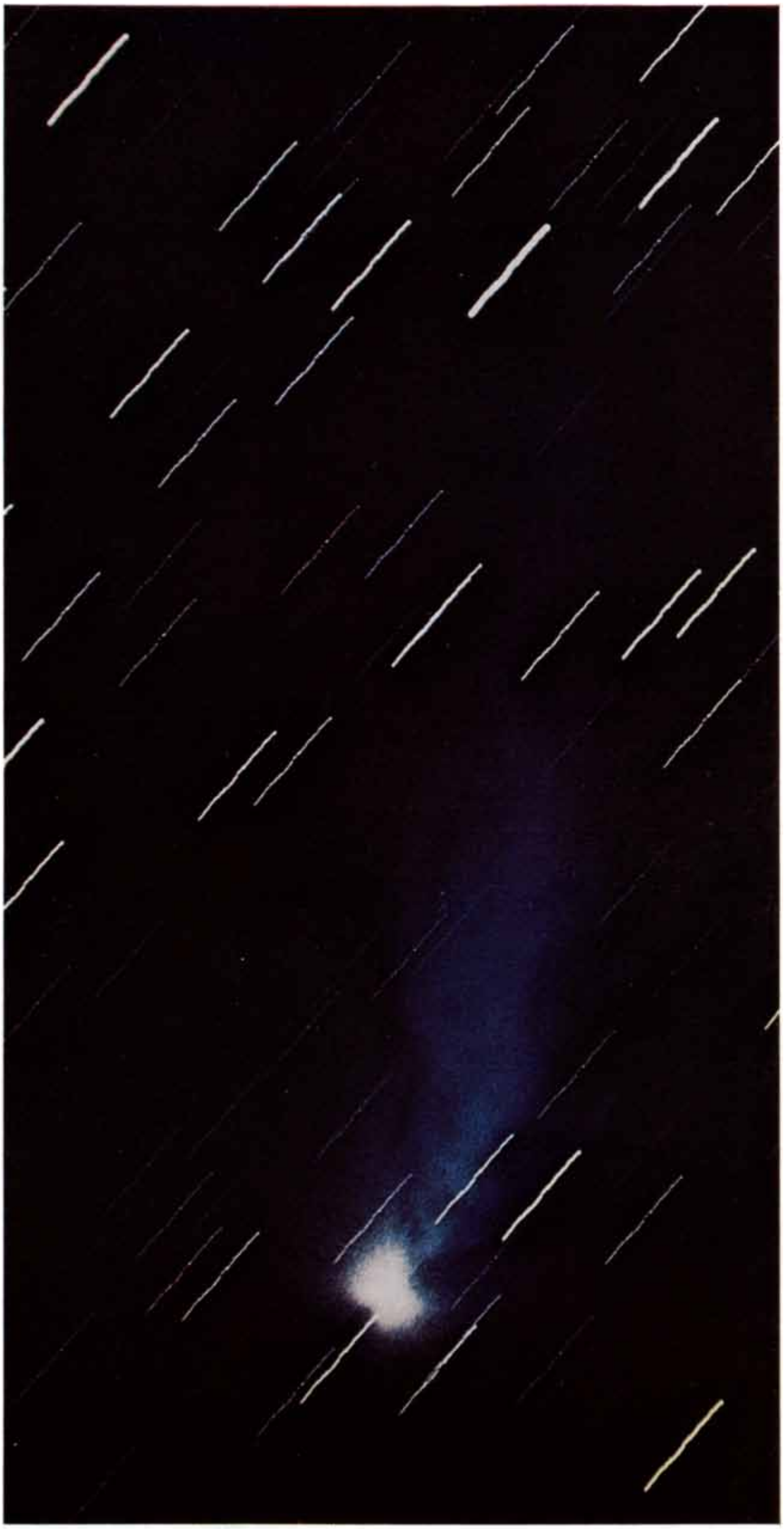
Are some of these exotic molecules the parent molecules of the simple radicals we see in comets? If they are, then comets probably formed in fragmentary interstellar clouds. If, on the other hand, comets originated within the orbits of the outer planets, one would expect the parent molecules to be the more stable ones such as ammonia and methane. As Pol Swings and Armand Delsemme of the Institut d'Astrophysique in Belgium have suggested, even though the temperature in the solar cloud may never have been low enough for methane to freeze out (some 20 degrees above absolute zero), such molecules were probably trapped in water snow.

The two possible birthplaces of comets might also give rise to marked differences in the structure and behavior of comets. Solar-cloud comets made of dust grains surrounded by ices could be expected to disintegrate to nothing as their substance sublimes. On the other hand, comets formed in the outer solar cloud might well have developed an earthlike core if the temperature of the cloud's great disk first rose during its collapse and then slowly fell. An old comet of this type might lose its outer shell of ices, slowly become inactive and finally appear starlike instead of fuzzy in our telescopes, looking very much like an asteroid. Indeed, many astronomers believe the small asteroids observed to cross the earth's orbit are really old comet nuclei.

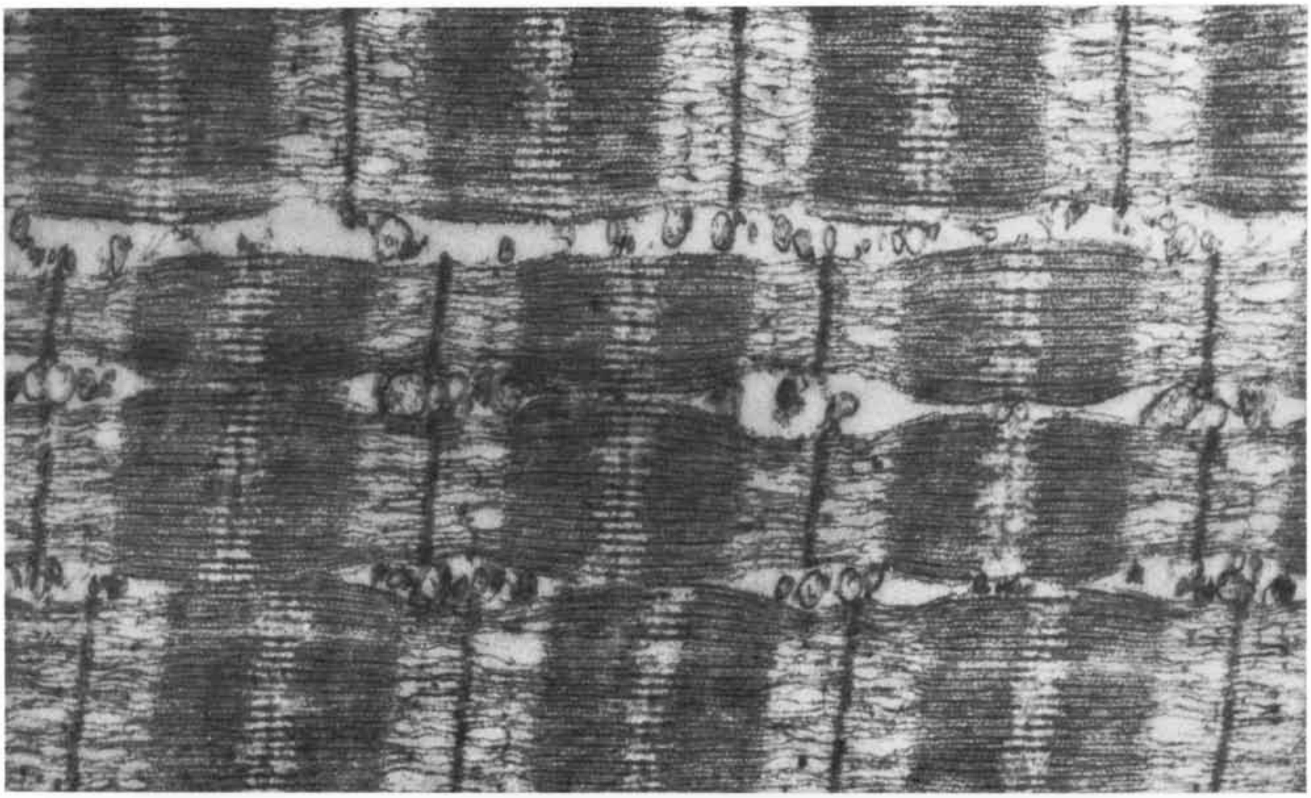
A Space Probe to a Comet

The best way to answer these questions about comets, and thereby questions basic to an understanding of the origin of the solar system, is to send space probes to the nucleus of a comet and to the surface of an asteroid. Such space probes, which would actually be unmanned observatories, are completely feasible today, both technologically and scientifically. They would be less expensive than several of the planetary probes that have already proved to be so successful.

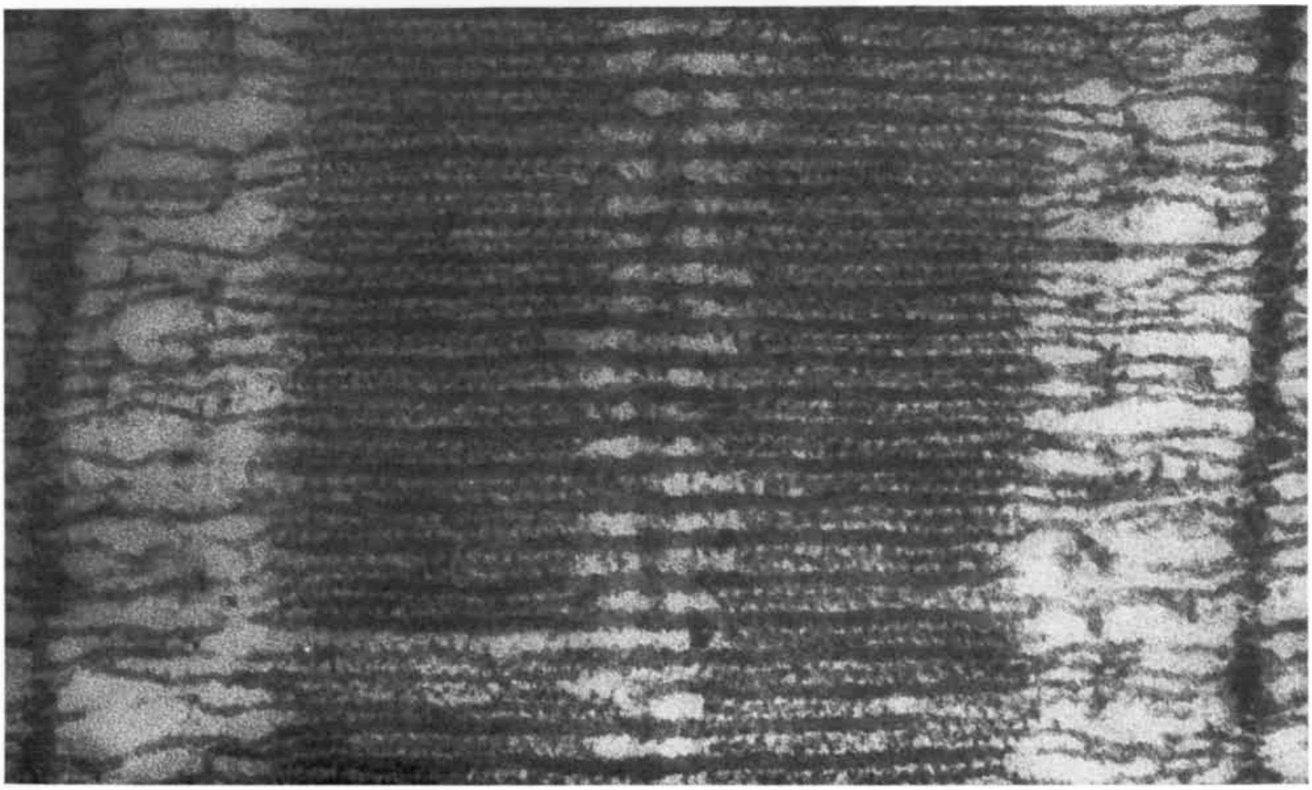
NASA is currently studying the possibilities of such space missions to comets and asteroids. The earliest mission now being considered would be to Comet Encke in 1980. Meanwhile the interest in comets renewed by Comet Kohoutek should result in new information of prime importance for our understanding the true nature of the most primitive bodies in the solar system.



COMET HUMASON, a relatively faint comet visible in 1960, was unique in that its spectra showed almost no dust at its great perihelion distance of some four astronomical units from the sun. The blue color is due to emission from ionized molecules such as carbon monoxide (CO^+), molecular nitrogen (N_2^+), the hydroxyl radical (OH^+), cyanogen (CN) and CH^+. Since the telescope (the 48-inch Schmidt telescope on Palomar Mountain) was following the comet's motion, the images of the background stars appear as streaks.



MUSCLE FIBER from vertebrate skeletal muscle is enlarged about 25,000 diameters in an electron micrograph made by H. E. Huxley of the Medical Research Council Laboratory of Molecular Biology in Cambridge, England. The specimen is a very thin longitudinal section. The fiber is an assembly of thin filaments (the finest horizontal elements), which extend in two directions from disklike protein structures called *Z* lines (*dark vertical lines*). Between the thin filaments, centered between *Z* lines, are the thick filaments.



FILAMENTS OF MUSCLE FIBER are further enlarged, to about 75,000 diameters. The thick filaments are seen to be studded with small projections, the cross bridges. By making contact with the thin filaments and swiveling, the cross bridges cause relative motion between thick and thin filaments and thus bring about muscle contraction, as shown in the illustration on the opposite page.

The Cooperative Action of Muscle Proteins

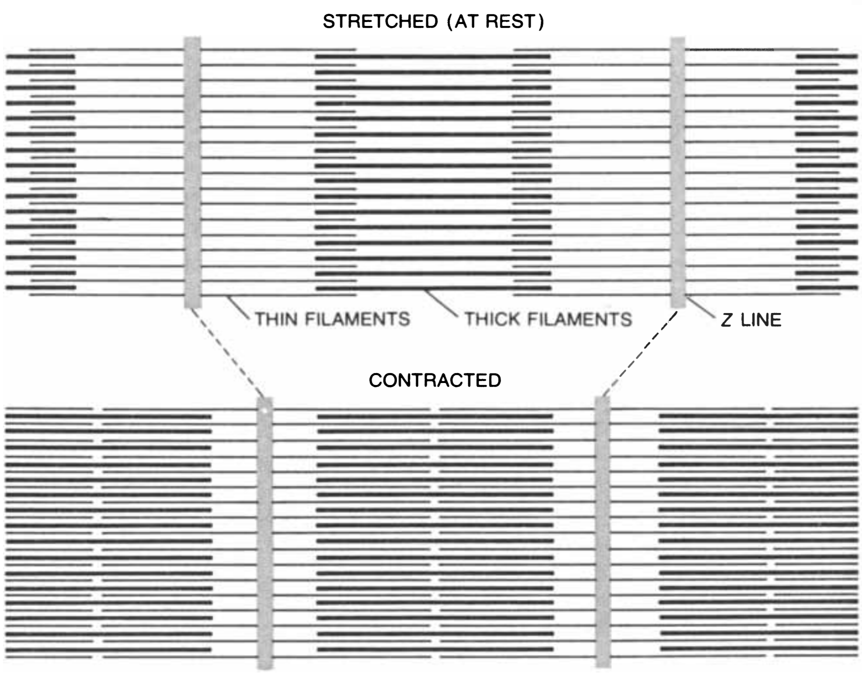
Muscle fibers are composed of four major proteins: myosin, actin, tropomyosin and troponin. The contraction of muscle is a function of their interactions, controlled by the presence of calcium ions

by John M. Murray and Annemarie Weber

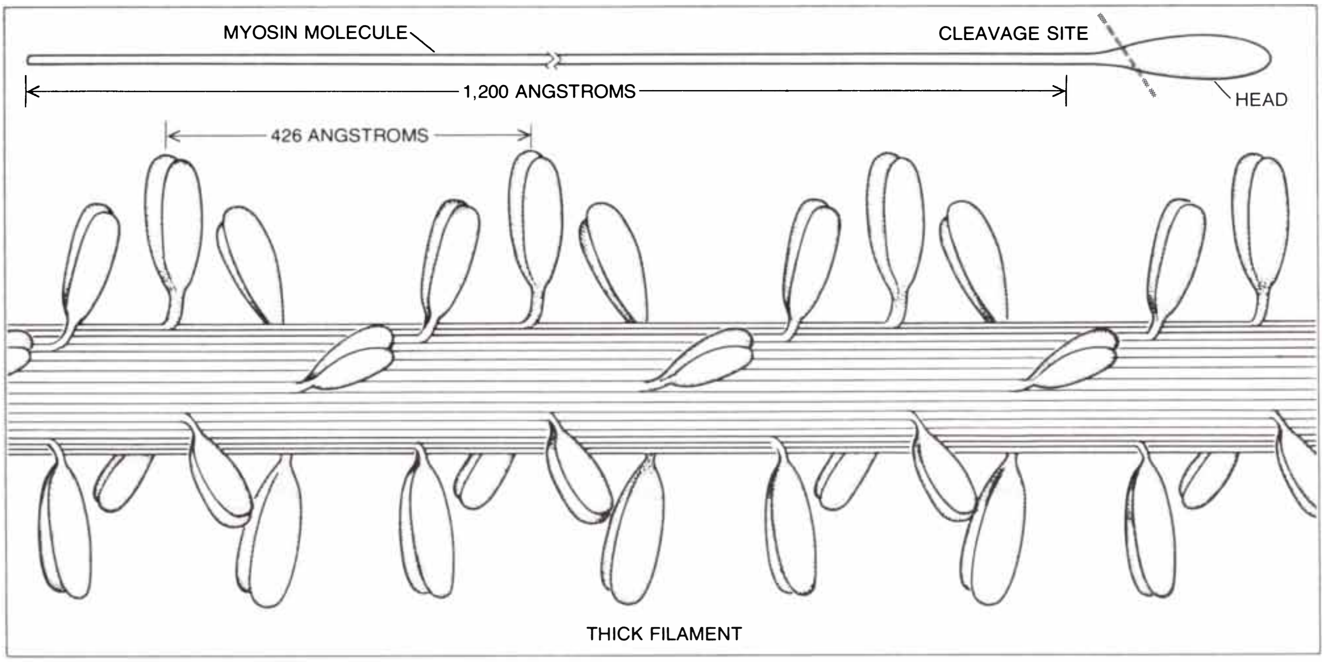
Each of the many activities of the living cell in large measure reflects the behavior of specific proteins in the cell. A striking illustration of this important generalization in biochemistry is provided by the contraction of muscle cells, which can be explained primarily in terms of the properties of just four proteins: myosin, actin, tropomyosin and troponin. Each of these proteins can be isolated from muscle in pure form and studied separately, but much of their behavior is a result of the highly ordered structures they form together within the muscle cell and of the cooperative biochemical interactions to which those structures lend themselves.

Major advances in the knowledge of the proteins and the structures they form have come from the work of Albert Szent-Györgyi in the 1940's and more recently from the laboratories of H. E. Huxley at the Medical Research Council Laboratory of Molecular Biology in Cambridge, England, the late Jean Hanson at Kings College London and Setsuro Ebashi at the University of Tokyo. In the muscle cell the four proteins are assembled into two multimolecular aggregates, the thick and thin filaments, which in turn are the basic components of the muscle fiber. The filaments can be prepared in the laboratory either by gently disrupting a muscle fiber into its components or by allowing the purified proteins to associate and reform the filaments. One way to understand the contribution each protein makes to the final structure is to perform the assembly of a muscle fiber verbally, as it were, by describing the individual molecules and how they fit together.

The thicker of the two kinds of filament observed in electron micrographs of muscle contains all the myosin of muscle, together with much smaller amounts of other proteins whose function is poorly understood and that we will neglect here. A myosin molecule resembles a thin rod with two small globular "heads" at one end. (Myosin can be decapitated with an enzyme that breaks the chemical bonds joining the heads to the rest of the molecule. This is the basis of a method developed by Susan Lowey at Brandeis University for preparing isolated heads. The isolated heads retain the chemical activity of intact myosin but are much easier to handle in experiments.) Within the thick filament the myosin molecules are arranged in a sheaf, forming a cigar-shaped structure that is studded with projections along its entire length except for a bare zone in the middle. In electron micrographs of muscle the projections appear as minute cross bridges that seem to link the thick and thin filaments. Apparently each projection is formed from the head region of one myosin molecule; the molecules are arranged within the thick filaments with their heads toward the two ends, which accounts for the bare zone in the middle. A thick filament of a normal muscle is



STRIATED PATTERN of muscle fiber is explained by the arrangement of thick and thin filaments. The diagram shows one sarcomere (the unit of muscle between two *Z* lines) and parts of the two adjacent ones, at a scale about the same as the top micrograph on the opposite page. Cross bridges swivel in opposite directions on each side of a central bare region, imparting the relative motion that causes a resting muscle (*top*) to contract (*bottom*).

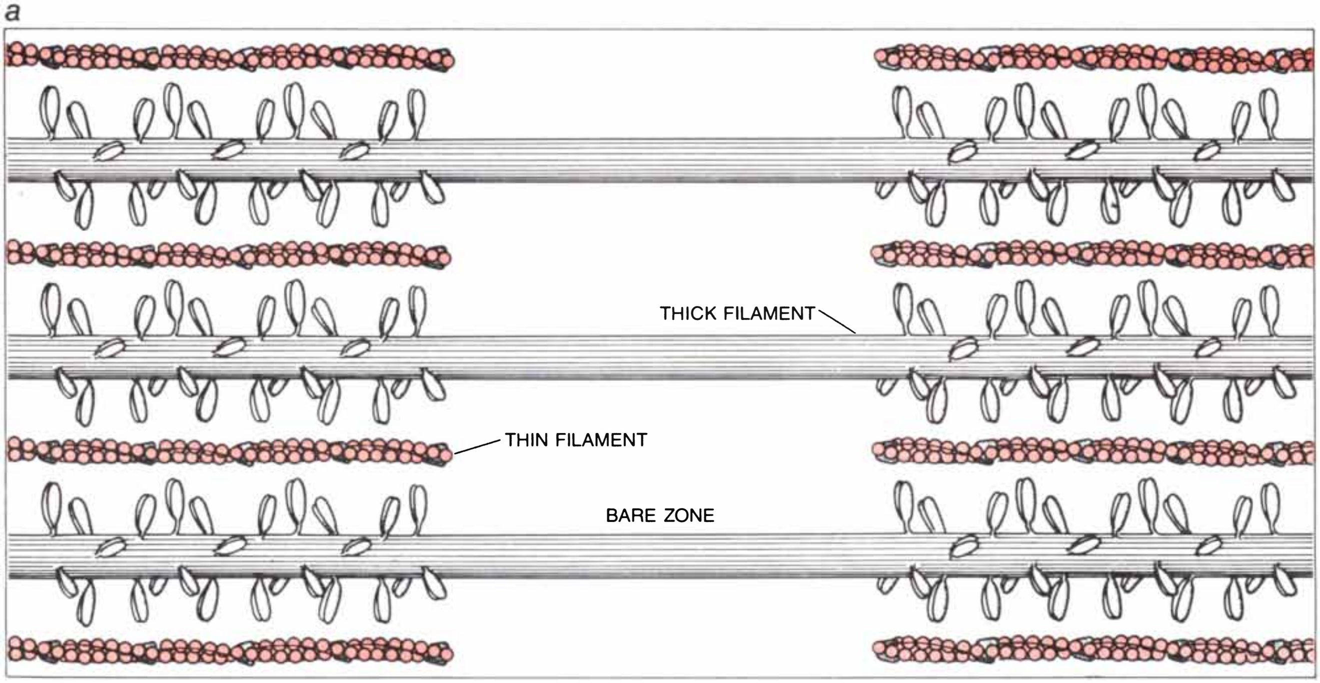


THICK FILAMENT is an assembly of myosin molecules, long rods with a double "head" at one end. The head has an active site where the chemical events involved in muscle contraction take place; purified heads are prepared for experiments by cleaving the molecule with an enzyme. In thick filaments myosins are bundled into a sheaf about 1.5 microns long with heads projecting in groups of three. Filament drawing includes pertinent aspects of a new model developed by John M. Squire of Imperial College London.

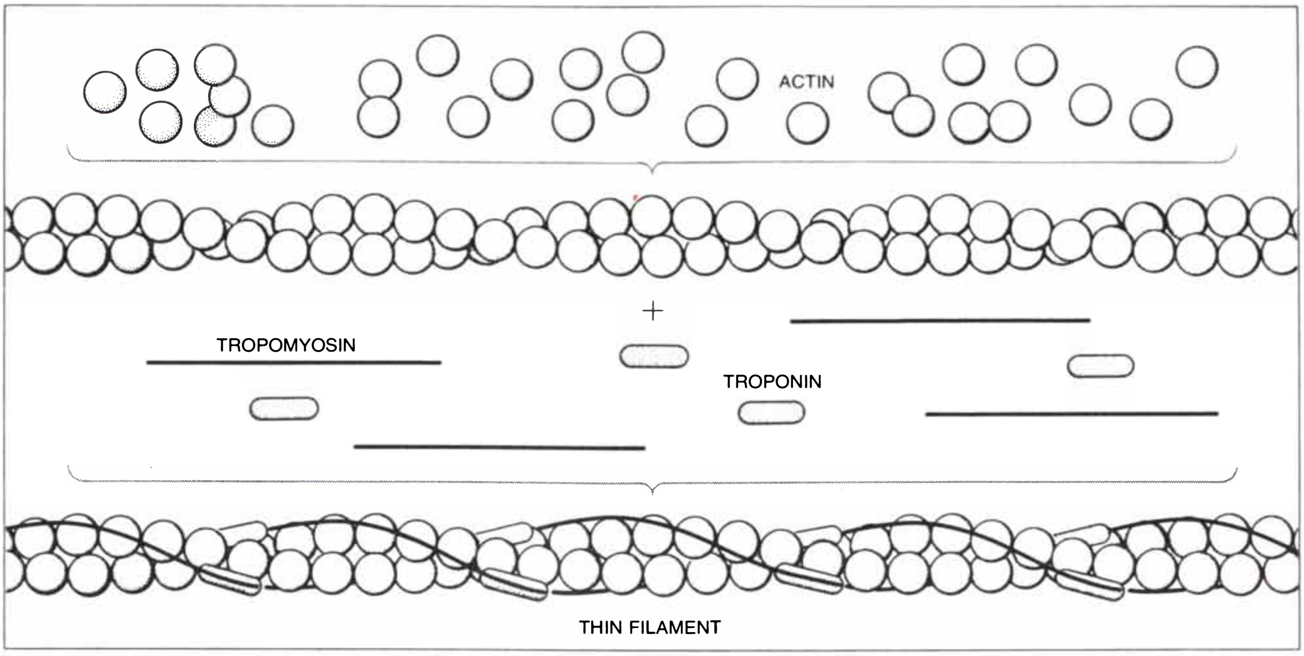
approximately 1.5 microns (thousandths of a millimeter) long and contains several hundred myosin molecules.

The other three major proteins involved in muscle contraction are incorporated in the thin filaments. Of the three actin is present in the largest amount. The actin molecules are small, roughly spherical particles that are arrayed in the thin filament as if to form a twisted double strand of beads. An important characteristic of the individual actin molecules is that they are not spherically symmetrical; each acts as if it had a distinguishable "front" and "back." Since the actin molecules are assembled into filaments in a "front to back" manner, the entire thin filament acquires a recognizable polarity, that is, the front sides of all the actins face in one direction and the back sides face in

a



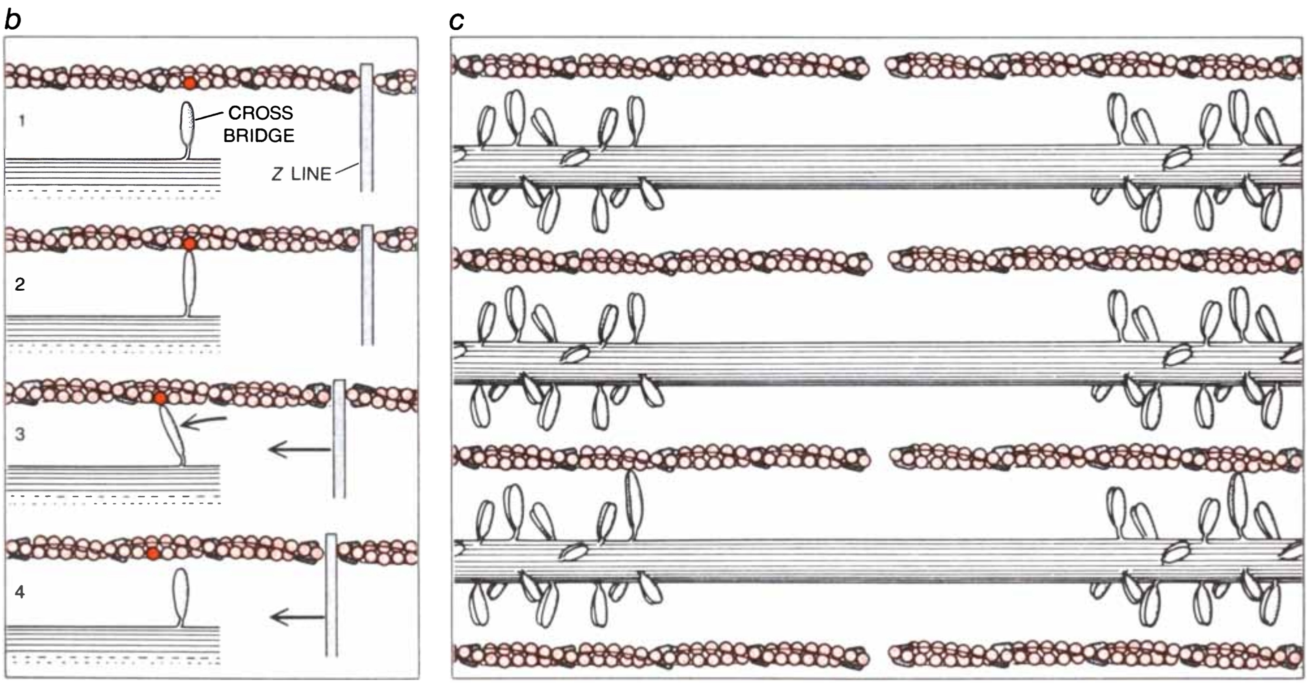
THICK AND THIN FILAMENTS interdigitate in an orderly array to form a muscle fiber. Two sets of thin filaments extend toward each other from adjacent *Z* lines. They lie between, and partly overlap, a set of thick filaments (*a*). The combination accounts for the striated appearance of muscle, as shown in the illustration on preceding page. In this arrangement myosin heads of thick fila-



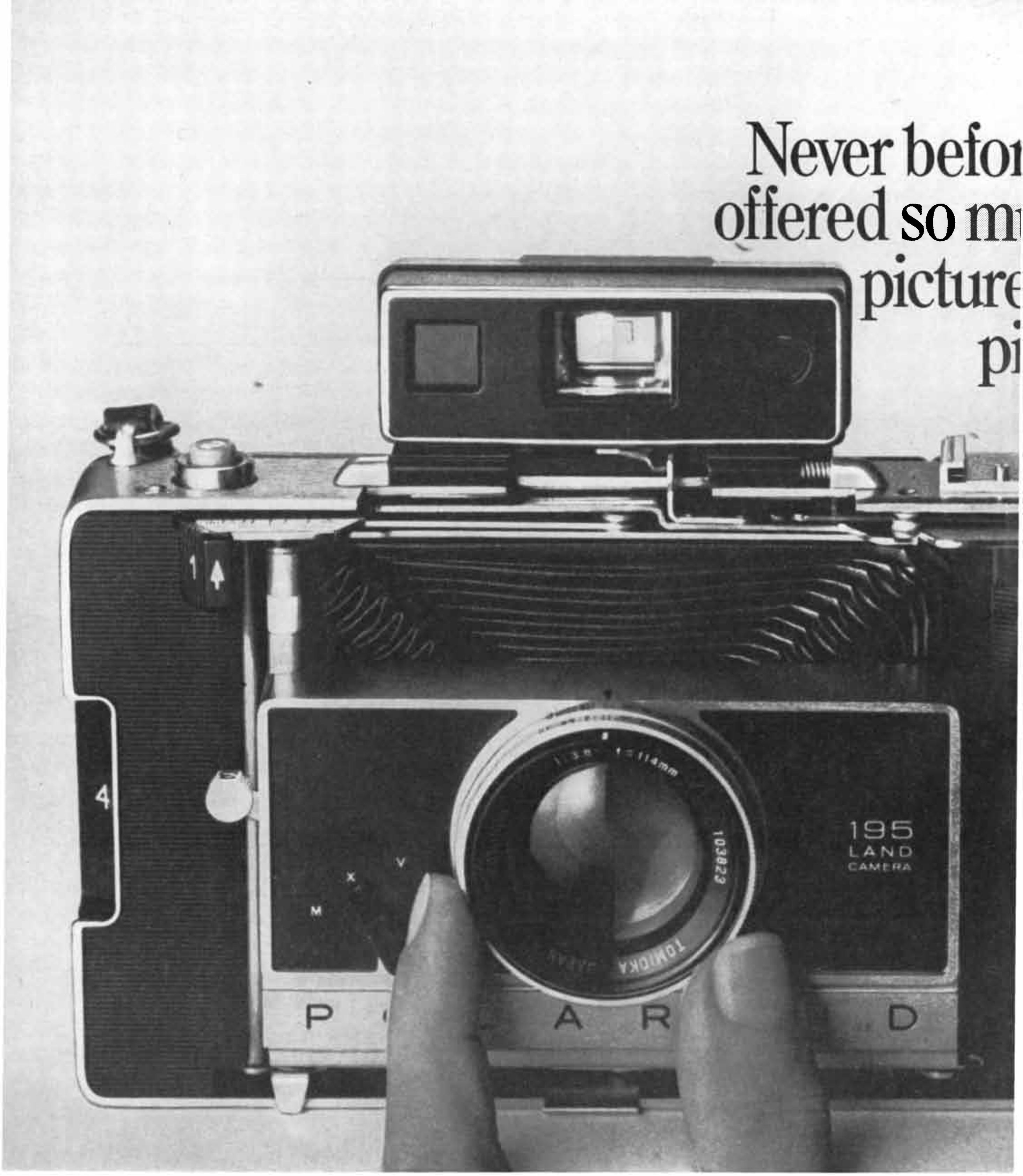
THIN FILAMENT is an assembly of actin, tropomyosin and troponin molecules. The actins, present in the largest amount, are small spheroidal molecules that are linked to form a double helix. Tropomyosin, a long, thin molecule, forms a continuous strand that sits on the string of actins alongside each groove of the double helix. A globular troponin molecule is affixed near one end of each tropomyosin. One tropomyosin extends over seven actin molecules and there are from 300 to 400 actins in the micron-long filament.

the opposite direction. This directionality is essential to muscular contraction.

In addition to actin the thin filaments in muscle contain the proteins tropomyosin and troponin. Tropomyosins are long, thin molecules that attach end to end, forming a very thin filament on the surface of an actin strand. Each strand carries its own filament, which lies near the groove between the paired strands. Troponin has a more or less globular shape and sits astride the tropomyosin molecule, probably a short distance from one end. In the thin filament one tropomyosin molecule extends over seven actins, with one troponin molecule on each tropomyosin. Thin filaments of skeletal muscle are commonly one micron in length and contain from 300 to 400 actin molecules and from perhaps 40 to 60 tropomyosins. In intact muscle one end



ments act as cross bridges that make contact with actin molecules in the thin filaments. The heads are oriented in opposite directions on each side of the central bare zone. The muscle contracts when a nerve signal initiates a sequence of events that causes myosin heads to attach to actins, swivel and then break contact (*b*), thus propelling the thin filaments past the thick and shortening the muscle (*c*).



You can set our new instant picture camera at any of 170 shutter speed/lens opening combinations to get a photograph with maximum depth-of-field, or to get the precise exposure effect you want.

And you can load our new camera with our new instant 3¼x4¼-inch black-and-white film to get a positive. *And* a negative that can produce an infinite variety of proportion and tone in copies or enlargements.

For serious picture-takers, the Polaroid Model 195 Land camera provides:

Ten shutter speeds from 1/500 to 1 second with additional settings for time exposure, flash synchronization and 10-second exposure delay.

Nine lens openings from f/3.8 to f/64 with clicks at each f-stop, and halfway in between.

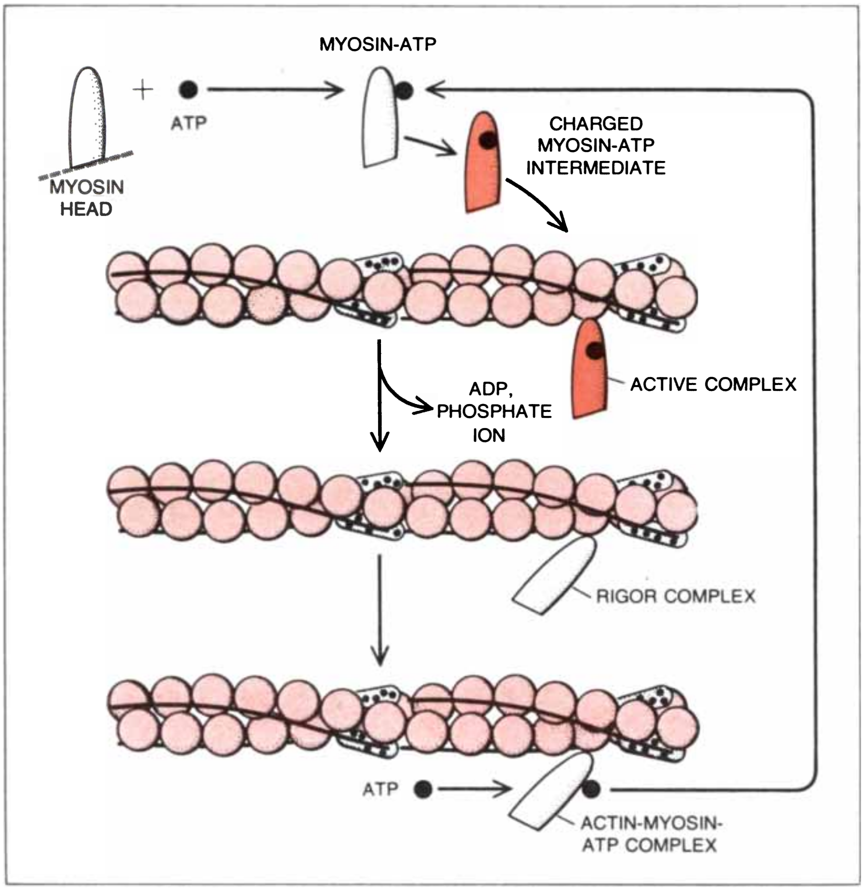
A 114mm, 4-element, coated, color-corrected Tominon lens (similar in design to a Tessar lens).

A Seiko between-the-lens shutter that gives the camera extraordinary exposure capabilities.

A superimposed-image rangefinder for quick, accurate focusing.

And a built-in timer that tells you when your picture is developed.

The Model 195 will make beautiful 3¼ x 4¼-inch



CHEMICAL EVENTS of the muscle-contraction cycle are outlined as they take place in the soluble experimental system described by the authors. A myosin head combines with a molecule of adenosine triphosphate (ATP). The myosin-ATP is somehow raised to a "charged" intermediate form that binds to an actin molecule of the thin filament. The combination, the "active complex," undergoes hydrolysis: the ATP splits into adenosine diphosphate (ADP) and inorganic phosphate and energy is released (which in intact muscle powers contraction). The resulting "rigor complex" persists until a new ATP molecule binds to the myosin head; the myosin-ATP is recycled, recharged and once again undergoes hydrolysis.

of the thin filaments is attached to the Z line, a flat protein structure. The thin filaments, which extend perpendicularly from both of its surfaces, have opposite polarity on opposite sides of the Z line. In a resting muscle the successive Z lines are separated by about 2.2 microns, and the thin filaments are aligned parallel to the long axis of the muscle fiber.

The last step in our mental construction of a muscle involves interdigitating a set of thick filaments in the spaces between the thin filaments projecting from adjacent Z lines. The two types of filaments are arranged in parallel arrays and overlap for part of their length [*see bottom illustration on pages 60 and 61*]. The spacing of the filaments is such that the two sets can slide past each other without restriction. It is this sliding motion that is responsible for the shortening of muscle contraction [see "The Mechanism of Muscular Contraction," by H. E. Huxley; SCIENTIFIC AMERICAN, December, 1965]. The driving force for the sliding comes from the cross bridges on the thick filament; they attach to the thin filament at a certain angle and then presumably swivel to a different angle, pulling the thin filaments past the thick. Cross bridges on opposite sides of the bare zone swivel in opposite directions, as if they were oars pulled by rowers facing in opposite directions at the two ends of a racing shell. The result of this rowing motion is to pull the filaments into greater overlap, thus decreasing the distance between Z lines and shortening the muscle. In order for significant shortening to occur, each cross bridge must act in a cycle: it must attach, swivel, detach and then reattach at a point farther along the thin filament. A single cycle of attachment, swiveling and detachment causes a relative movement of the two filaments of about 100 angstroms. This means that if all the cross bridges of a muscle fiber went through the cycle just once, the muscle would shorten by about 1 percent of its length.

The energy for muscle contraction, as for most other work in the living cell, is provided by the hydrolysis of adenosine triphosphate (ATP). This breakdown reaction splits the high-energy ATP into lower-energy products: adenosine diphosphate (ADP) and inorganic phosphate. The difference in the energy content of these compounds is available to do useful work, such as forcing a cross bridge to swivel, thus pulling a thin filament along past a thick one and shortening a muscle.

For muscle to contract something else must be present in addition to ATP and the four proteins. The additional requirement is calcium ion and, as in many other cellular processes, the calcium provides a means of regulation. Since the reactions of muscle contraction (the cyclic attachment and detachment of cross bridges powered by the hydrolysis of ATP) do not proceed in the absence of calcium, one can turn these reactions on or off at will in the laboratory merely by supplying and removing calcium ion. The contraction of muscle is regulated in just this way in the living organism [see "The Sarcoplasmic Reticulum," by Keith R. Porter and Clara Franzini-Armstrong; SCIENTIFIC AMERICAN, March, 1965]. When the nerve signal to initiate contraction arrives at the muscle cell, it causes a release of calcium into the fluid surrounding the filaments from special storage vesicles in the sarcoplasmic reticulum (an intracellular system of tubules formed from membranes), allowing the cross bridges to attach and contraction to occur. The calcium is quickly removed and returned to the storage vesicles by a calcium "pump" situated in the membranes of the sarcoplasmic reticulum. The removal of calcium prevents the further cycling of cross bridges and the muscle relaxes. Both the release of calcium and its subsequent removal are very fast processes, normally requiring only a fraction of a second.

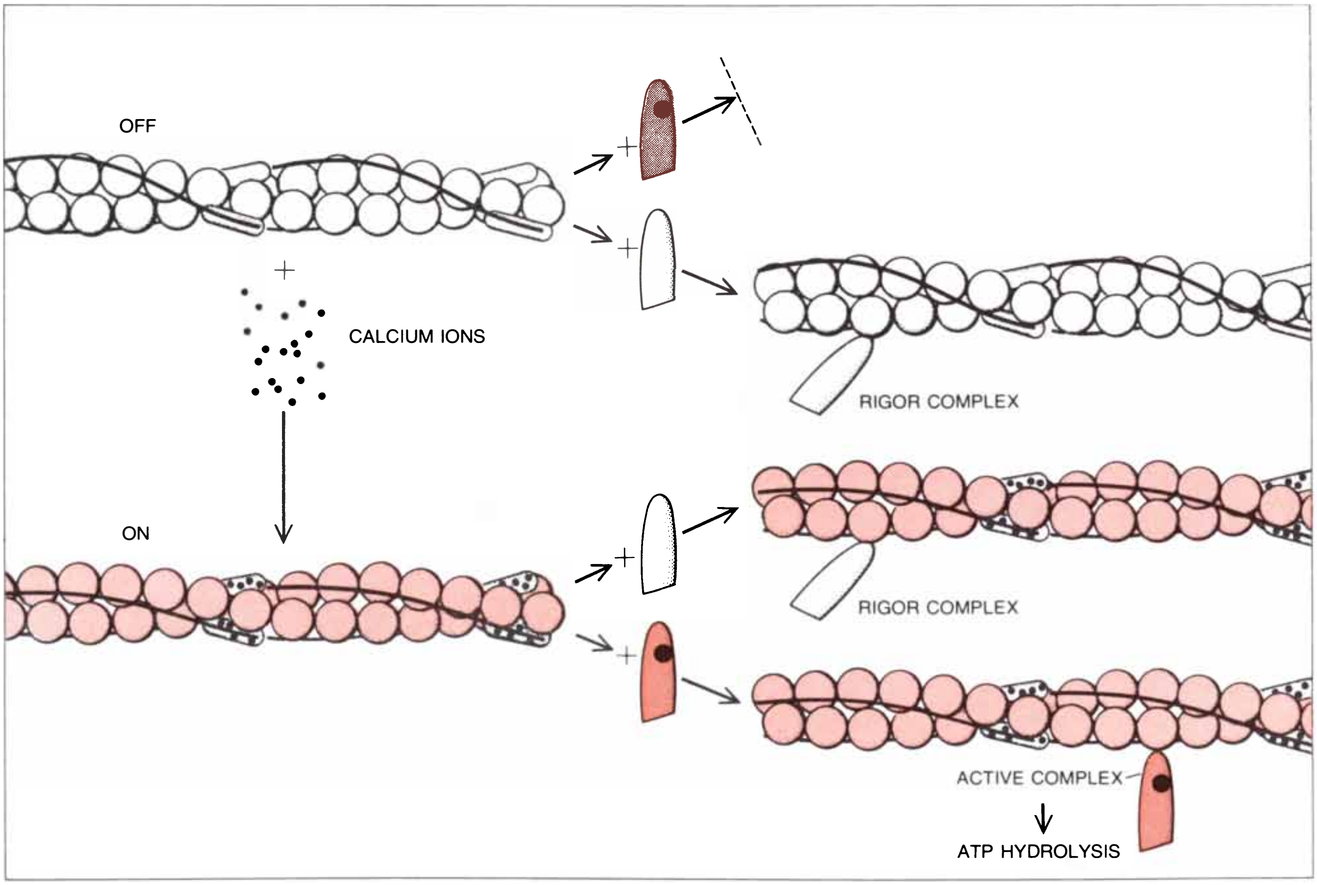
Many biochemical techniques are admirably suited to the study of proteins in solutions of randomly distributed individual molecules, which can be easily prepared from cells of tissues such as kidney or liver. The techniques are less suitable for studying structured aggregates of protein such as actin and myosin filaments. Even though the proteins of muscle can be isolated in pure form, the shortening of a muscle depends on a very special three-dimensional arrangement of these proteins; when they are isolated in solution and dispersed as single molecules for more convenient study, the cardinal manifestations of their activity—shortening and the generation of tension—are irreversibly lost.

As a result neither the complete retention of structure nor the complete randomization of structure yields a system that is easy to study. The method we adopted with Robert D. Bremel at St. Louis University Medical School and now apply at the University of Pennsylvania School of Medicine is a compromise. We partially disassemble the muscle, taking apart the thick filaments and isolating the head regions for our work, and removing the Z-line material but leaving the thin filaments themselves intact, although in disarray. In this way the proteins are dispersed evenly enough so that most of the techniques developed for homogeneous solutions can be applied. The fortunate property of muscle proteins that makes them amenable to detailed studies is that although this dispersal results in the loss of the mechanical manifestations of their activity (the ability to shorten and develop tension), the chemical manifestation (the splitting of ATP) survives intact. The influences of various artificial conditions on contractile activity can therefore be studied indirectly, by measuring their effect on the rate of hydrolysis of ATP, rather than directly as an effect on tension or shortening.

One of the first properties of muscle to be studied in this soluble system was the regulatory action of calcium ion. As one would expect from the behavior of intact muscle, the splitting of ATP in a solution of myosin heads and thin filaments was absolutely dependent on the presence of calcium ion. Ebashi had discovered that calcium control resulted from the presence of the proteins troponin and tropomyosin on the actin filament. When actin filaments were prepared that lacked these two proteins, contraction became completely insensitive to calcium, and the hydrolysis of ATP continued in an uncontrolled manner until the ATP was all gone. (This behavior is common to all vertebrate muscles but is not the only form of regulation found in nature. Andrew G. Szent-Györgyi discovered that among invertebrates an alternative form of control by calcium ion has been developed.) In order to understand how troponin and tropomyosin confer calcium sensitivity on the system, we must first discuss in a little more detail the chemical events of muscle contraction.

The chemical reactions resulting in the hydrolysis of ATP take place on the head regions of the myosin molecules, which form the cross bridges projecting from the thick filament. The splitting proceeds in several distinct steps. First an ATP molecule becomes bound to a particular site on the surface of the protein. The tendency for ATP to bind to this region is so great that under the conditions existing in a healthy muscle just about every myosin head has an ATP bound to it. The second step in the hydrolysis of ATP is the conversion of this complex of myosin and a bound ATP into a "charged" intermediate form. In contrast to the myosin-ATP from which it is derived, this charged intermediate shows a great tendency to bind to an actin molecule in the thin filament.

Although the charged structure is fairly stable when left on its own, it decomposes rapidly, with the splitting of ATP and the consequent release of energy, as soon as it becomes attached to

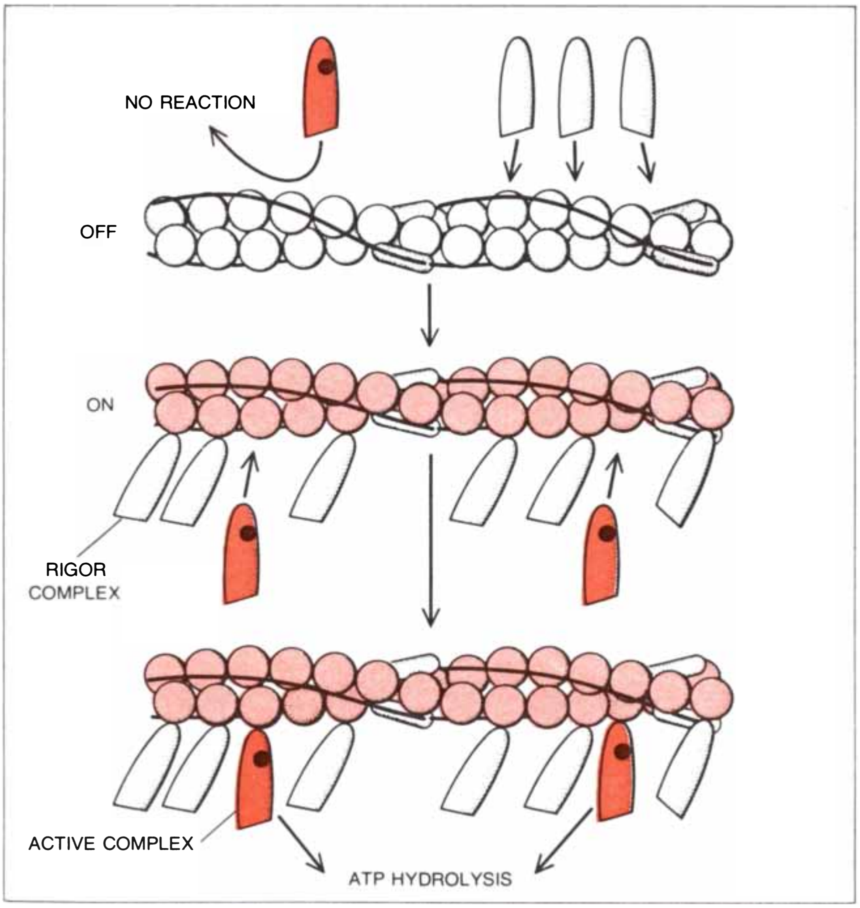


REGULATION OF MUSCLE CONTRACTION is accomplished primarily by calcium ions, which change the thin filament from an "off" to an "on" state, apparently by binding to troponin. When the thin filament is turned off (*top left*), a charged myosin-ATP intermediate cannot combine with it; when the filament has been turned on by calcium (*bottom left*), an active complex can be formed (*bottom right*) and hydrolysis proceeds. Formation of rigor complexes, however, is not sensitive to calcium control. Whether the thin filament is on or off, a myosin head can combine with an actin in the filament to form a rigor complex (*middle right*).

an actin molecule. In the soluble experimental system the liberated energy is wasted as heat. In the intact muscle, however, it is this energy that forces the cross bridge to swivel to a new angle, pulling the thick filament along with respect to the thin and thereby shortening the muscle. The final stage in the hydrolysis reaction, detachment of the cross bridge, is reached only after a new ATP has bound to the actin-myosin complex. The resulting actin-myosin-ATP complex rapidly dissociates to yield a free actin molecule and an uncharged myosin-ATP [*see illustration on page 64*].

By way of analogy, the process of liberating the energy of ATP hydrolysis can be likened to the firing of a gun. The gun must first be loaded by placing an appropriate cartridge (ATP) in a specific chamber (the myosin head). This combination (the myosin-ATP) is converted to a special metastable form by cocking the gun (the second step in the hydrolysis). If the cocked gun (or charged form) is left alone, it is more or less stable. If the trigger is squeezed (or if an actin molecule is available), however, the stored energy is rapidly released and work is done on a bullet (or a cross bridge). The process is completed by ejecting the spent cartridge (the products of hydrolysis, ADP and phosphate) and reloading.

The reader will perceive that two kinds of myosin-actin complex are formed in the course of hydrolysis, one high-energy and the other low-energy. The high-energy "active complex" is formed when the charged myosin-ATP intermediate links up with an actin molecule. It is short-lived: within about a hundredth of a second its ATP splits, energy is released and the complex decays to a low-energy state. This low-energy complex remains intact until a new molecule of ATP comes along. That cannot be long (perhaps a thousandth of a second) in the living muscle cell, where there is plenty of ATP. When there is no ATP, however, the low-energy complex is very stable indeed. We can see that in the laboratory, where such complexes can be formed by mixing myosin and actin molecules in the absence of ATP. In this situation the low-energy complex is formed not by decay of a high-energy complex but by the direct combination of actin with a myosin molecule that contains no bound ATP. Low-energy complexes formed in this way can be maintained as long as desired and then dissociated at will by the addition of ATP.



RIGOR COMPLEXES can usurp calcium's regulatory role when the ATP concentration is low. Charged myosin-ATP intermediates will not combine with a thin filament in the off state, but myosin heads will (*top*). If many rigor complexes are formed (which can only happen if there is not much ATP, which would detach the myosin heads), the filament is turned on (*middle*), charged intermediates are bound and hydrolysis proceeds (*bottom*).

The stability of the low-energy complex in the absence of ATP accounts for *rigor mortis*, the extreme rigidity that develops in muscle after death. The gradual disappearance of ATP following death prevents more and more of the myosin from binding an ATP molecule and thus leads to the formation of more and more rigor complexes between thick and thin filaments. Since there is very little ATP present to dissociate them once they are formed, the average lifetime of these complexes becomes very long and the muscle gradually becomes completely inextensible. Low-energy complexes are therefore referred to as "rigor" complexes. We refer to the combination of actin with the high-energy charged myosin-ATP intermediate as an "active complex" to suggest its role as the force-generating complex.

How is all of this regulated by calcium ion, in the absence of which the entire sequence of reactions comes to a standstill? Since the separate steps all proceed as part of a cycle, blocking any one of the individual reactions would inhibit the entire process. Data from a number of laboratories have shown that the step that is sensitive to calcium is the linking of the charged form of myosin-ATP to actin, that is, the formation of active complexes. Since tropomyosin and troponin are required for calcium to play its regulatory role, these two proteins are presumably responsible for blocking the formation of the active complex in the absence of calcium. If one recalls the relative amounts of the three proteins that go into making a thin filament, this responsibility raises a question: How do one tropomyosin and one troponin molecule control the behavior of seven actin molecules?

The location of the tropomyosin molecule in the filament suggests a possible mechanism for the blockage. Calcium ion is known to bind to troponin; if the response of troponin to losing its bound calcium (when calcium is removed from the fluid surrounding the filaments by the membrane calcium pump) is a movement such that tropomyosin is pushed farther away from the groove, toward the periphery, then the effect of this movement might be to physically block

Conversation Pieces

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

Stars and Atoms The molecules and atoms we have here on earth are made up of large volumes of space flecked with tiny dots of matter. For example, if you were sitting in the top row of a large football stadium (e.g., the Los Angeles Coliseum), the upper tier of seats would represent the orbit of a marble-sized electron. The atom's nucleus would be a BB sitting on the fifty-yard line. Everything in between the two would be empty space.

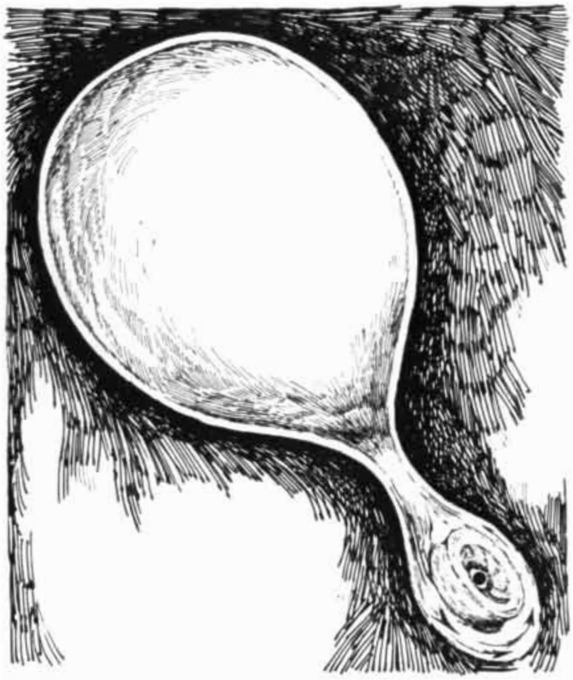
In the interior of stars, matter can be much more dense. For example, when a large star runs out of hydrogen fuel, the immense forces of gravity which have been held at bay by thermonuclear burning within the star suddenly become dominant. As Fred Hoyle puts it, the star has to pay all of its back gravitational taxes at once. The forces of this violent, almost instantaneous collapse are sometimes so great that the electron whizzing around the stadium is driven into the BB sized nucleus on the fifty-yard line. The two opposite charges cancel one another to form a neutron. Then, under the crushing force of gravity, the entire stadium fills up with marble-like neutrons.

Matter of this density exists in the heavens in the form of neutron stars or pulsars. In effect, pulsars are giant atomic nuclei in which the interatomic spaces of matter here on earth have been spectacularly reduced by gravitational collapse. We can learn about the structure of matter in such stars from the high energy radiation they emit.

Imagine now a stellar collapse so violent that the marble-like neutrons themselves are smashed together by the gravitational crush. The matter produced by such a collapse is unimaginably dense. The gravitational field of the resulting stellar object is so intense that no light (or any other kind of radiation) can escape its surface. Hence it is called a black hole. If you shined a flashlight directly on a black hole, you would see nothing for the photons of light would be sucked down its gravitational drain, never to return to your eye.

While black holes cannot be observed directly, their effects on stars unfortunate enough to be near them can be seen. Cygnus X-1 (see illustration) contains the first black hole tentatively identified. The hole is an invisible but dominant component of a binary pair of stars. It is sucking the material of its visible companion into a rotating disk. The violence of the transfer and shredding action heats up the atoms being sucked out of the visible star until they emit x-rays near the black hole, thus indirectly revealing its presence.

Today many physicists are interested in astronomy because much that we have to learn about the fundamentals of matter and energy can only be learned from the stars. That is why TRW Systems is building the High Energy Astronomy Observatory (HEAO) for NASA. The information this observatory will gather beginning in 1977 may well cause us to revise major portions of contemporary physics.



Cygnus X-1. Kip Thorne of the California Institute of Technology performed calculations leading to this model of the black hole. Our illustration is based on a painting of his model by Lois Cohen of the Griffith Observatory.

For further information, write on your company letterhead to:

TRW®
SYSTEMS GROUP

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

the approach of the myosin-ATP charged intermediate. This attractively simple model, which was proposed recently by Huxley and John Haselgrove on the basis of X-ray-diffraction studies, must be modified slightly to account for one additional observation: in contrast to the formation of active complexes, the formation of rigor complexes is insensitive to calcium, so that when myosin and actin are mixed in the absence of ATP, rigor complexes are formed whether calcium is present or not. Although the thin filament must be in a special "on" state (ordinarily achieved by the binding of calcium to troponin) before it can form an active complex, rigor complexes can be formed in the off as well as the on state; myosin and the charged myosin-ATP intermediate are chemically distinct species [*see illustration on page 65*].

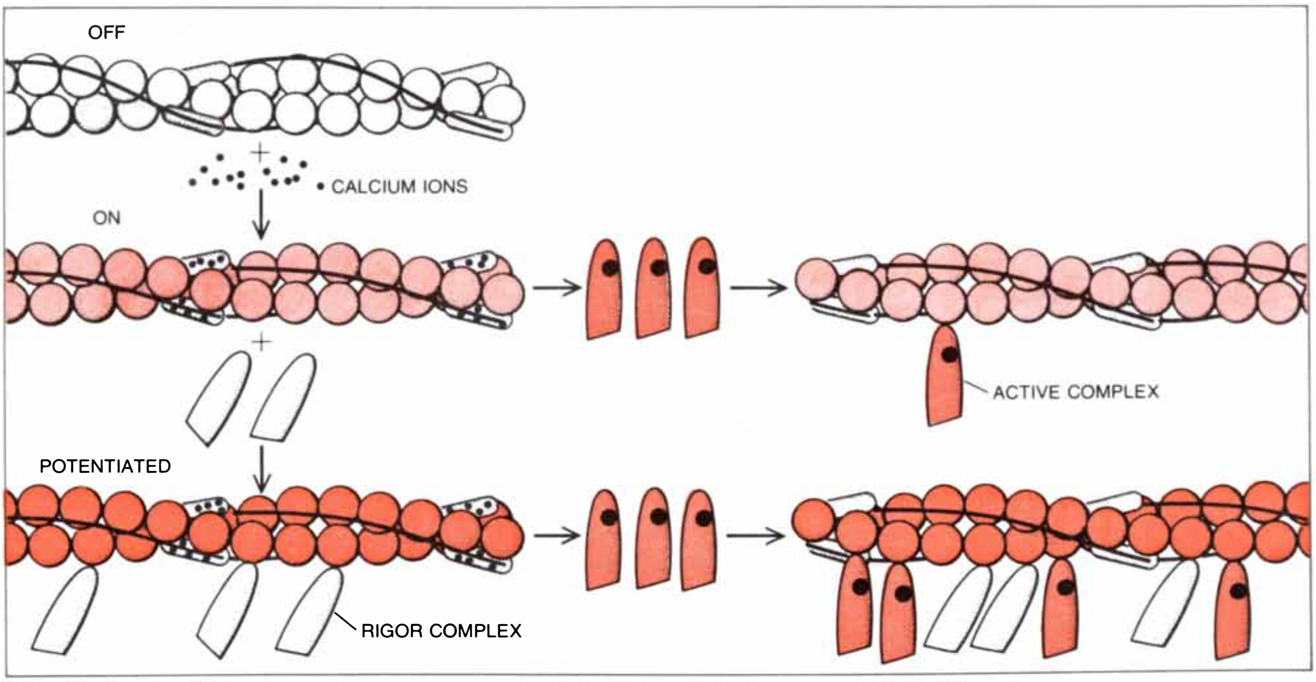
Apparently the actin molecules in the thin filament are able to distinguish a myosin molecule without a bound ATP (which would react to form a rigor complex) from one that has been converted into a charged intermediate (which would react to form an active complex). We reasoned that under experimental conditions exposing thin filaments to both forms of myosin (as is the case when the ATP concentration is low with respect to the myosin concentration) in the absence of calcium only the rigor complexes should be formed, and as a result there should be no hydrolysis of ATP. We carried out this experiment many times in several different ways, with the same result in every case: contrary to our expectation, ATP hydrolysis proceeded quite well! This puzzling observation was only the first of many indications that we had greatly oversimplified the behavior of the thin filament in muscle contraction.

The observation represents a loss of regulatory control. Since calcium was not present, the thin filament should have been "turned off" with respect to interaction with the charged form of myosin-ATP, but the observed ATP hydrolysis indicated unmistakably that the two proteins were interacting. Our first clue to the reasons for this apparent loss of control came from an experiment in which we compared the extent of loss of regulatory control at several different ATP concentrations. We found that at high ATP concentrations the regulation of contractile activity by calcium was normal; as the ATP level was lowered, the loss of control became increasingly severe until finally, at very low ATP concentrations, calcium had no effect on activity at all.

Why should this be? A direct effect of ATP on calcium binding did not seem likely because ATP does not interact at all with either of the regulatory proteins troponin and tropomyosin. We were therefore obliged to look for some indirect effect of ATP. We recognized that since the effect of ATP's binding to a rigor complex is to dissociate it quickly into myosin-ATP and actin, the higher the ATP concentration is, the smaller the number of rigor complexes on the thin filament must be at any one time. Could it be that the formation of rigor complexes on the thin filament somehow switched it on—that is, allowed the formation of active complexes, and thus the hydrolysis of ATP, to proceed—in the absence of calcium?

This possibility called for a new experiment. If in fact the loss of calcium sensitivity at low ATP concentrations was due only to the presence of greater numbers of rigor complexes on the thin filament and not to some other unknown effect of ATP itself, then varying the number of rigor complexes while keeping the ATP concentration constant should change the extent of escape from calcium control. The number of rigor complexes on the thin filament can be most easily altered by varying the ratio of myosin to actin in the experiment, and this can be done at constant ATP concentration. When we did it, we found that calcium control disappears at high myosin-to-actin ratios and reappears at low myosin-to-actin ratios, exactly as predicted on the assumption that rigor complexes turn on the filament in the absence of calcium.

It was satisfying to have the prediction confirmed, but we were a little startled by the implications. Control of the thin filament could no longer be re-



POTENTIATED STATE is a third state of actin. Molecules in the off state are turned on by calcium ions and in the on state they can form active complexes (*right*). At low ATP concentrations rigor complexes are formed. This raises actins that have already been turned on to a potentiated state in which more charged intermediates are bound and more contractile activity is shown.

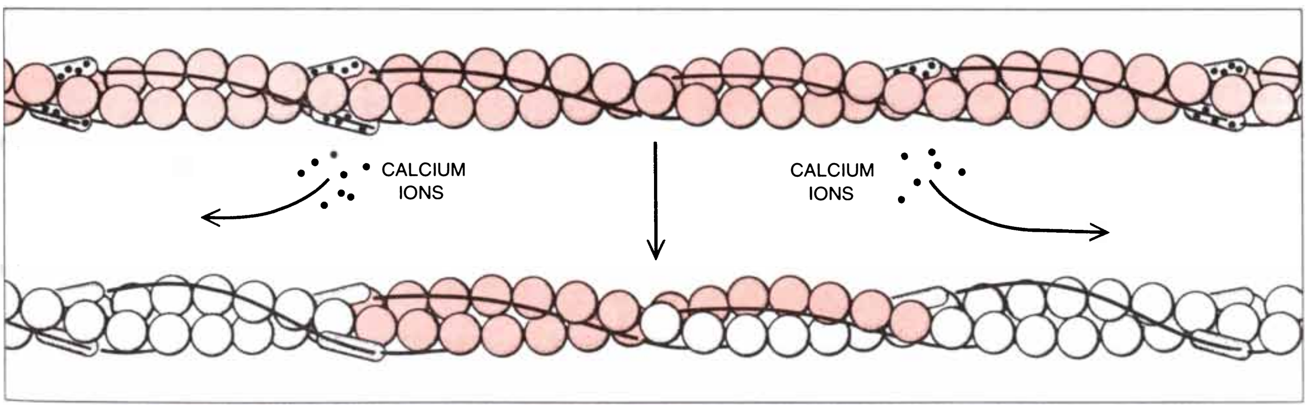
garded as a simple switching process mediated only by troponin and responsive only to calcium. This was demonstrated in other experiments as well. For one thing, it developed that the level of actin activity is modulated. When the effects of rigor complexes were examined more closely, we found they not only were turning on actin molecules that would otherwise have been off (in the absence of calcium) but also, in the presence of calcium, were acting to modify the behavior of any actins that were already turned on. The modification was manifested as a much higher than normal rate of ATP hydrolysis, which in our system represents the level of contractile activity. We call this state of actin a "potentiated" state, as distinguished from the on state, in which the actin molecules are turned on (because calcium has been bound to troponin) but exhibit lower activity because no rigor complexes are present. Further investigation revealed that the higher level of contractile activity in the potentiated state is the result of much faster formation of active complexes from actin and the myosin-ATP charged intermediate [*see illustration on preceding page*].

These experiments and others prompted us to reconsider our ideas about the behavior of actin molecules in the thin filament. The many actins making up a filament had been regarded as more or less independent entities that were arranged in a single structure merely to serve as convenient handholds for the cross bridges of the thick filament to grab as they pulled the thin filaments along and shortened the muscle. Accessibility to these handholds was controlled, in groups of seven, by a troponin in conjunction with a tropomyosin. Apart from the fact that they happened to be within the range of influence of the same troponin switch, actin molecules close together on a filament had been viewed as having no special relation to one another, nor had they appeared likely to have any "knowledge" of one another's actions. Clearly these concepts were no longer justified. If rigor complexes could influence the activity of the thin filament, there had to be communication between actin molecules. That was because the actin directly involved in a rigor complex could not itself be carrying out hydrolytic activity (an actin molecule can bind at one time only one myosin head, either charged or uncharged); it must be passing the message about the presence of a rigor complex along to an actin that was free to engage in hydrolysis.

The new concept of communication between actins along the thin filament proved to be a fruitful one, immediately suggesting a new set of questions (or at least new ways of phrasing old questions). How is the information transferred along the filament? How far along the filament can a message be communicated—from end to end or only to actins controlled by the same troponin? Can other messages be carried or only information about the current number of rigor complexes?

As we pondered the last question it occurred to us that perhaps regulation by calcium binding to troponin was part of the same process as regulation by number of rigor complexes. Possibly troponin conveyed the message "Calcium has been bound, turn on!" to the seven actins in just the same way that the message of rigor-complex formation is conveyed from one actin to another. If so, then perhaps messages could be transmitted in the reverse direction as well, from actin to troponin. On investigation we found that such reverse messages were in fact being sent to troponin, so that when rigor complexes were formed, the troponin got the message and responded by altering its behavior in a measurable way: it bound more calcium. Once we knew that more than one kind of message was being transmitted along the thin filament we looked for and found evidence of additional messages. It turned out that the actins on the filament were also informed about the number of active complexes present at any one time, and they responded to this information much as they had responded to rigor complexes on actin filaments in the on state (achieved by calcium): by potentiating the activity of the on state.

Having now determined that several kinds of messages could be sent along the thin filament, we considered how these messages were carried. The most obvious possibility was that actins spoke directly to neighboring actins. This could be tested, since it is possible to prepare filaments that contain only actin because their normal complement of troponin and tropomyosin is completely removed; such filaments have been studied extensively by Evan Eisenberg at the National Heart and Lung Institute and Carl Moos at the State University of New York at Stony Brook. Lacking troponin, these filaments are not under control by calcium concentration; they are always turned on. We could therefore not look for the switching-on message in these unregulated filaments but we could look for the potentiation message. It was completely absent. Neither rigor nor active complexes exerted any effect on the neighboring actins; individual actins appeared to be completely independent of



MESSAGE OF CALCIUM'S PRESENCE is conveyed from the troponin to which it binds, via tropomyosin, to the actins that the tropomyosin molecule reaches. This was established by an experiment in which the number of troponin molecules was reduced. Initially (*top*) calcium was present in the solution. Actins within the reach of tropomyosin molecules carrying troponin were therefore on. Actins beyond the reach of troponin were also on because actin is released from calcium control in the absence of troponin. When calcium was removed (*bottom*), only actins within the reach of tropomyosin molecules that carried troponins were turned off.

the rest of the filament. When we replaced tropomyosin (but without the troponin), the potentiation message returned. This made it clear that tropomyosin plays an important role in the transmission of messages.

Is tropomyosin itself the carrier, however, or does it merely modify the actins, so that in its presence they can communicate whereas in its absence they are isolated? We have not yet been able to obtain a final answer to this question, but we have made some progress toward it. We prepared thin filaments that lacked troponin but contained the normal amount of tropomyosin. Then we added troponin in very small increments and at the same time measured the spreading of calcium control along the actin filament. We reasoned that if the message of calcium binding was carried by actin, with tropomyosin as merely a necessary accomplice, then the message might spread for distances that were large compared with the length of one troponin-covered tropomyosin. On the other hand, if tropomyosin itself carried the message, then it would be expected to carry it only to actins covered by tropomyosin molecules on which a troponin was located. The latter expectation was fulfilled. The data on the relative amounts of troponin and of hydrolytic activity showed that only the actins associated with both a tropomyosin and a troponin exhibited sensitivity to calcium. The other actins, even though they were all associated with tropomyosin, did not appear to be aware of the troponin that was present at a distance. In this case at least the simplest explanation is that actins do not communicate directly but only indirectly via tropomyosin. Such an explanation fits well with an observation by Haselgrove and Huxley that tropomyosin is pushed back toward the groove, out of the way of the myosin heads, when the actin filament is turned on by calcium or by rigor complexes.

All the contractile systems studied to date, even those that lack troponin, have been found to contain tropomyosin. This has been somewhat puzzling, since tropomyosin alone has not seemed to influence the activity of the thin filament. The central role of tropomyosin in the phenomena described above suggests that those phenomena may be integral parts of the contraction process. Perhaps one of the reasons tropomyosin is invariably associated with contractile systems is that it has the ability to convert the independent actin molecules of a thin filament into a highly cooperative unit.

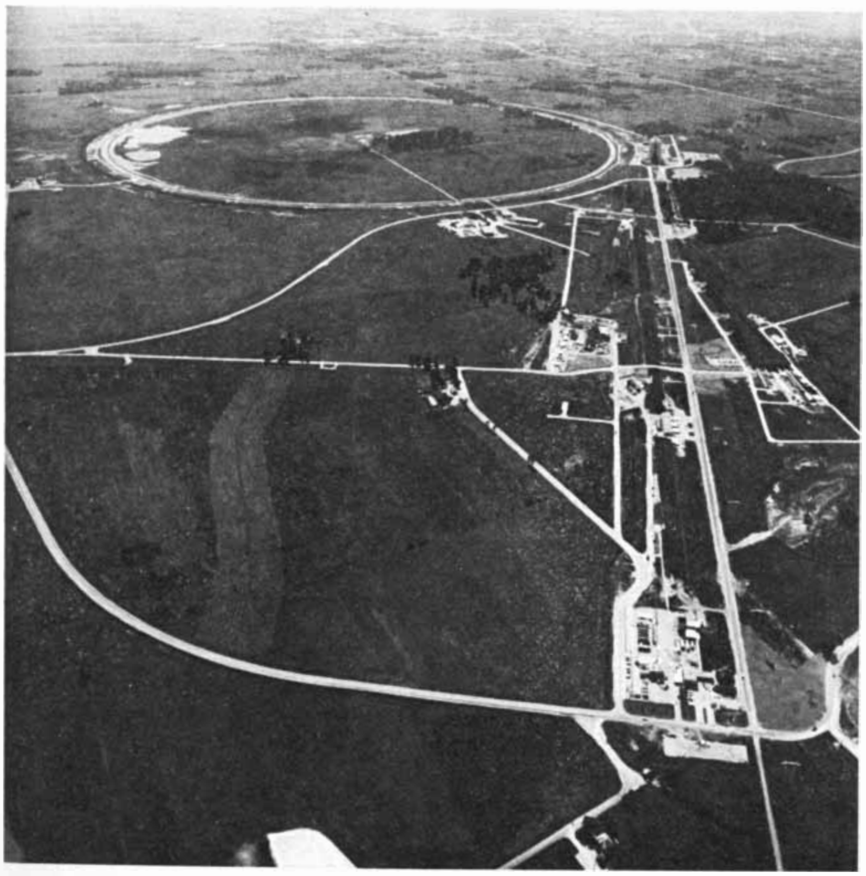
The Batavia Accelerator

The great Illinois synchrotron is currently accelerating protons to energies up to 400 GeV (billion electron volts). Its designers hope to reach 1,000 GeV with a new superconducting-magnet ring

by R. R. Wilson

From the most ancient times philosophers, and recently physicists, have been preoccupied by the pursuit of the ultimate atom. How elusive it has turned out to be! We are reminded of Milton's words: "And in the lowest deep a lower deep... opens wide." Yet the search goes on and what has been found thus far is both strange and wonderful.

During the past half-century physicists have built a succession of particle accelerators for probing the structure of the atomic, nuclear and subnuclear worlds. From the first Cockcroft-Walton generator of 1928 to the newest proton synchrotron at Batavia, Ill., the power of these instruments has increased millions of times. Yet in spite of all that has been learned the simple questions remain unanswered. Why is matter stable? Why is the electric charge on the electron exactly equal to the opposite charge on the proton? Why is that charge 1.6×10^{-19} coulomb and why is the mass of the electron 9×10^{-28} gram? Why is the proton 1,846 times heavier? We do understand now that electric and magnetic forces are merely different manifestations of the same thing: electric charge. We have not, however, found the same unifying relation between the other kinds of forces. Thus the gravitational force, the electric force, the nuclear force and the weak force could each be merely a different aspect of some analogous unifying thing. In fact, tentative new theories advanced by Steven Weinberg and Abdus Salam indicate that the weak force is another aspect of the electric force. Even if these theories are borne out, however, that would still leave the relation among the other forces a profound mystery.



MAIN RING OF WORLD'S LARGEST ACCELERATOR is housed in a circular tunnel two kilometers in diameter buried some 20 feet underground at the National Accelerator Laboratory at Batavia, Ill., about 35 miles west of Chicago. The machine accelerates protons to an energy of some 400 GeV, or to more than five times the energy of the second-largest particle accelerator, the 76-GeV proton synchrotron located at Serpukhov in the U.S.S.R.

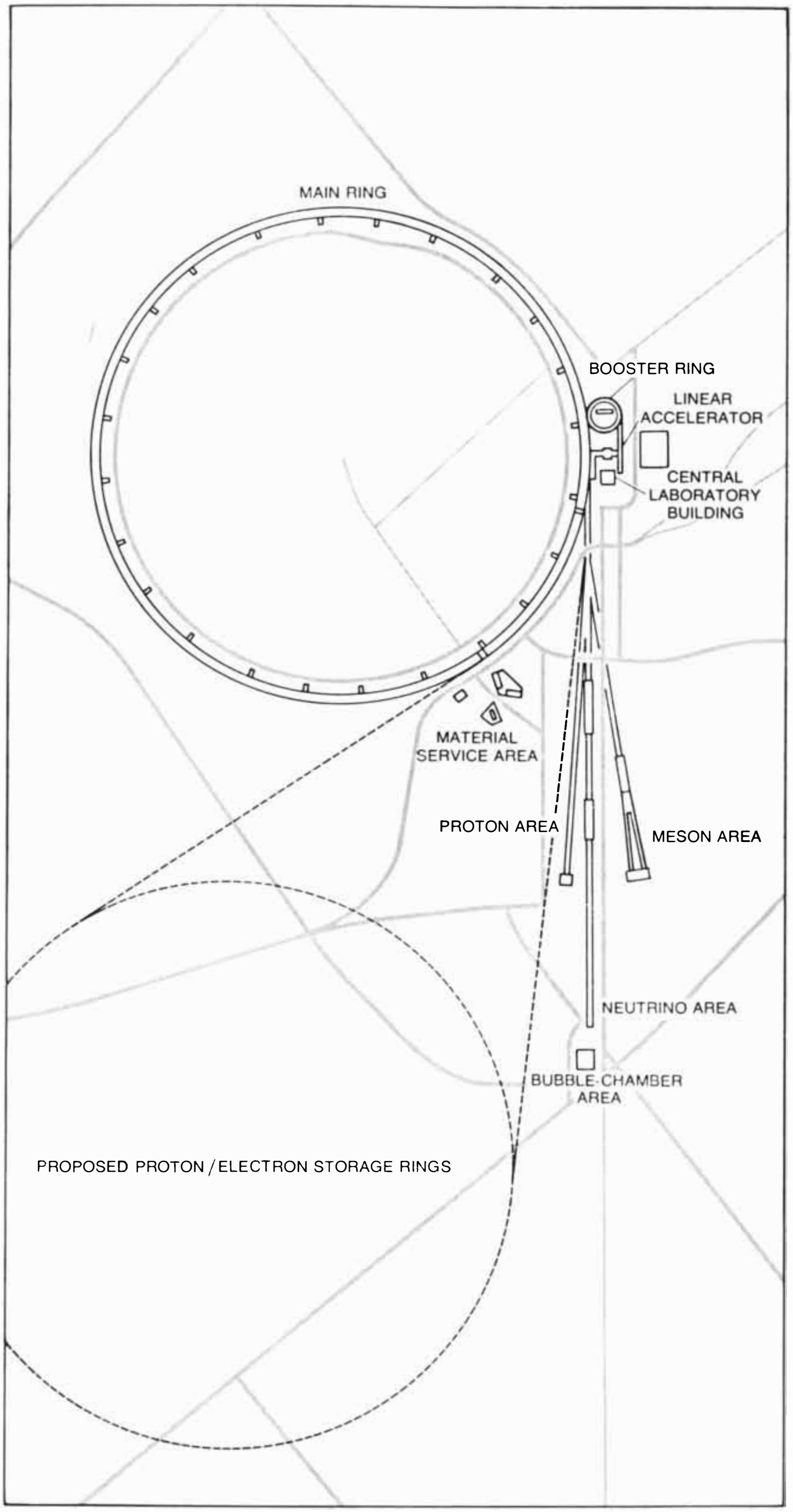
Nevertheless, considerable theoretical progress has been made. The "eightfold way" of Murray Gell-Mann and Yuval Ne'eman successfully explains many puzzling relations among the properties, masses and lifetimes of the scores of particles uncovered in the search for the ultimate units of matter. Carrying the eightfold way a step further, Gell-Mann and George Zweig proposed that all the hadrons (the particles that respond to the nuclear force) consist of combinations of a curious triad of new particles with fractional electric charge: quarks. In spite of a diligent search, however, the quarks have not been found.

A striking characteristic of the subnuclear world is that many of the properties of particles, designated by quantum numbers, do not change when the particles collide with each other. These observations have led physicists to formulate laws expressing deep symmetries in nature. Having discovered these majestic symmetries, physicists must explain how it is that some of them can be broken, even if ever so slightly. At this stage, then, the subnuclear world is rich in phenomena. It is a beautiful world, but it is still a mysterious one.

The new machine at Batavia was built in the expectation that clarity, perhaps even simplicity, might emerge from a deeper look into the subnuclear world. The Batavia synchrotron makes it possible for the first time to accelerate protons to an energy measured in hundreds of billions of electron volts (GeV) instead of tens of billions. One electron volt (eV) is the energy that a particle carrying the charge of one electron receives when it is accelerated through a field of one volt. The *G* in GeV stands for giga, the equivalent of 10^9, or a billion.

The most striking feature of the new synchrotron is a ring of magnets two kilometers in diameter that provides the final stage of a four-stage acceleration process. Protons, obtained by ionizing hydrogen in a discharge tube, receive their first push in an electric field of 750,000 volts produced by a Cockcroft-Walton generator, a direct descendant of the original particle accelerator. The protons are then directed into a linear accelerator, a copper vacuum tube about a meter in diameter and 145 meters long. It contains hundreds of pairs of copper electrodes between which oscillating electric fields are induced, tuned in such a way as to generate the electrical equivalent of a surf wave. The protons are accelerated as they ride this electrical wave down the tube, and when they emerge, their energy has been increased to 200 million electron volts (MeV).

The protons are then led into a medium-size synchrotron, the booster, a descendant of two famous machines that in their day were the most powerful in the world: the 6-GeV Bevatron at the University of California at Berkeley and the 28-GeV proton synchrotron at the Brookhaven National Laboratory. The booster, which can raise the proton energy to 8 GeV, consists of about 100 electromagnets arranged in a ring 150 meters in diameter. The magnets are pulsed from low to high magnetic fields at 15 cycles per second. Protons leaving the linear accelerator at 200 MeV are injected into the evacuated aperture of the magnets when the magnetic field is low and are guided by the field so that



MAP OF NATIONAL ACCELERATOR LABORATORY locates the main ring and other principal features. Plans are being made for two large storage rings, one to hold protons of 400 to 1,000 GeV, the other electrons of 20 GeV. Extremely energetic collisions could be obtained by diverting protons of 400 (or 1,000) GeV from the main ring into a bypass where they would run head on into either the stored protons or the stored electrons. Collisions between the protons in two 1,000-GeV beams would correspond in the amount of energy released to the protons in a single beam of 2.2 million GeV striking a target of protons at rest.

as it rises they make some 20,000 revolutions of the circular orbit. At each turn the protons pass through electrodes between which an oscillating electric field is generated. The frequency of the field is synchronized with the orbiting frequency so that at each turn the protons gain an amount of energy equivalent to being acted on by about .4 MeV. As a result after 20,000 successive boosts the protons are raised in energy from 200 MeV to 8,000 MeV, or 8 GeV, when the magnetic field has reached its full value.

At this point the protons are nudged out of the booster by a magnetic pulse and are guided magnetically down a long vacuum tube to the main accelerator, which is simply a larger version of the booster. It is 6.3 kilometers in circumference and incorporates 954 magnets. The protons are injected into the vacuum chamber of the main ring at 8 GeV and are accelerated in exactly the same way as in the booster. Physicists commonly refer to the circular vacuum chamber of a synchrotron as the "donut," preferring the simplified spelling. After 200,000 revolutions in the donut the protons have acquired an energy of 400 GeV. At the end of the acceleration process the protons are deflected out of the main ring and led to the experimental areas, where they strike the targets chosen for various investigations.

The Batavia accelerator had its origins a dozen years ago at the Lawrence Radiation Laboratory of the University of California. It was Ernest O. Lawrence who conceived the cyclotron principle for accelerating particles about 1928. During the 1930's and 1940's he participated in the design of a series of ever more powerful machines, culminating in the Bevatron, completed in 1954. By the early 1960's the Berkeley group had drawn up preliminary plans for a proton synchrotron of 200 GeV. At about the same time several Federal agencies had appointed a panel under Norman F. Ramsey of Harvard University to consider the next step in the nation's high-energy physics program. The Ramsey panel recommended that first priority be given to the construction of a 200-GeV machine based on the Berkeley design.

There was a prolonged controversy over the location of the machine. It was obviously a rich plum, politically and culturally as well as scientifically. Some 125 sites from one end of the country to the other were proposed. A committee of the National Academy of Sciences narrowed the list to six locations and left it to the Atomic Energy commissioners to make the final selection. They picked an area near Batavia known to the local inhabitants as Coon Hollow. Physicists quickly dubbed it "Dirksen Junction" in reference to the role played by Illinois's late senator in the selection of the site. Meanwhile a nationwide consortium of universities, known as the Universities Research Association, had been formed to build and operate the machine.

Congress gave the go-ahead for the construction of the accelerator in 1967 but limited the funding to only $250 million for a "reduced scope" installation instead of the $350 million the Berkeley study had concluded would be needed for a "full scope" effort. The reduction in scope meant a reduction of 10 percent in beam intensity (to 1.5×10^{12} protons per second), a reduction from five experimental areas to three and from one large bubble chamber to none. All these decisions provoked sharp debate behind the scenes. Thus was born the National Accelerator Laboratory, a fragile project for which many prophesied catastrophe.

At this point I was asked to head the project. In accepting the assignment I believed it might be possible to exceed the specified rating of 200 GeV and to mitigate the reductions in scope without exceeding the cost of $250 million. This was close to bravado, but only by such a challenge could I expect to attract the right kind of people to join in what to many seemed an exercise in frustration.

We decided to shoot for 500 GeV, to design for a full intensity of 1.5×10^{13} protons per second, to lay out experimental areas matched to the needs of critical experiments (including a big bubble chamber) and to try to produce a 200-GeV beam within five years. On June 15, 1967, a small group of stouthearted men rallied with me near Dirksen Junction to do just that.

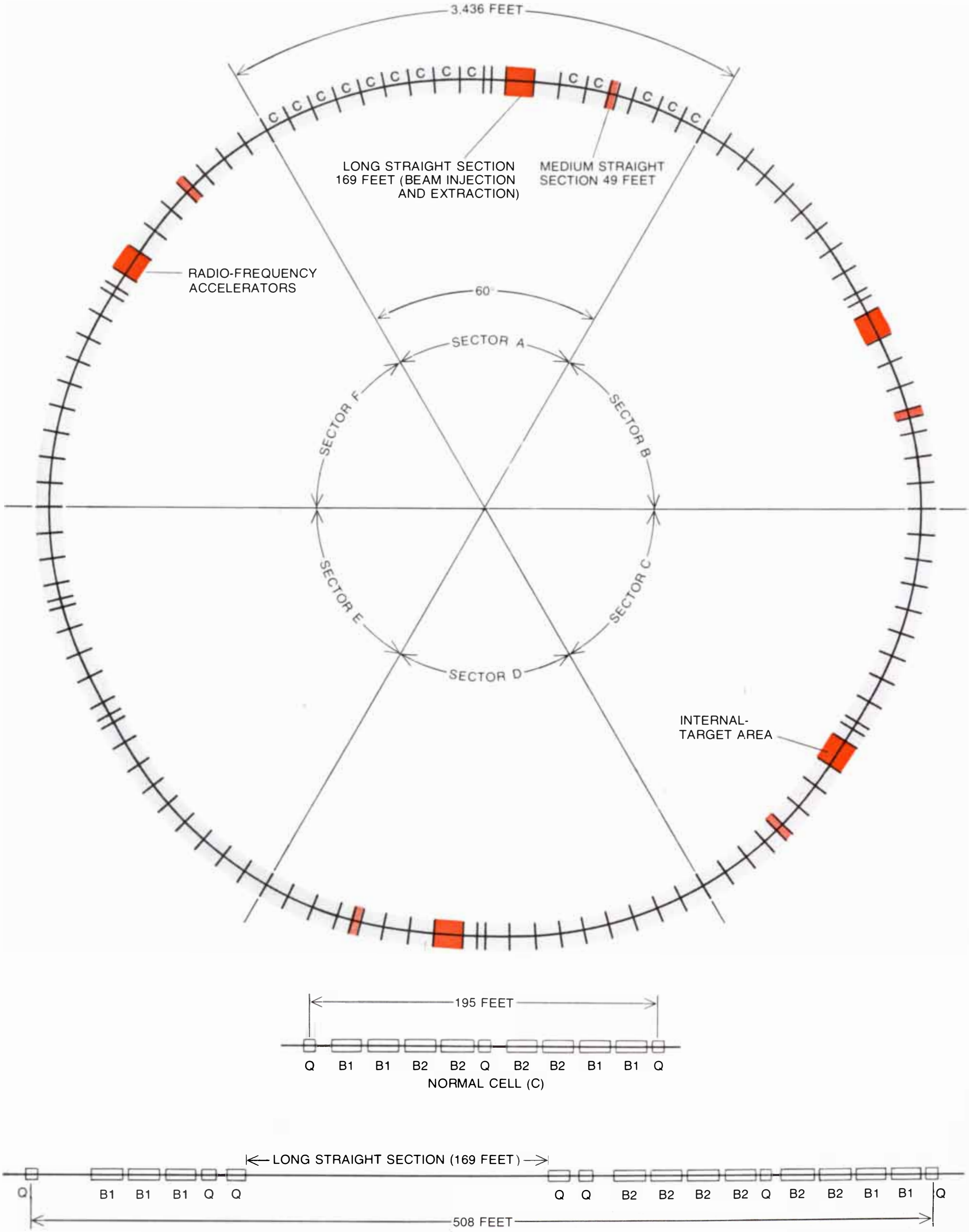
Where does an accelerator design start? We began by choosing the size of the main ring. The site, nearly 10 square miles, suggested a large ring. The larger the ring was, the weaker would be the magnetic field needed to bend protons of a given energy to the appropriate curvature. On the other hand, the larger the ring was, the higher would be the construction cost. If we were to reach 500 GeV, the minimum size of the ring would be set by the maximum practical strength of the magnets (depending also on their spacing).

Somewhat arbitrarily we selected a ring radius of exactly one kilometer. About 25 percent of the path along the orbit was made free of magnets to allow the injection and extraction of the proton beam, to provide room for orbit-measuring and -correcting equipment and to allow the placement of the electrical accelerating devices. With the magnets restricted to 75 percent of the ring circumference, we needed bending magnets that would generate a field of nine kilogauss to hold protons of 200 GeV inside a donut whose radius was one kilometer. With fairly conventional electromagnets of iron and copper one can reach a magnetic field of about 18 kilogauss. Accordingly, if we were able to reach that value, our ring would hold protons accelerated to 400 GeV. And 500 GeV would be within reach if we could attain a magnetic field of 22.5 kilogauss.

With this challenging prospect in mind we decided to design electromagnets ultimately capable of reaching the highest magnetic field attainable with conventional techniques. After a certain amount of soul-searching we rejected the possibility of installing pulsed superconducting magnets that theoretically could generate fields of several times 18 kilogauss. The art of building such magnets was not then sufficiently advanced for us to gamble on them.

Since the magnets and the tunnel in which they were to be housed would be among the most costly parts of the entire accelerator, we designed them for maximum economy. A critical dimension is the aperture between the poles of the magnet through which the evacuated donut is to pass. The larger the cross section of the donut is, the easier it is to keep the protons from striking the walls and being lost. Then too the protons electrically repel one another. Thus the beam tends to expand at very high intensities, with the result that the size of the magnet aperture is closely related to the ultimate intensity that can be achieved by a particular machine. On the other hand, the size of the magnet, the size of the power supply and the size of the tunnel grow directly with the dimensions of the donut, and so, of course, does the cost.

Balancing these factors, we chose an aperture two inches high and four inches wide for half of the magnets and 1.5 inches high and five inches wide for the other half. The choice was made largely on the basis of experience and intuition. The two different openings correspond to the undulating size of the proton beam that is characteristic of the "strong focusing" design. Lengthy calculations showed that the selected apertures should be adequate to accommodate the



"LATTICE" OF BATAVIA SYNCHROTRON (*top*) specifies the arrangement of magnets around the ring. In the Batavia design the functions of bending the beam and focusing it are performed separately by specialized magnets. The ring is divided into six identical sectors, each containing one long straight section (*dark color*), one medium straight section (*light color*) and 14 cells (*C*). The long straight sections are for beam injection and extraction, for the radio-frequency acceleration system and for the internal-target area. A normal cell and a long straight section are shown in elevation below the ring. Bending magnets are *B1* (narrow gap) and *B2* (wide gap). The quadrupole, or focusing, magnets are *Q*. A normal cell has eight bending magnets and two quadrupole magnets.

electrical repulsion of beams whose intensity was several times 10^{14} protons per pulse; hence the apertures should be more than adequate for our expected beam of 1.5×10^{13} protons per second.

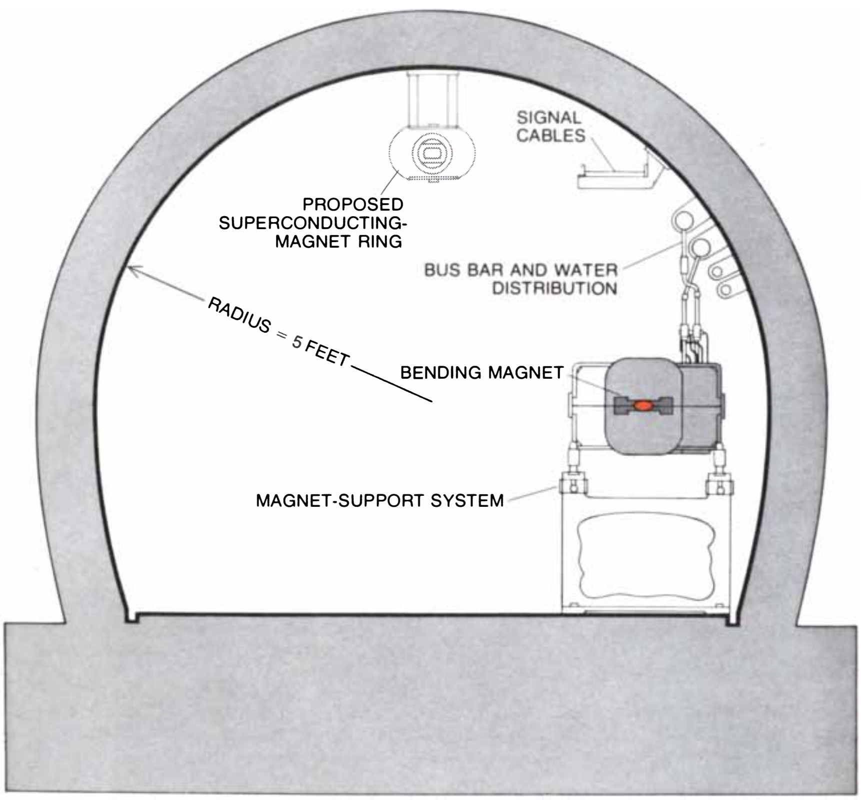
The protons repel one another most strongly at the time of injection, when the magnetic field is at a minimum. An even more important consideration in choosing the size of the aperture has to do with extracting the proton beam after it has been accelerated to maximum energy, when the field is at its maximum value. To extract the beam with high efficiency it is desirable to have the separation between two successive orbits as large as possible at the time of extraction so that the protons can most easily skip into a deflection channel. The separation is directly proportional to the horizontal width of the magnet and amounts to three-eighths of an inch for the apertures we have chosen.

In the accelerator design selected for Batavia the bending and the focusing of the beam are separate functions, performed by two different kinds of magnet. In all the accelerators that had been built up to this time the field ot each magnet was shaped so that the guiding effect and the focusing one were simultaneous. Such magnets are saturated at field strengths too low for the Batavia accelerator. The Batavia bending magnets are 20 feet long, one foot high and two and a half feet wide [*see illustration on opposite page*]. There are 774 of them, each weighing 11 tons. The focusing magnets, 180 in all, are only seven feet long, weigh five tons and are spaced around the ring at intervals of 100 feet. Called quadrupole magnets because they have four poles, the focusing magnets act on the beam of protons exactly as a series of lenses might act on a beam of light: they keep the beam focused along the center of the donut. The magnetic field generated by the four poles is canceled to zero at the center of the donut but increases in every direction away from the center. Thus a proton that wanders away from the center "feels" a force that on the average directs it back to the central orbit. The focusing forces cause the paths of the deviant particles to oscillate around an imaginary line that passes down the center of the donut.

If anything can be said to be new about the Batavia synchrotron apart from its size, it is the separation of the magnet functions. Also innovative, however, is the compact design of the magnets, which enables them to provide a good field at a much higher strength than had previously been attained. New too is the power supply, which ties the magnets directly to the power line through large selenium-controlled rectifiers; such rectifiers were just coming on the market as we started construction. Earlier proton synchrotrons had all drawn their power from huge motor-generator sets, which were less reliable and much more expensive.

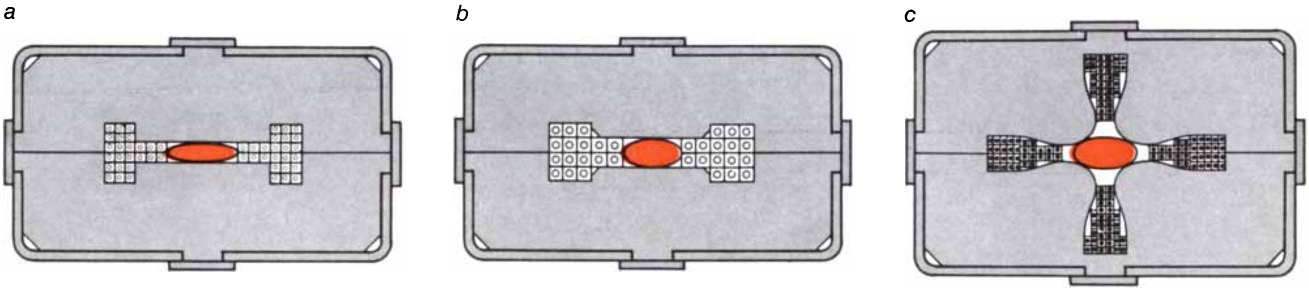
In the main ring the bending magnets are clustered in groups of four, each group being separated by a quadrupole focusing magnet [*see illustration on preceding page*]. All the magnets, particularly the focusing magnets, must be located with extreme accuracy: to within about a hundredth of an inch of the ideal central orbit. At first it seemed that to achieve such precision the magnets would have to be supported on concrete caissons resting on bedrock about 100 feet down. Since such caissons might have cost several million dollars, we decided simply to place the magnets directly on the floor of the tunnel. To provide against any slight settling the magnets are mounted on adjustable bases that can be realigned as needed.

Similar cost-saving decisions were made throughout the design period. For example, by making the magnets compact we were able to limit the tunnel to a width of 10 feet and a height of eight feet. The tunnel trench was dug by the open-excavation method. The tunnel itself was fabricated from precast concrete sections emplaced on a concrete floor slab. The top of the tunnel is 17 feet underground; the depth was chosen so that the floor slab could rest directly on hard glacial till.



MAIN-RING TUNNEL, shown in cross section, is 10 feet wide and eight feet high. The middle of the evacuated donut (*color*) in which protons are accelerated is three feet above the floor. The magnet-support system makes it possible to align the axis of the donut with an accuracy of .01 inch. Plans are being drawn for a second acceleration ring, to be suspended from the ceiling, which would be energized by superconducting magnets. Protons from the existing ring would be injected into the new ring, where their energy would be doubled.

Radiation damage has been a serious and continuing problem with the synchrotrons at Brookhaven and at the European Organization for Nuclear Research (CERN) in Switzerland. It has caused some of the critical parts of those machines, such as the insulation around the magnet coils, literally to crumble away. Other parts have become so radioactive that maintenance has been difficult and time-consuming. We feared that the problem at Batavia would be much worse because the intensity as well as the energy of our beam would exceed



THREE TYPES OF MAGNET USED AT BATAVIA are shown in cross section: the narrow-gap bending magnet (*a*), the wide-gap bending magnet (*b*) and the quadrupole, or focusing, magnet (*c*). The elliptical aperture in the narrow-gap magnet is 1.5 inches high and five inches wide; in the wide-gap magnet it is two inches by four inches. Both magnets are 20 feet long and weigh about 11 tons. Focusing magnets are seven feet long and weigh five tons. Spaced around the ring are 774 bending magnets and 180 focusing magnets.

previous levels by a factor of 10. This meant that the total energy in our beam, nearly a megajoule, would be greater by a factor of about 100, implying deleterious radiation effects in proportion.

Radiation damage is caused by protons that escape from the machine and dissipate their energy on whatever they happen to hit. If a machine could be built with such precision that all the protons were confined to the center of the donut until they reached their intended target, radiation damage would be nil. One of our primary design goals was therefore to accelerate and to utilize without loss every proton injected into our machine. For that reason we decided to have just one extraction point instead of several. We believed that we could tune the machine to optimize the extraction at one point, and that if we provided several extraction regions we might compromise the extraction efficiency at all of them.

We are confident eventually of being able to extract about 99 percent of the protons injected into the main ring, leaving only 1 percent to cause radiation damage or other residual radiation effects. Even this loss can be controlled by use of "scrapers": thick replaceable obstacles installed at intervals so close to the circulating protons that they are the first things to be struck by any errant particles.

It is always hard to get contractors to finish structures rapidly and on time. Our chief stratagem was to have the bids for construction made on the basis of a normal completion date. After the bid had been awarded we would sometimes offer the contractor a "deal" according to which he would receive a reward for each day he finished the building ahead of the schedule written in the bid. To arouse his interest a graduated system was devised such that if he saved only a few days from his schedule, he would receive, say, $100 per day. If, however, he could shave more than 10 days from his schedule, the bonus would go up to $300 per day. If he could advance his completion date to the "impossible" date in our private schedule, he might get as much as $1,000 per day for every day saved; savings in time beyond that date were not rewarded. On a $1-million building the maximum bonus could amount to only a few tenths of a percent; nevertheless, it appealed strongly to our contractors' sporting instinct, and the synchrotron enclosures were finished on schedule.

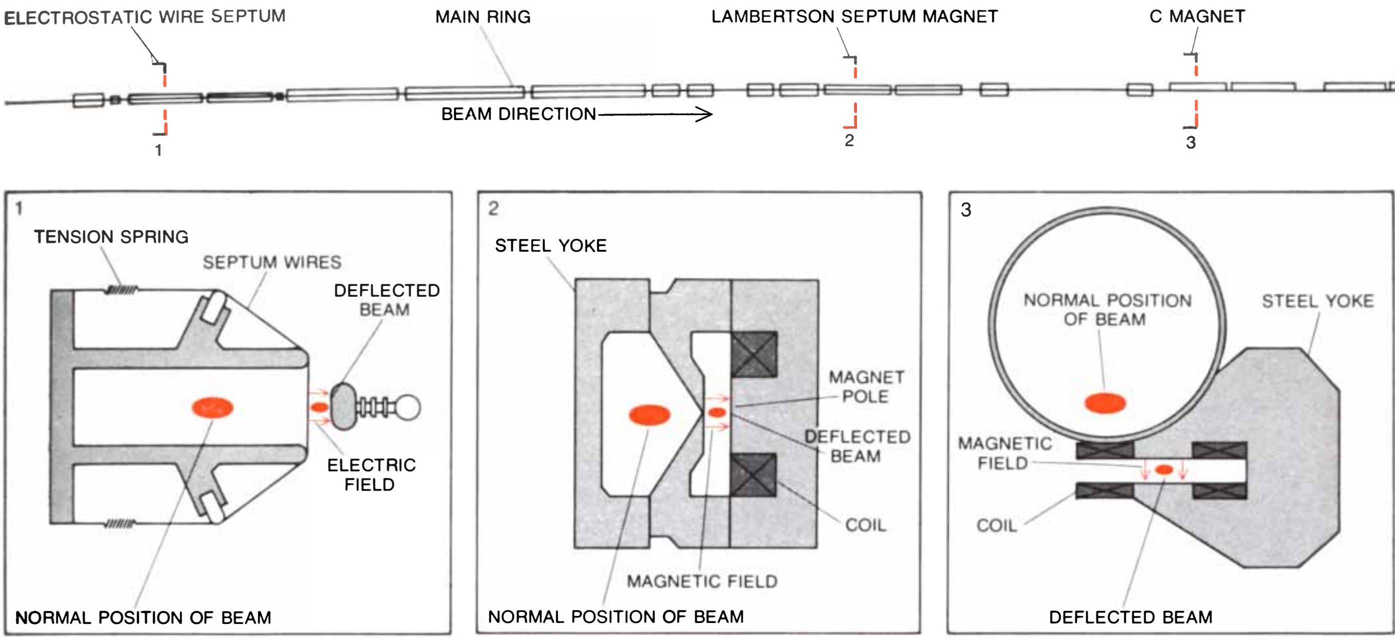
For the large orders of technical components we found it best not to be at the mercy of a single vendor. Sometimes a third of an order would be given to one manufacturer and a third to a second with the understanding that the final third would be allocated between the two on the basis of performance and cost. Of course, no mere bag of tricks is enough to keep a large project on schedule. There is no substitute for an aggressive attitude, a "gung ho" spirit and careful attention to detail on the part of a small cadre of devoted people.

Things went swimmingly—up to a point. One unexpected hitch was that Federal funds for the project came through at a much slower pace than had been scheduled. Since salaries represented a sizable fraction of total costs, my own response to funding delays was to make the deadlines earlier rather than later. Finally a combination of desperation and hubris led me to announce in 1969 the possibility that, with good luck and adequate funding, a beam of protons might be produced by July of 1971, a full year ahead of the original rash five-year schedule and only two and a half years after the groundbreaking of December, 1968.

This had one good effect: it set a goal for the men on the job to shoot for. Without drama it is hard to communicate with a large group. Almost all the other effects were bad. For instance, many of the physicists around the country who were beginning to plan experiments with the machine took the new date seriously and started to arrange their lives accordingly. We almost broke our backs trying, in spite of (or perhaps because of) reduced funds, to meet the deadline of July, 1971. We had all the main-ring magnets in place, all the power supplies on, all the radio-frequency circuits ready, the linear accelerator injecting protons into the booster and the booster injecting protons into the main ring, and none of these things came easily. We even succeeded at that point in making the protons circulate a few times around the main ring. Then the whole thing just seemed to fall apart, and our dream disintegrated.

What had happened? For one thing there were unexpected and unknown obstructions in the donut. The ring is more than four miles around and is made of about 1,000 stainless-steel sections welded together. A tiny piece of wire a few thousandths of an inch thick anywhere inside the donut could be enough to spoil the beam Discovering such a small impediment inside those miles of magnets boggled the imagination. The very first magnet we opened was found to contain a stray piece of plastic partially obstructing the donut.

How does one inspect the interior of a four-mile steel donut? As one measure of our desperation we tried to train a ferret to enter the tube, trailing behind it a long thread of wire to which we could then attach a magnet or a miniature vacuum cleaner. The ferret, Felicia by name, was happy to pull wire through short sample lengths of the tube, but it turned out that she was terrified by the



EXTRACTION OF PROTON BEAM is accomplished at the end of a long straight section of the main ring by a sequence of four devices: one electrostatic (*1*) and three magnetic (*2, 3, 4*). In the detailed cross sections the proton beam (*color*) is viewed head on in its normal circulating position and in subsequent deflected positions. The septum of fine wires in the electrostatic device acts as a shield so that when the beam is circulating normally the protons do not feel the force of the electric field that guides the beam when it

real machine. Eventually she earned her keep by helping to clean out long, straight stretches of pipe in the experimental area.

The task of cleaning the donut had to be solved by other means. We resorted to blasts of air, shock waves and magnets drawn through by wire (finally inserted without Felicia's help). In all we located and removed several hundred potential obstructions. Even after the machine went into operation, however, we continued to find bits of dirt and debris in the tube. They are now easier to find because they become radioactive when they are struck by protons and, if they are not vaporized, they can be detected with a portable ionization chamber. How did the obstructions get there? Mostly through carelessness, I think. For example, a defective "can opener" that was used to cut open the donut was responsible for leaving behind many steel slivers.

The worst problem and the biggest delay had to do with the magnets. They had been installed during the winter, and when they were tested individually, they worked fine. The tunnel had been completed during an exceptionally cold winter and was covered by frozen earth. With the coming of summer, humid air entered the tunnel; the water vapor condensed, causing a veritable rainstorm that soaked the magnets. Since the magnets were not designed to work under water many of them sparked over and failed when power was fed into them in July, 1971. We spent the rest of that year searching out not only the magnets that had failed but those that had been damaged to the extent that they might fail. Eventually we had to rebuild or replace almost half of them.

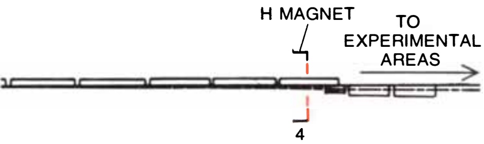
There were other difficulties. The entire machine, consisting of four accelerators in series, was controlled by an intricate electronic network that stretched out over a distance of more than five miles. All the enclosures had to be absolutely safe against radiation and electric hazard, and all the major components, perhaps 10,000 of them, and safety devices had to work at the same time. Now, if each component has a reliability of 99 percent, the probability that the assembly of 10,000 will work at the same time is only .0001 percent.

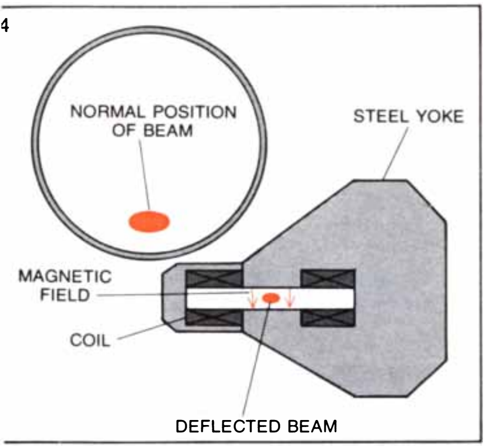
If the machine was to work hours on end without failure, it was necessary for each component to have an average reliability of better than 99.9999 percent. We knew from the start that this kind of reliability was an order of magnitude greater than had been necessary for any earlier accelerator. Although we had invented a new computer control system to provide an element of redundancy, we were still not quite up to building components with that kind of reliability at the beginning of the project.

Ultimately we solved these problems, along with many others, and achieved a beam of 200-GeV protons by March, 1972. That was later than I had hoped but still earlier than we had originally planned.

Making high-energy protons in itself is not very useful. What counts is how the protons are used as a probe of nature. At about the time we made our first batches of energetic protons we welcomed at the laboratory seven Russian physicists and their wives. The Russians proposed to study in collaboration with a group of American physicists the size of the proton as viewed by the protons from our accelerator. A similar experiment had just been conducted by members of the group at the 76-GeV accelerator at Serpukhov in the U.S.S.R., the most powerful machine in the world up to that time.

In preparation for the new experiment the Russian group had brought with it a device for shooting a jet of hydrogen into the donut to intersect the proton beam. If a high-energy proton collides with a proton that is nearly at rest, and if the collision is a glancing one, the proton at rest is deflected nearly at right angles to the path of the high-energy proton. By measuring the distribution pattern of the deflected protons the experimenters can deduce the proton's radius. The new results nicely extended





is in the deflected position. The stratagems that are used to drive the beam across the septum for the purpose of extraction are shown in illustration at bottom of the page.

the curve plotted earlier at Serpukhov: as the incident proton increases in energy it seems to "see" a larger and larger target proton [see "Proton Interactions at High Energies," by Ugo Amaldi; SCIENTIFIC AMERICAN, November, 1973]. We were delighted that our very first experiment was conducted by an international team.

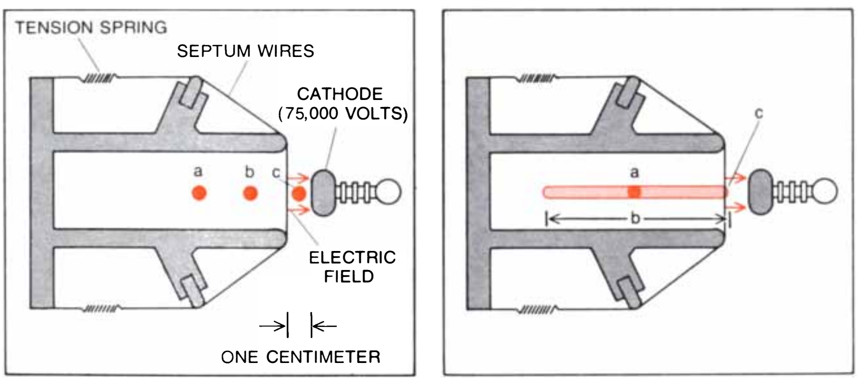
Not all experiments can be done as this one was, directly in the circulating beam of the machine (although the procedure has a number of advantages, such as making it possible to vary the energy continuously). For the great majority of experiments it is necessary to extract the beam and guide it to various large experimental areas. The extraction of a proton beam as energetic as ours presented novel problems. In one method (the first to be used) the beam is brought rapidly out of the main ring by magnetically kicking it into a region of electrostatic field provided by a septum consisting of fine wires [*see illustrations on these two pages*]. The electric field deflects the protons only slightly, but enough to direct them to a region of magnetic field that is shielded from the donut area, which in turn deflects the particles completely out of the accelerator.

A second mode of extraction is to allow the accelerator to resonate, that is, to allow the currents in the focusing magnets to change enough with respect to the bending magnets so that "betatron" oscillations are excited in the orbits of the protons as they circulate. At a rate of about 20 oscillations per turn the oscillations increase in amplitude, swinging the orbits farther and farther away from the center of the donut; finally the protons enter the same region of electrostatic field used in the first extraction method, and with the same result. Resonance extraction brings the beam out slowly over many turns, so that particles in the extracted beam can be spread out in time by as much as a second. This is useful for certain experiments where the means of counting events are limited. The two extraction methods, alone or in combination, should make it possible to get an extracted beam whose pulses last anywhere from 20 microseconds to one second.

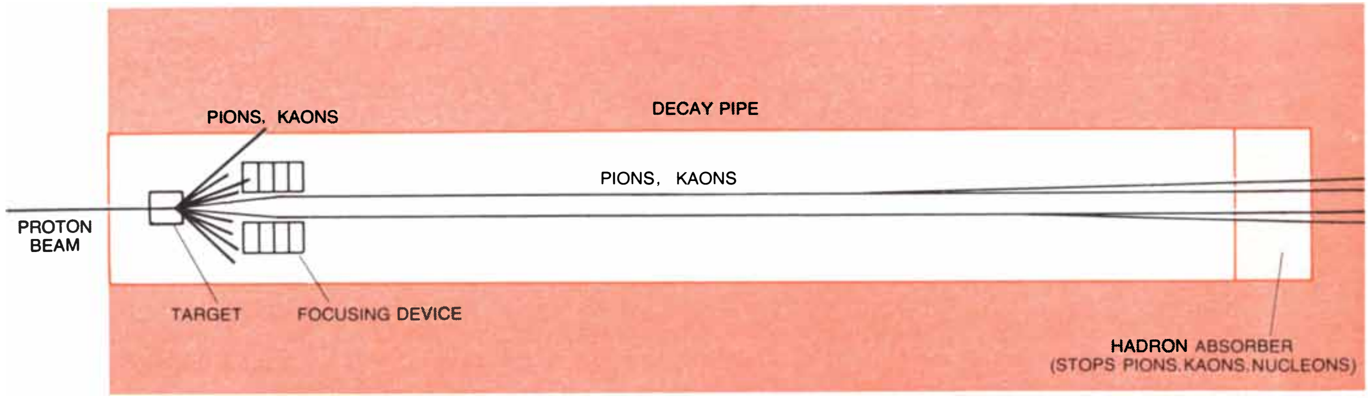
After the protons have been extracted from the accelerator they are guided to the experimental areas by means of a "switchyard," an arrangement of about 100 bending and focusing magnets almost as complicated as the synchrotron itself. Thus the protons can proceed straight ahead to the neutrino area, or they can be switched to the left (west) to the meson area or to the right to the proton area. Although the names are somewhat arbitrary, they are indicative of the principal activity in each area. All these experimental areas are roughly a mile from the extraction point.

Let us first consider the neutrino area. The protons are guided to the surface of the ground by magnets in the switchyard and strike a target, usually a piece of steel in a large mound of earth shielding. In the shower of particles created when the protons collide with the atomic nuclei in the target there are many different kinds of particle, but in this area we are interested primarily in the pi mesons, or pions, because they will decay into neutrinos. Most of the pions travel in a narrow cone in the forward direction and can be magnetically steered into an evacuated pipe that is about 300 meters long. As the pions travel down the pipe they decay into muons and neutrinos. The muons (and all the other particles present except neutrinos) strike the closed end of the pipe and are eventually absorbed in a mound of earth that is about a kilometer long. The neutrinos, which interact only weakly with matter, easily pass through the mound of earth and emerge as a pure beam at the far end.

If pions of all energies enter the decay pipe, then the neutrinos that emerge are also of all energies. If, however, the pions are magnetically sorted so that only pions with a narrow range of energies enter the pipe, then the neutrinos that emerge are comparably monoenergetic. In practice kaons are also present in the pion beam and likewise decay into neutrinos. Since the pion-neutrinos and kaon-neutrinos differ in their energy, the neutrinos in the emerging beam lie in two narrow energy bands. Such neutrino beams, with energies of up to several hundred GeV, provide one of the principal research tools at Batavia [see "Experiments with Neutrino Beams," by



TWO EXTRACTION TECHNIQUES enable experimenters to choose between fast extraction (*left*), which delivers the entire beam in 20 microseconds, and slow extraction (*right*), in which the beam is slowly peeled out over the course of one second. The two can also be mixed in any combination. In fast extractions steering magnets move the proton beam from *a* to *b* in about .01 second. A fast-acting magnet then "kicks" the beam to *c* in about one microsecond, so that all the protons in the main ring are emptied out in one turn. In slow extraction the beam is widened (*b*) by resonant excitation of horizontal "betatron" oscillations. As beam width grows, the part (*c*) that extends beyond wire septum is peeled off by a 75,000-volt electric field. Removal of entire beam is spread out over 50,000 turns, or about one second. In many experiments instrumentation would be saturated by fast extraction.



PURE NEUTRINO BEAMS are created at Batavia by directing the extracted proton beam against a steel target. The collision produces a shower containing all kinds of particles, but particularly pions and kaons. The charged pions and kaons are focused and directed down a 300-meter evacuated pipe. During flight both decay into muons and neutrinos. The muons, as well as other strongly interact-

Barry C. Barish; SCIENTIFIC AMERICAN, August, 1973].

From work on the interaction of neutrinos with protons done at other laboratories it is known that the cross section for interaction increases almost linearly with energy up to about 5 GeV. The linear rise seems to continue up to about 10 GeV, but the evidence at these higher energies is still not very reliable because the rate of events has been low. Our experimenters are now recording thousands of interaction events in the 100-GeV range, partly because of the greater intensity of our neutrino beam and partly because of an increase of a factor of 10 in the number of interactions. We are anxiously awaiting the analysis of an experiment conducted by a group from the California Institute of Technology to learn whether the linear rise continues or whether the curve bends over. On the basis of the first indications from another experiment it appears that the curve continues to rise as a nearly straight line up to 100 GeV.

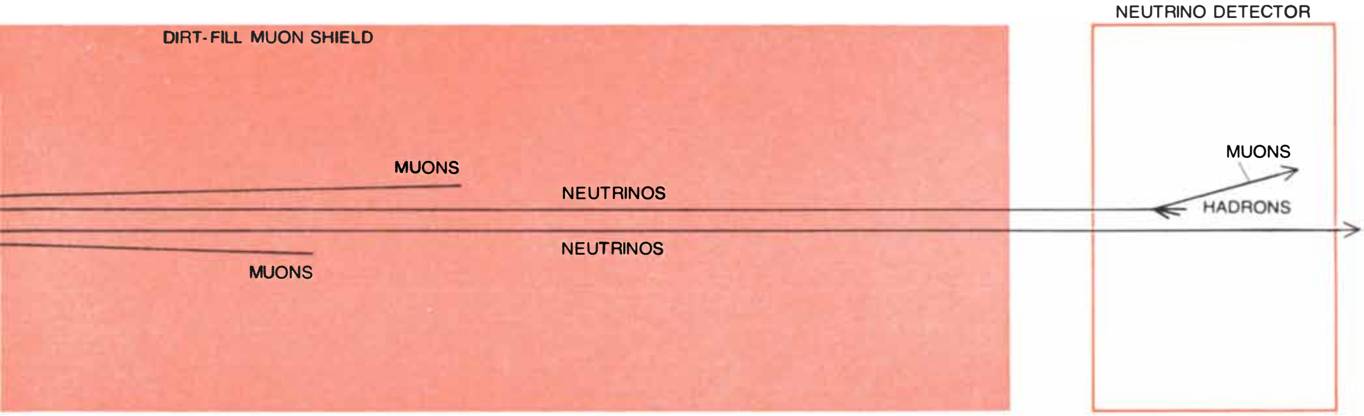
One reason for the curve eventually to bend over would be the existence of a still hypothetical particle called the intermediate boson or the *W* particle, carrying either a positive or a negative charge. The old data indicated that this particle would need to have a mass equivalent to more than 2 GeV. The data gathered by our experimenters show that the mass is greater than 5 GeV and probably greater than 10 GeV. If $W^{\pm}$ particles exist, a few of them should reveal their presence by simultaneously giving rise to two muons, one positive and one negative. So far there has been no definite indication that such an event has been observed.

Just as exciting as the discovery of the $W^{\pm}$ particle would be the demonstration of a neutral intermediate boson, because the existence of such a particle is predicted by Weinberg's theory that unifies the electric force and the weak force. Before this theory was developed it had been expected that a neutrino would interact with a proton and yield either a muon or an electron, depending on the kind of neutrino involved. On the basis of the new theory the neutrino could interact not by means of a charged current but by means of a neutral current, with the current being carried by the hypothetical neutral intermediate boson. The first evidence of muonless (or electronless) neutrino-induced events has come from CERN. When the large bubble chamber "Gargamelle" was exposed to neutrinos from the 28-GeV CERN proton synchrotron (with neutrino energies of some 12 GeV), events were observed that did not show muons or electrons.

At Batavia similar events were apparently being observed at higher neutrino energies in a huge electronic detector set up by a group of physicists from Harvard University, the University of Pennsylvania and the University of Wisconsin. Their results were subject to the criticism that perhaps a muon was created but not observed. They have repeated the experiment with better arrangements that should allow for a much higher probability of observing muons. The new results are still being analyzed. Meanwhile another run has been started by the Cal Tech group, whose apparatus is similar but perhaps even better suited to this kind of work. Moreover, the 15-foot Batavia bubble chamber is now coming into operation. Therefore we should soon have some definitive results in this area.

As I have indicated, the collisions that supply neutrinos for these investigations also yield a rich flux of muons as a result of the decay of pions. Some of the muons are led magnetically from the decay pipe to a passage extending along the west side of the earthen mound that filters the neutrino beam. In the muon laboratory we observe how muons of 100 GeV and up interact with matter. The magnet from a cyclotron at the University of Chicago has been moved to the muon laboratory, where it is being used as a giant spectrometer for detailed studies of muon interactions. Muons have also been utilized as a high-energy probe in experiments at the Stanford Linear Accelerator Center (SLAC). The muons produced at Batavia have a higher energy than those produced at SLAC, so that the SLAC studies can be considerably extended. The mere existence of the muon, which differs from the electron solely in having a mass 200 times greater, is a mystery that might be solved by such experiments.

The hydrogen bubble chamber at Batavia is located at the end of the neutrino area. Even though it was not included in the reduced-scope budget, it has recently been completed in the path of the neutrino beam. Nearby a 30-inch hydrogen chamber, moved from the Argonne National Laboratory, has been in service since we got our first external beam in the summer of 1972. Protons or secondary particles from the neutrino target can be steered into either chamber. Preliminary runs have been made with the 30-inch chamber using protons of 100, 200, 300 and 400 GeV and negative pions of 200 GeV. The results of the first exposures at 200 GeV, which were examined by the Argonne group that had provided the chamber, constituted the first scientific publication of the National Accelerator Laboratory: August, 1972. They showed that earlier results on the



ing particles present, strike the end of the pipe and are eventually absorbed in a mound of earth one kilometer long. The neutrinos, almost totally oblivious to matter of any kind, survive and enter the neutrino detection area, which includes a 15-foot hydrogen bubble chamber. Out of every 10^{12} neutrinos that enter the chamber, two, on the average, interact with a hydrogen nucleus, or proton.

multiplicity of particles produced in cosmic ray collisions were quite wrong. In the Batavia experiments 20 or even 30 particles materialized in a single collision. The bubble-chamber photographs offer a direct and aesthetically pleasing revelation of particle interactions at high energies.

Now let me turn briefly to the meson area. Here 300-GeV protons strike a target from which the secondary particles (usually mesons) are divided into six separated beam lines that terminate in experimental equipment. In this area more than elsewhere the experiments tend to resemble studies at lower energies in other laboratories. For example, several searches for the hypothetical quarks have already been made here. Although no such particle has been seen, new limits have been set on the probability of its materializing. In a secondary beam of neutral particles the neutron-proton interaction is being measured along with studies involving the regeneration of the neutral kaon. One of the secondary beam lines carrying charged particles is part of an elaborate spectrometer that is being used for precise studies of meson interactions at high energies. In a separate beam line the scattering of pions by protons and of protons by protons is under investigation, both at large scattering angles in order to reveal more about the structure of the proton.

On the east side of the neutrino area is the proton experimental area. The proton beam, instead of being brought up to the surface as in the other two areas, is raised from a depth of 20 feet (the level of the donut in the main ring) to a depth of 10 feet. In this way the natural shielding of the earth can serve to absorb the scattered radiation. A series of pits have been dug to hold the targets against which the full intensity of the proton beam can be directed. The area is designed to receive the most energetic protons we can make, perhaps as high as 500 GeV, compared with the meson area, which is now limited to a maximum energy of about 300 GeV.

A wide variety of experiments are being planned for the proton area. One will measure proton-proton scattering at larger angles than is possible in the other areas. Already in progress is an experiment being conducted by a group from Columbia University looking for electrons that may be emitted at large angles. A related experiment undertaken by physicists from Princeton University and the University of Chicago will make an exploratory search for anything emitted at large angles and high energy. The two experiments are designed to plumb the deepest regions of the proton. Also in the proton area we plan to create a beam of ultraenergetic photons to study the photoproduction of particles in the 100-GeV range. Later this year an electron beam will be produced to carry electron physics into the same unexplored range of energies.

Today the synchrotron and most of the experimental facilities can be said to be finished in the sense that most of the components are in place and have been operated. We can make protons with energies of up to 400 GeV. We have not yet succeeded, however, in getting the machine to work reliably or in producing intense beams of protons. The machine delivers a usable beam only about half of the time. The record beam intensity is 6×10^{12} protons per pulse, short of our design goal by nearly a factor of 10. Nevertheless, some 30 experiments have already been completed and another 30 are in progress or are being set up. Fifteen more experiments are scheduled for installation this year, and some 20 experiments have been approved but are awaiting either space or money.

Perhaps the most striking thing revealed by the first round of experiments is that when protons are struck by high-energy particles such as neutrinos, the high-energy particles created by the collision emerge at large angles. This is reminiscent of Ernest Rutherford's discovery some 60 years ago that when matter is bombarded with alpha particles (helium nuclei), the particles rebound at large angles. From this observation Rutherford immediately concluded that the atom has a massive and compact nucleus, thereby ruling out the "plum pudding" model suggested by J. J. Thomson, in which the mass of atomic matter was diffusely distributed.

Until recently physicists pictured the proton as being more or less homogeneous. By analogy with Rutherford's finding, the new observation that high-energy particles are emitted at large angles from the proton implies that parts of the proton are "hard." The observation holds true whether the proton is bombarded with neutrinos, with muons or with other protons, and whether the emerging particles are muons, pions, photons or other protons. In short, there is little doubt that we are beginning to perceive an inner structure of the proton.

A highly simplified way of looking at the proton, as it is now revealed, is to regard it as being composed of three (or more) hard little particles. Richard P. Feynman of Cal Tech, who originally referred to these constituent particles as partons, now thinks they are probably quarks. On this view one imagines that when a high-energy proton collides with a target proton, what actually happens is that a single quark in the incident proton collides with a single quark in the target proton [*see illustration below*]. The quarks' that result from the quark-quark impact should then hurtle off at large angles. Since we do not directly observe quarks, we assume that for some reason we do not yet understand a quark magically turns into an ordinary particle on the way out of the proton. Perhaps it does so by associating itself with other quarks in an ambient quark sea, or perhaps a quark-antiquark pair materializes in the quark-quark collision. Thus if the quark were to pick up one other quark on the way out, we would see a meson; if it were to pick up two other quarks, we would see a neutron, a proton or some more exotic heavy particle.

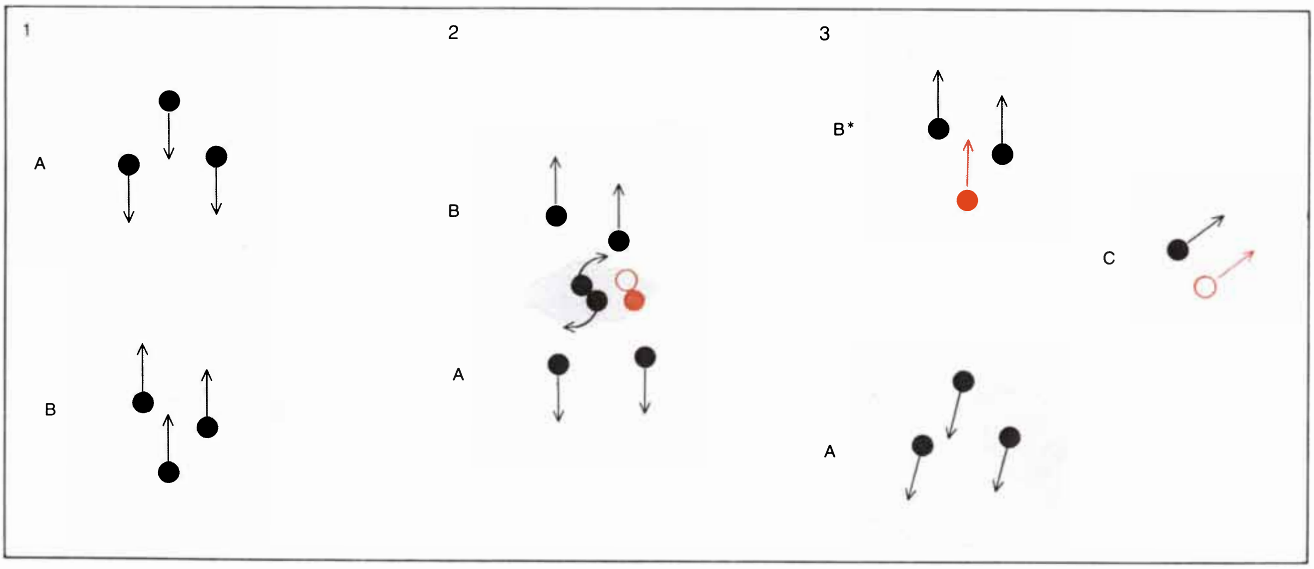
Since certain intrinsic properties, characterized by quantum numbers, are conserved in any interaction, one would expect at least two quarks with opposite (or canceling) quantum numbers to be created in a high-energy quark-quark collision. They might emerge as two particles or they might appear as two particle "jets": groups of particles coming out of the proton all within a small conic angle. There is actually some evidence for such jets. Experiments are now being designed to see if they really do occur, and if they do, whether they occur in correlated pairs as the quark model predicts. That we can begin to discuss such remote possibilities represents a tremendous advance over our thinking of even a few years ago.

Because the phenomena I have been describing become more dominant as the energy of the collision increases, we have naturally considered how one might press beyond 400 GeV at Batavia. The machine now runs about as well at 400 GeV as it did at 300 GeV, but nearly twice the electric power is required to achieve the higher level at the same pulse rate. As a result the Commonwealth Edison power lines show a 1 percent drop in voltage. At 300 GeV we had already been pressing the agreed-on limit of a .5 percent drop. Thus the voltage drop sets the present limit to higher proton-beam energies. To eliminate the drop we are planning to install appropriate capacitors in series with the secondary winding of the power transformer. In a few months we should learn if the capacitors will enable us to approach 500 GeV. Other effects of the large surging load may conceivably present a new limitation.

An energy of 1,000 GeV (one tera-electron-volt, or TeV) does not seem an impossible goal. To reach that level we could install a second donut fitted with superconducting magnets in the same tunnel as the present main ring. The old ring would then serve as a booster accelerator from which protons would be transferred to the new superconducting ring.

We have built a number of short superconducting magnets. Our tests make us optimistic that such magnets can generate a field of 45 kilogauss, which is what would be needed to reach 1,000 GeV. At this point we still have some $30 million in unexpended funds, which might be barely enough to build a superconducting ring next to the existing ring, provided that we do not spend more than about $10 million for improving our experimental areas or building new ones.

We are impelled to build the superconducting ring not only by technical considerations but also by financial ones. The power at the present level of operation is about 40 megawatts, and we would like to double or even quadruple that figure to achieve the kind of operation we desire (for example to generate pulses with a long flat top). Our present power bill is several million dollars per year, and it would be multiplied accordingly. We calculate that, if the transfer of the protons to the superconducting ring were to occur at 100 GeV, the total energy required by both the main ring and the cryogenic apparatus of the superconducting ring would be less than 20 mega-



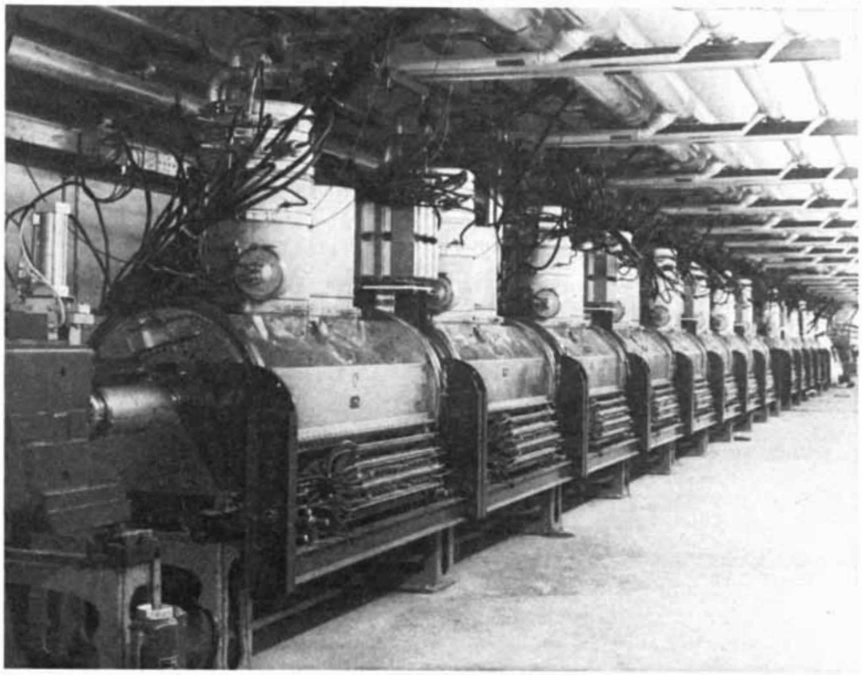
QUARK MODEL, here highly simplified, may explain how a secondary particle, such as a pion, might materialize and be ejected at a large angle in a high-energy proton-proton collision. The diagram at left shows two protons, *A* and *B*, approaching at high velocity. Each is depicted as containing a cluster of three quarks (or "partons"), whose direction of motion is shown by arrows. The gray region surrounding the quarks indicates the remainder of the proton, conceivably the "glue," or force field, that holds the quarks together. In the middle diagram proton *A* has passed almost completely through proton *B*. Although four of the quarks have passed by without "seeing" one another, two of the quarks have collided. The force of the impact has produced a quark-antiquark pair (*color*). In the diagram at right the quarks have regrouped into proton *A* and particle *B**, which can be regarded as an excited state of proton *B*. The new antiquark joins one of the original quarks to form a pion (*C*), which flies off at a large angle. Proton *A* exhibits a slight deflection, necessary to conserve momentum. Actual events in such a collision are undoubtedly far more complex than this.

watts. That would hold true for all modes of operation, either in stretching the pulses for experiments limited by counting rate, in attaining the highest energy of 1,000 GeV or doing both at the same time.

There are persuasive arguments, however, for also moving in a totally different direction. It is well known that as one increases the energy of a beam of protons proportionately less energy is available when the protons are simply allowed to slam into protons at rest. Actually the available energy increases only as the square root of the incident energy. Thus at 1 GeV only .43 GeV is available, at 10 GeV only 2.9 GeV, at 100 GeV only 10.5 GeV and at 1,000 GeV less than 35 GeV.

The answer is to direct one beam against another of the same energy, thereby making all the energy available for the collision of particles. This is being done successfully in the intersecting-storage-ring facility at CERN. There a proton synchrotron first fills one ring with 30 GeV protons going in one direction, then fills another ring with protons of the same energy going in the opposite direction. The rings are built so that the beams are interwoven, intersecting at eight points where 60 GeV are available for particle collisions. If only a single beam were used, 1,800 GeV would be necessary to get the same collision energy.

Now, there are some decided advantages in working with a simple high-energy beam. For one thing, there are millions of times more collisions when protons strike a target of, say, liquid hydrogen than there are when the protons in two tenuous colliding beams meet. For another, many secondary particles are made in a collision of a high-energy proton with one at rest, and the energy of the moving system is shared among them. Thus high-energy secondary beams of pions and kaons, as well as tertiary beams of neutrinos and photons, are available from a single-beam high-energy machine but not from a colliding-beam facility. These secondary particles provide for simpler collisions than those between two apparently complex protons. Finally, the beam from the Batavia machine could be used to fill a set of separate intersecting storage rings. For example, collisions between two 100-GeV beams would correspond to a single beam of 22,000 GeV striking a target at rest, collisions between two 400-GeV beams would correspond to 320,000 GeV and collisions between two 1,000-GeV beams would correspond to 2.2 million GeV.



RADIO-FREQUENCY STATION is located in one of the long straight sections in the main ring of the Batavia accelerator. The 15 identical units in the photograph are resonant electromagnetic cavities that impart energy to the proton beam in increments of about 2 MeV per cycle. As the frequency of revolution increases, the frequency of the accelerating voltage must increase in step. As the protons gain energy in each cycle they increase in velocity from an initial value of 99.6 percent of the speed of light at the time they enter the main ring at 8 GeV to a final value exceeding 99.999 percent of the speed of light at 400 GeV.

Although there are a number of intermediate steps we might contemplate, such as to arrange for the 8-GeV protons from the booster synchrotron to collide with the 400-GeV protons from the main accelerator (equivalent energy 6,400 GeV), we have been advised by our Long-Range Planning Committee to move directly toward a more ambitious utilization of our accelerator and our site. That would be to build a colliding-beam facility that would be filled by the 1,000-GeV protons from the superconducting ring. The committee also recommends that an electron ring capable of providing a stored beam of, say, 20 GeV, be built in the same tunnel. The idea is not to arrange for esoteric collisions between an electron and two protons but to provide for the possibility of studying either 1,000-GeV protons colliding with 1,000-GeV protons or 1,000-GeV protons colliding with 20-GeV electrons.

In this three-ring circus some of the magnets in the vicinity of the straight sections where the collisions occur might have to be moved in order to achieve one or the other of these alternatives, but that is not a serious difficulty. The reason for the committee's recommending such a course is that it is known that at high enough energies the increasing strength of the weak force will overcome the decreasing electric and nuclear forces, leading to a logical inconsistency where "something will have to give," that is, where new phenomena will have to reveal themselves. It is not known whether electron-proton collisions or proton-proton ones will be best for this purpose; that is the reason for seeking to provide both alternatives. We are moving ahead to a serious design study that will show us what is possible and how much it will cost.

The ultimate capability of the National Accelerator Laboratory is limited only by the size of the site. The largest superconducting ring we could build within our present boundaries would have a circumference of about 10 miles. If the facility were designed as an intersecting-storage-ring system, it might enable us to reach collision energies of several million GeV. If the experiments we are now capable of doing do not yield the knowledge we desire, or, what is more likely, if the new knowledge makes it irresistible to discover what happens at very much higher energies, we are confident those energies can be achieved at our laboratory on the Illinois plain.

NUTRITION AND THE BRAIN

When a meal rich in carbohydrate is eaten, the brain makes more of the nerve-impulse transmitter serotonin. The mechanism may be part of a closed circle in which diet influences food consumption

by John D. Fernstrom and Richard J. Wurtman

Does eating influence brain function? To put the question more specifically, do the changes in blood chemistry that follow the intake of nutrients produce corresponding changes in the tissues of the brain? If so, could such diet-induced changes affect the functional activity of the brain? Finally, might the "hungry" brain feed itself, so to speak, by "ordering" the body to eat?

In the past the usual answer to these questions has been no. The brain extracts oxygen, glucose and other nutrients, such as amino acids, from the bloodstream at a rate that seems unrelated to the concentration of those substances in the blood. For this reason transient fluctuations in the composition of the blood following the ingestion of food or during a period of fasting have not been thought to influence the brain. There is, however, an exception to this general rule and we and our colleagues in the Department of Nutrition and Food Science at the Massachusetts Institute of Technology have been investigating it for several years.

The exception is as follows. After a meal rich in carbohydrate there is an increase in the rate at which the brain synthesizes the neurotransmitter serotonin. Neurotransmitters are the substances present in the neurons of mammals that, when released, transmit signals across synapses to other neurons within the brain or to muscle cells or secretory cells outside the brain. Serotonin is one of six compounds that are reasonably well established as being neurotransmitters in the brains of mammals. The others are acetylcholine, gamma-amino butyric acid and the three catecholamines: epinephrine (or adrenalin), norepinephrine and dopamine.

In laboratory rats the brain's response to the ingestion of carbohydrate begins within an hour after eating. It is preceded by a change in the normal concentration of amino acids in the blood. The concentration of most of the amino acids decreases, but the concentration of the amino acid tryptophan increases. As a result there is a proportional increase in the brain's uptake of tryptophan. Now, tryptophan is a precursor in the synthesis of serotonin. Therefore after the ingestion of carbohydrate the brain neurons that convert tryptophan to serotonin are able to do so at a higher rate, and the level of serotonin in the brain rises. Since an increased amount of neurotransmitter will probably be released whenever the neurons that store serotonin are excited, it appears likely that brain function will be modified.

That the brain should be sensitive to diet-induced changes in the blood seems at first not only surprising but also disconcerting. What physiological advantage can there be to the individual when an organ as important as the brain can be affected by the vagaries of food intake? If in fact there is no advantage, how has such an open-ended mechanism for modifying brain function survived the tests of natural selection? Whatever the answer may be, there is no escaping the conclusion that nutrient intake does alter the amount of serotonin present in the serotoninergic neurons of the brain.

Like the three catecholamines, serotonin is a monoamine; specifically it is 5-hydroxytryptamine. It was first discovered in 1948 by three workers at the Cleveland Clinic: Maurice M. Rapport, Arda Green and Irvine H. Page. They isolated the substance from the blood of mammals, primarily the blood platelets. Soon afterward it was learned that relatively large quantities of serotonin are present in the mammalian brain, and within a decade the Swedish neuroanatomists Annica Dahlström and Kjell Fuxe, working with the fluorescence microscope, had located the neurons where serotonin is stored. They found that virtually all the serotonin present in the brain is confined to a distinct group of neurons known as raphe nuclei. The cell bodies of the raphe-nuclei neurons are situated in the brain stem; their fibers ascend into the rest of the brain and descend through the spinal cord to form widely distributed synapses.

Tryptophan is one of the "essential" amino acids, that is, it cannot be synthesized by mammalian cells. Whatever store of tryptophan an animal has can be obtained only by the ingestion of protein that contains the substance. This is not the case for many of the other amino acids, for example tyrosine, the precursor of both norepinephrine and dopamine. Tyrosine can be synthesized in the mammalian liver, with the amino acid phenylalanine as a precursor. Tryptophan not only has to be obtained by ingestion; it is also the least abundant amino acid in dietary protein. As a result the mammalian body's supply of tryptophan is usually quite small.

The transformation of tryptophan and tyrosine into neurotransmitters begins in the same way: by the addition of a hydroxyl group (OH) through the catalytic action of an enzyme. The two enzymes involved, however, differ significantly. Tryptophan hydroxylase is a low-affinity enzyme. This is to say that a molecule of tryptophan and a molecule of the enzyme have no strong inclination to combine. Therefore it is only when the concentration of tryptophan is much higher than normal that the enzyme can function at the maximum rate. In contrast, tyrosine hydroxylase is a high-affinity enzyme; it operates at maximum efficiency even when the concentration of tyrosine is no greater than normal.

We first became interested in the relation between nutrition and serotonin synthesis in 1968, when a graduate student in our laboratory, William Shoemaker, found that the brains of weanling rats given a diet deficient in protein contained abnormally low levels of dopamine and norepinephrine. In view of the fact that tryptophan is present only in small amounts even in a normal diet and, unlike the precursor of the two catecholamines, is available only from proteins, it seemed to us that a tryptophan-poor diet should depress the level of brain serotonin even more strongly.

We proceeded to feed rats for several weeks a diet that had only one source of protein: corn, a food known to be particularly deficient in tryptophan. When we measured the levels of tryptophan and serotonin in the rats' brains, the levels proved to be far below normal. With the long-term dependence of brain-serotonin levels on the levels of brain tryptophan thus established, we wondered whether the same dependence could be demonstrated when the change was a rapid, short-term one. It seemed possible that such acute changes might occur spontaneously during the course of a day; we had earlier observed that in man the level of blood tryptophan increases and decreases with a characteristic daily rhythm.

Working with rats, we found that animals fed a normal diet and assayed at four-hour intervals day and night also exhibited rhythmic changes in the level of blood tryptophan. Moreover, a brain-tryptophan rhythm existed and paralleled the blood-tryptophan rhythm.

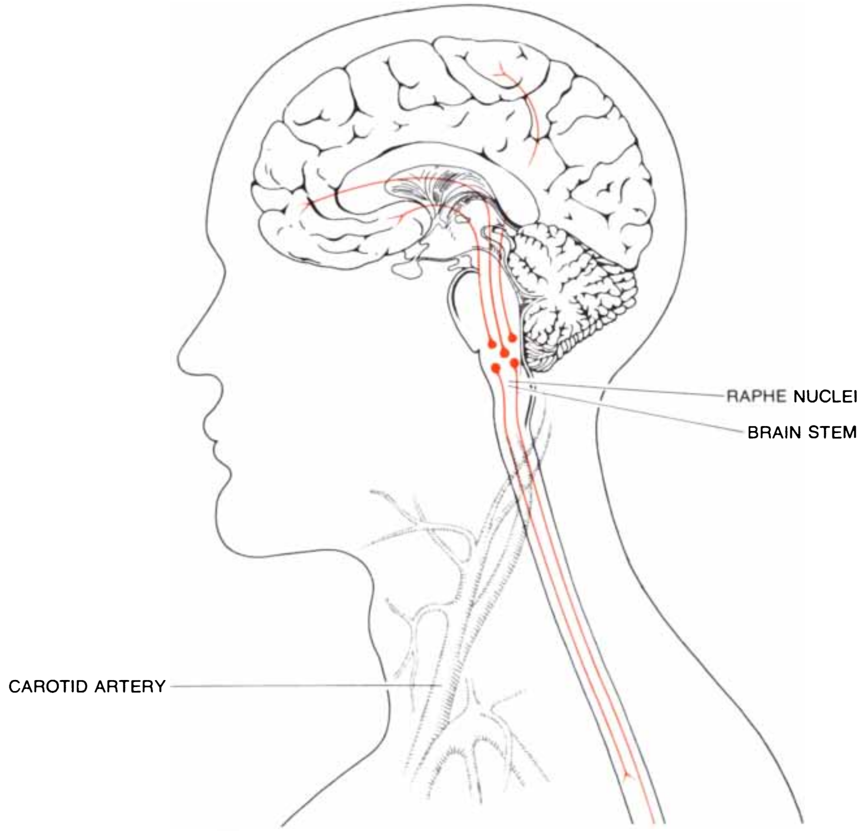
While the study was in progress we learned of a new chemical assay for serotonin that was far more sensitive than the methods that had been available. Using the new assay, we compared brain-serotonin and brain-tryptophan concentrations in groups of rats sacrificed as before. Because the level of serotonin in the brain would depend as much on its rate of utilization as on its rate of synthesis, we did not anticipate that the rhythms would be perfectly in phase. Indeed, they were not, but they were nevertheless quite close. A slight afternoon increase in the brain-tryptophan concentration was not matched by a corresponding change in the brain-serotonin concentration; the brain serotonin actually decreased. Thereafter, however, the two rhythms were generally in phase. (Being nocturnal feeders, rats have a daily rhythm that is opposite in phase to the rhythm in diurnal feeders such as man.)

We completed these rhythm studies in 1970 and were satisfied that the natural variations in rats' brain-tryptophan concentrations influenced the brain's synthesis of serotonin. Shortly thereafter we undertook a search of the scientific literature on factors affecting the synthesis of other neurotransmitters. Having worked with daily changes, we were somewhat surprised to find that, at least insofar as studies of the effect of precursor availability were concerned, most other experiments had dealt only with long-term changes. It was our view that, if availability of precursors was a significant physiological factor affecting neurotransmitter synthesis, then the rapid short-term changes in the level of precursor concentrations would be the most important ones. We decided to see if a rapid artificial elevation of the brain-tryptophan level in rats, produced by injection of the amino acid, would increase the brain-serotonin concentration.

Our first injections were large enough to immediately supply the rat with about half the amount of tryptophan in a day's food ration. Within an hour the tryptophan concentration in the rat's brain reached nine times its normal level and the serotonin concentration nearly doubled. Even when we reduced the amount of tryptophan injected to a tenth of the original dose, the levels of brain tryptophan and serotonin still rose.

Having demonstrated that even small increases in brain tryptophan cause a rise in the concentration of brain serotonin, we decided to see if a decrease in the level of brain tryptophan would have the effect of diminishing the brain's rate of serotonin synthesis. Injections of insulin are known to lower the concentrations of amino acids in the bloodstream of both mammals and reptiles, and so we set about testing the possibility of lowering the brain-serotonin level in rats by injecting them with insulin. Far from lowering the concentration of tryptophan in the bloodstream, the injection increased it. Two hours after the injection the concentration of blood tryptophan reached a maximum and the levels of brain tryptophan and brain serotonin were similarly elevated. Not only was the result of our experiment paradoxical; it also was the first experimental evidence that a hormone (in this instance insulin) can affect the synthesis of a neurotransmitter in the brain.

But was our analysis of the experiment correct? Perhaps the increase in brain serotonin was unrelated to the increase

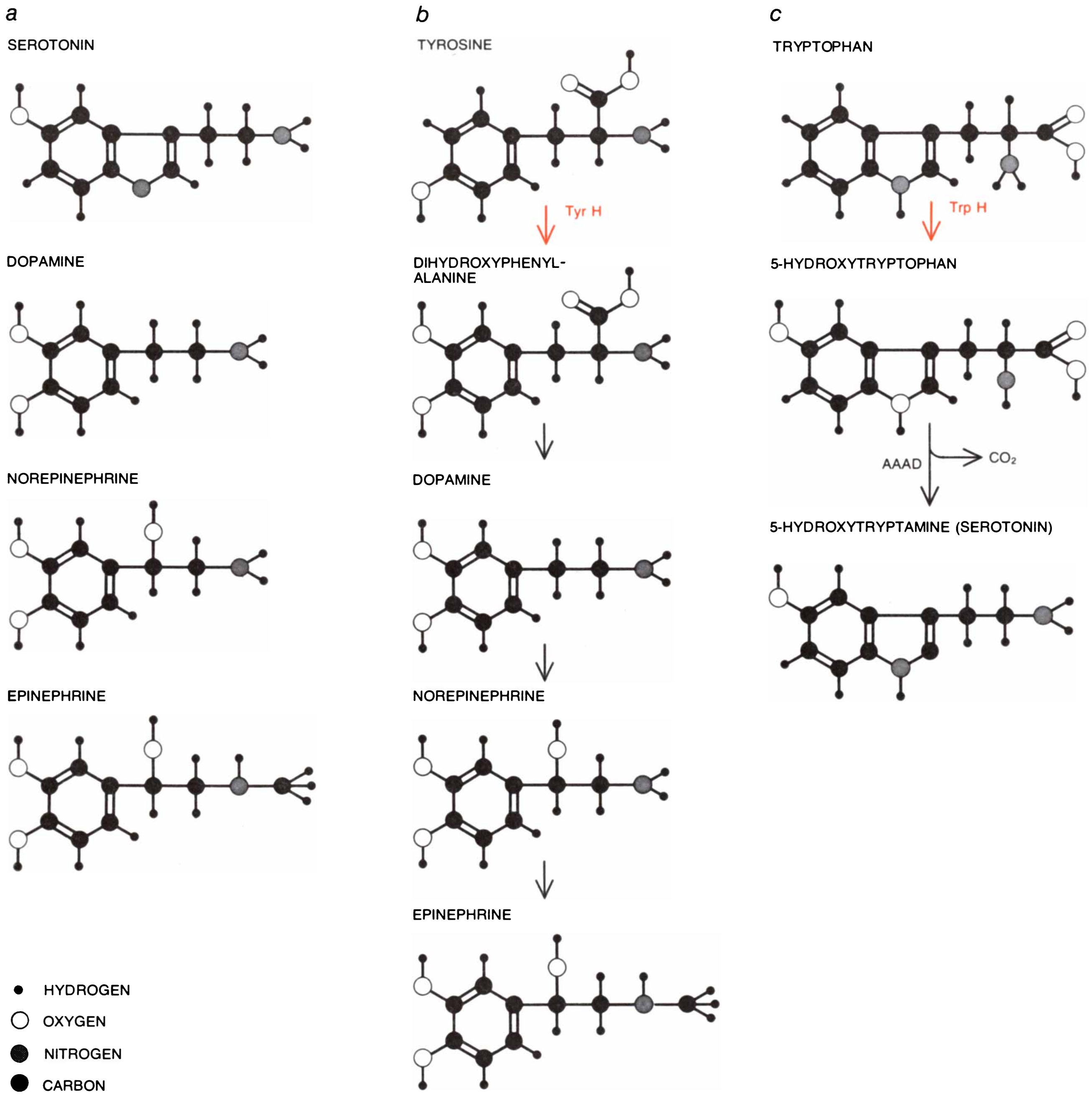


BRAIN NEURONS where most of the neurotransmitter serotonin is stored are a group known as the raphe nuclei. The neuron cell bodies are situated in the brain stem; the nerve fibers ascend into the rest of the brain and descend through the spinal column (*color*).

in brain tryptophan. It might be due to reflex changes in neuronal activity caused by the precipitous fall in the blood-sugar level induced by the insulin injection. To test this possibility we decided to let the rats introduce the insulin into their bloodstream themselves. We would allow the rats to eat a carbohydrate-rich diet, thereby stimulating the pancreas to secrete insulin naturally.

We fasted groups of rats overnight and then gave them a ration containing carbohydrate and fat but no protein. The natural secretion of insulin that followed produced the same effects that injections of insulin had: tryptophan levels in the blood rose significantly within an hour and reached a peak in two hours. During the first hour the level of brain tryptophan rose more than 20 percent and after two hours it reached a peak of 65 percent above normal. The concentration of brain serotonin increased in the first hour and was nearly 20 percent above normal after two hours [*see illustration on page 89*].

Reviewing our results, we reached two conclusions and ventured a prediction. First, the natural release of a hormone (insulin) in response to the consumption of carbohydrates elevated the level of a neurotransmitter (serotonin) in the brain of the rat; evidently it was the increase in the concentration of brain tryptophan that accelerated the brain's synthesis of serotonin. Further,



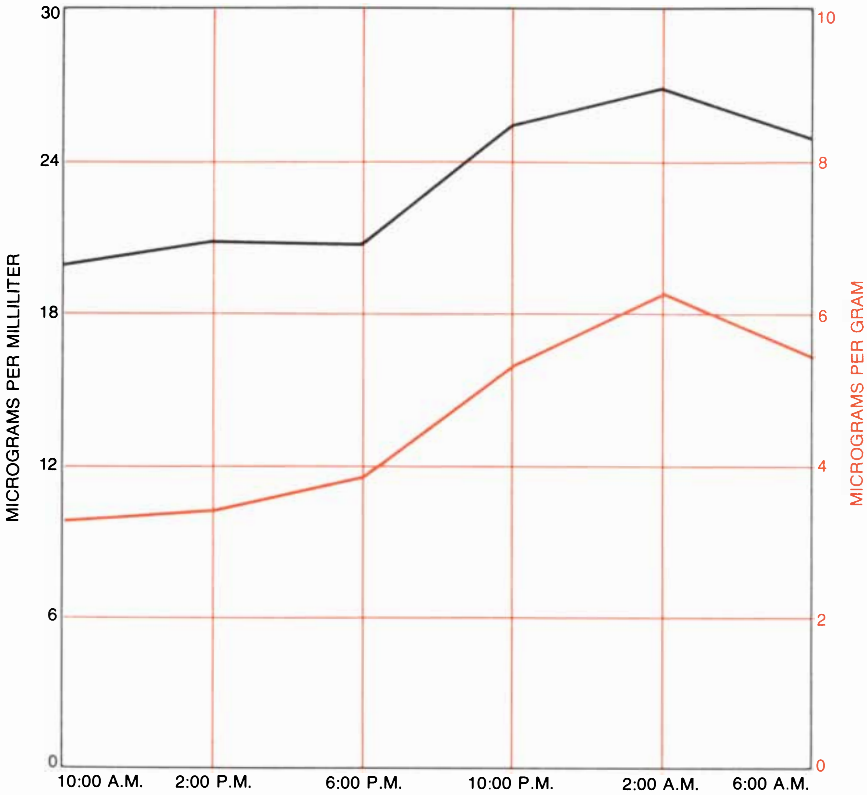
FOUR NEUROTRANSMITTERS found in the brain of mammals (*a, left*) are serotonin, dopamine, norepinephrine and epinephrine. The precursor of serotonin is the amino acid tryptophan (*c, right*); the first step in the synthesis is catalyzed (*colored arrow*) by the enzyme tryptophan hydroxylase. An "essential" amino acid, tryptophan is available only in dietary protein. The other neurotransmitters (*b, center*) are formed in successive steps from the same precursor amino acid, tyrosine; the initial reaction (*colored arrow*) is catalyzed by the enzyme tyrosine hydroxylase. Not an essential amino acid, tyrosine can be synthesized in the liver of mammals.

not only this "natural" insulin stimulus but also the artificial provision of insulin or of tryptophan by injection gave rise to identical effects: an increase in the level of tryptophan in the blood, paralleled by an increase in the levels of tryptophan and serotonin in the brain. On the basis of these conclusions we predicted that any increase in the level of blood tryptophan would be followed by an increase in the concentration of brain serotonin. For example, if we fed the rats a diet that included protein (the natural source of tryptophan) as well as carbohydrate and fat, we would expect the levels of both brain tryptophan and brain serotonin to increase in proportion to the increase in blood tryptophan.

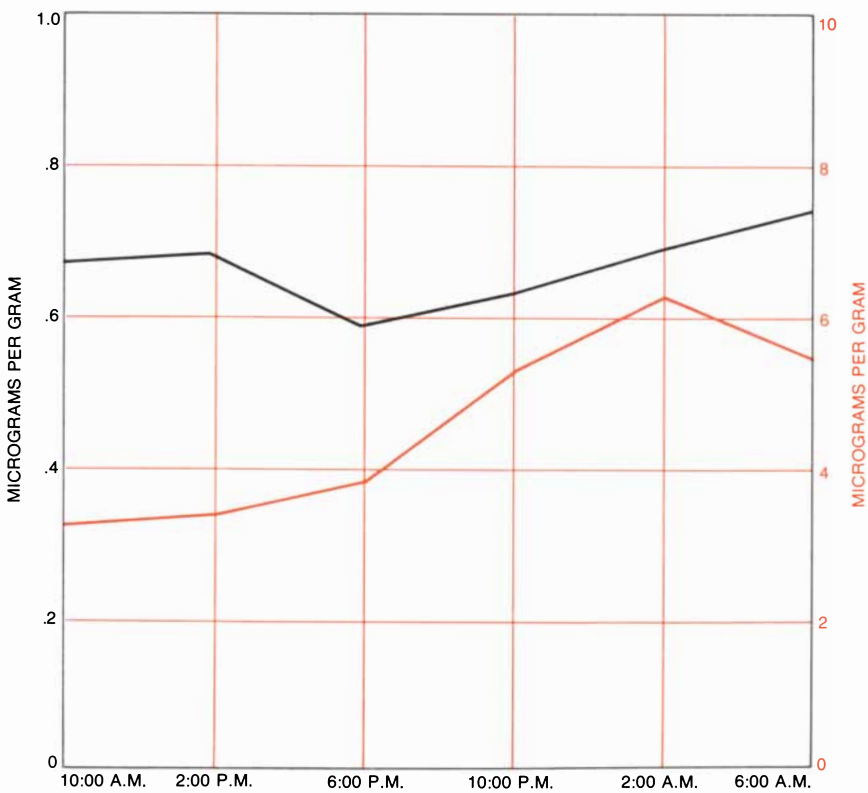
We tested our prediction by feeding special diets to two groups of rats. For one group we fortified a ration of carbohydrate and fat with natural protein by adding 18 percent of the milk protein casein. For the other group in place of the casein we substituted a complete mixture of amino acids. Only one pair of feedings was needed to show that our prediction was wrong. On either of the fortified diets the level of tryptophan in the rats' bloodstream increased but neither brain tryptophan nor brain serotonin was elevated at all!

We soon learned what was wrong with our prediction. As two groups of independent investigators had already established, amino acids are conveyed from the bloodstream to the brain by a carrier system that embraces several separate subsystems. For example, one carrier transfers only amino acids that are not ionized at normal blood acidity (pH 7). A second carrier transfers only those that are positively ionized and a third transfers only those that are negatively ionized. As it happens, tryptophan is not ionized at normal blood acidity. Once present in the bloodstream following the ingestion of protein, it competes with five similarly neutral amino acids (leucine, isoleucine, valine, phenylalanine and tyrosine) for transport. The levels of the five competing amino acids in the bloodstream of our rats rose sharply when they received a diet including protein. The rise in the level of tryptophan, however, was much smaller because tryptophan is far less abundant in protein than the other amino acids are. Evidently when the concentration of the competing amino acids became high enough, it impeded the passage of tryptophan into the brain by swamping the transport system.

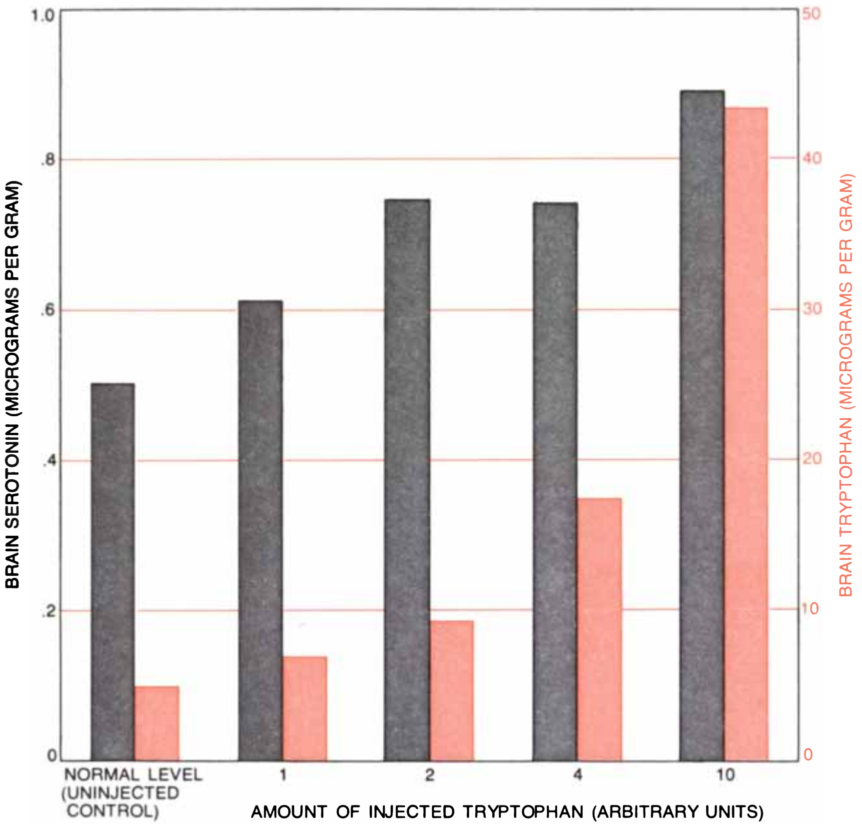
This sequence of events could also account for the action of injected and se-



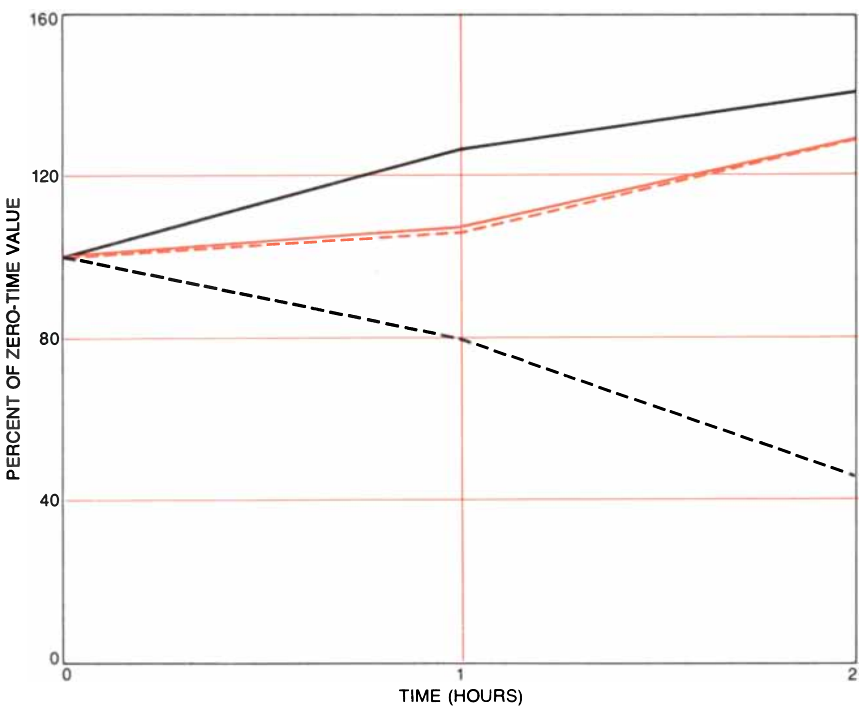
DAILY FLUCTUATION was noted in the level of tryptophan in the bloodstream of laboratory rats (*black curve*). The rhythm was correlated with the rats' time of food intake. The level of brain tryptophan also fluctuated (*colored curve*); the rhythms were in phase.



SEROTONIN CONCENTRATION in rats' brains also fluctuates (*black curve*). The rhythm reflects two interacting processes: accumulation of the neurotransmitter by tryptophan synthesis and its expenditure in neuron discharge. The rhythm is not quite in phase with the brain-tryptophan rhythm (*colored curve*), but a correlation between the two is clear.



BODY INJECTIONS of different amounts of tryptophan rapidly raised the level of brain tryptophan (*colored bars*) and brain serotonin (*gray bars*) in rats. In one hour a dose equivalent to half a day's normal intake in food (*far right*) raised the tryptophan level ninefold and serotonin by 40 percent compared with an uninjected rat (*far left*). As little as one-tenth that dose still raised both precursor and neurotransmitter levels significantly.



UNFORESEEN RESULT followed an attempt to lower the levels of brain tryptophan and brain serotonin in rats. In general the injection of insulin decreases the amount of amino acids normally found in the bloodstream, as it does the amount of blood sugar (*broken black line*). The injection instead increased the tryptophan level in both blood and brain (*solid black and solid colored line*) and also the brain-serotonin level (*broken colored line*).

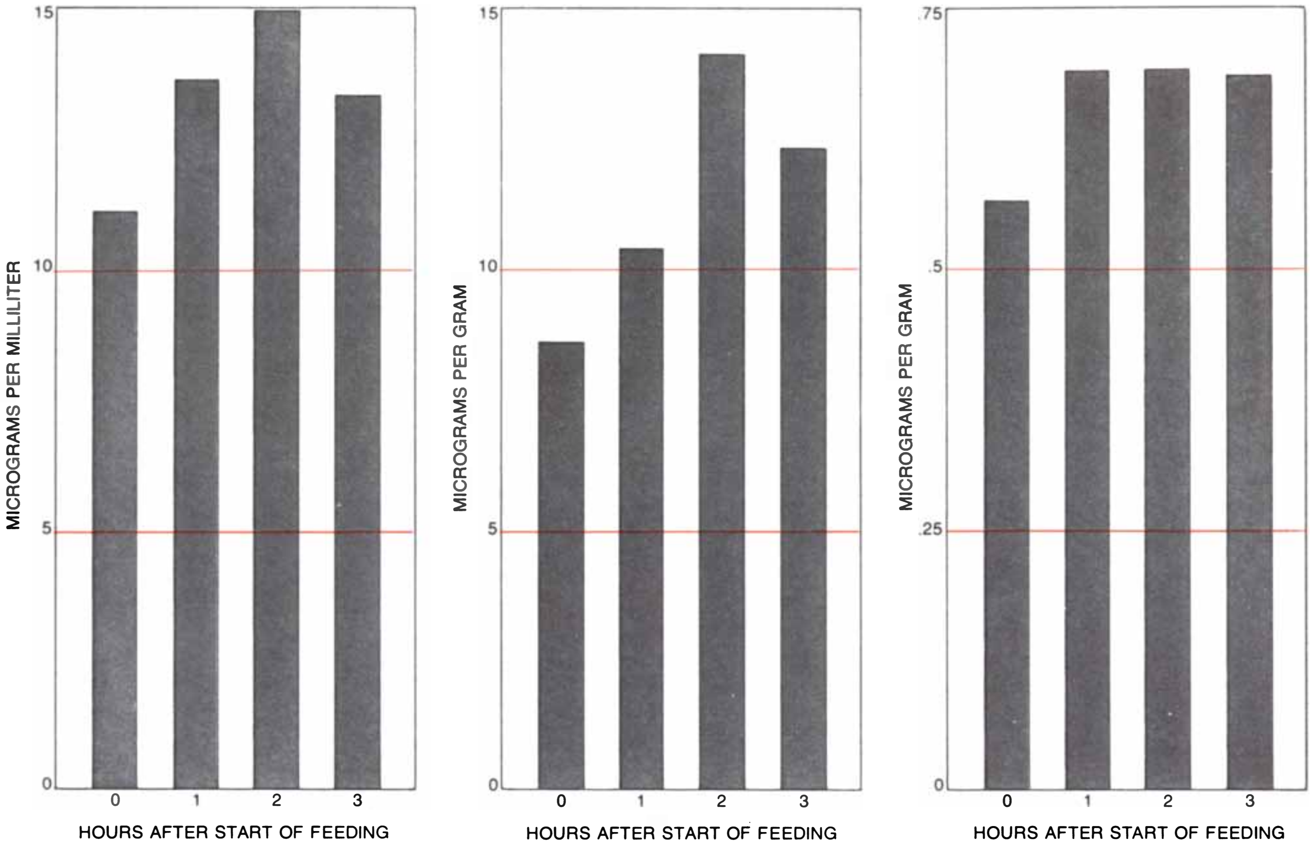
creted insulin on brain tryptophan. By lowering the bloodstream concentration of other neutral amino acids that compete with tryptophan for transport into the brain the insulin would facilitate a rise in the level of brain tryptophan. The simultaneous insulin-induced rise in the level of blood tryptophan would contribute further to the increase in brain tryptophan.

To test this interpretation we prepared two new diets to be fed to further groups of rats. One ration contained carbohydrate, fat and all the amino acids that would normally be present in an 18 percent casein diet. The second ration was identical except that the five neutral amino acid competitors of tryptophan were omitted. We found that the level of tryptophan in the bloodstream increased among the groups of rats fed either ration. Only among rats fed the ration without competing amino acids, however, did brain-tryptophan and brain-serotonin levels increase substantially [*see top illustration on page 90*].

As a further test we fed groups of rats a diet omitting two other amino acids, aspartate and glutamine, whose charge is such that they do not compete with tryptophan for transport. The blood-tryptophan level was elevated but neither brain tryptophan nor brain serotonin increased. It was clear that when it is relatively unimpeded by competitors, the transport of tryptophan from bloodstream to brain is accelerated.

We summarized our new conclusions as follows. The concentration of tryptophan in the brain (and the resulting synthesis of serotonin) reflects the ratio of tryptophan to other neutral amino acids in the bloodstream more accurately than it does the quantity of blood tryptophan alone. A carbohydrate-rich diet, by stimulating insulin secretion, will tend to raise that ratio; a protein-rich diet, by elevating the level of competitive amino acids in the bloodstream, will tend to lower it. We recently tested this conclusion further by feeding a group of rats a diet containing a very high concentration of protein: 40 percent casein, about the proportion of protein found in beefsteak. The rate of serotonin synthesis in the brain of the rats actually decreased.

Nonetheless, the paradox with respect to insulin was not entirely dispelled. Our measurements showed that in rats insulin acted in the bloodstream not only to elevate the ratio of tryptophan to the other amino acids but also to increase the amount of tryptophan present. How could this be? The search for an answer to the question proved to involve a di-



INSULIN PARADOX was investigated further by feeding a high-carbohydrate diet to hungry rats so that they would secrete their own insulin. From normal levels (*zero bar in each group*) the blood tryptophan (*left*), brain tryptophan (*center*) and brain serotonin (*right*) increased substantially in the first two hours after the rats were given access to food that induced insulin secretion.

etary component we had more or less ignored: fat.

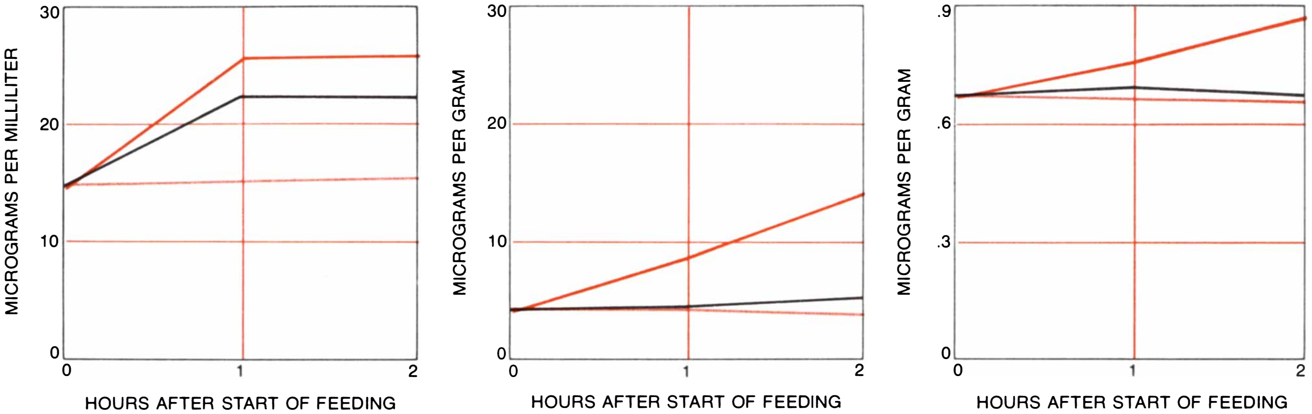
In common with a number of other substances that circulate in the bloodstream (bilirubin, certain hormones and free fatty acids), tryptophan has a tendency to bind to albumin, a protein that is normally present in the blood serum. Roughly 80 percent of all the tryptophan molecules in the bloodstream are bound in this way; the phenomenon was first demonstrated in 1958 by Rapier H. McMenamy and John L. Oncley at the Harvard Medical School. Ray W. Fuller of the Eli Lilly & Company Research Laboratories, who has followed our work and made many helpful suggestions, pointed out in 1971 that some of our observations might be the result of changes in the ratios of albumin-bound and free tryptophan in the subjects' bloodstream.

We discussed this possibility with our colleague Hamish Munro and spent the next year working with Bertha Madras and Deborah Lipsett in search of the most efficient way to separate free tryptophan in the blood serum from bound tryptophan. We then began an experiment with human volunteers aimed at testing Fuller's suggestion. After an overnight fast each volunteer drank 75 grams of glucose; blood samples were then taken at regular intervals. As we expected, the ingestion of carbohydrate caused the volunteers to secrete insulin, and the levels of all the amino acids in their blood except tryptophan fell markedly. (Unlike the response in rats, the human response to insulin does not include an actual increase in the amount of tryptophan in the bloodstream.) Although the total amount of the volunteers' blood tryptophan remained the same, the amount of free tryptophan in the blood samples fell by as much as 40 percent. The decrease was matched by a proportional, albeit smaller, increase in the amount of bound blood tryptophan (which of course had originally made up the greater part of the total). What was causing the change?

Like the concentration of amino acids in the blood, the concentration of free fatty acids is known to be lowered by the presence of insulin in the bloodstream. It seemed possible that, if the insulin forced some molecules of free fatty acid to give up their bonds with molecules of albumin and leave the bloodstream, the increased availability of unbound albumin might allow more molecules of tryptophan to shift from the free state to the albumin-bound one.

To test this conjecture we measured the concentration of free fatty acids in the blood serum of our volunteers. The concentration changed in a pattern that was strikingly like the change in the concentration of free tryptophan. After the ingestion of glucose both concentrations fell within a similar length of time and in similar proportions. When we repeated the experiment with rats instead of human volunteers, the result was the same.

This led us to the conclusion that the free fatty acids in the blood serum are intermediaries in the effect of insulin on tryptophan circulating in the bloodstream. On this hypothesis tryptophan does not decrease in the bloodstream in the presence of insulin as the other neutral amino acids do because the action of insulin simultaneously frees albumin from existing fatty-acid bonds. The freeing of albumin, in turn, allows additional molecules of free tryptophan to become albumin-bound. As the concentration of free tryptophan in the blood falls, it is partially replenished by additional molecules of tryptophan that diffuse out of other tissues, such as skeletal



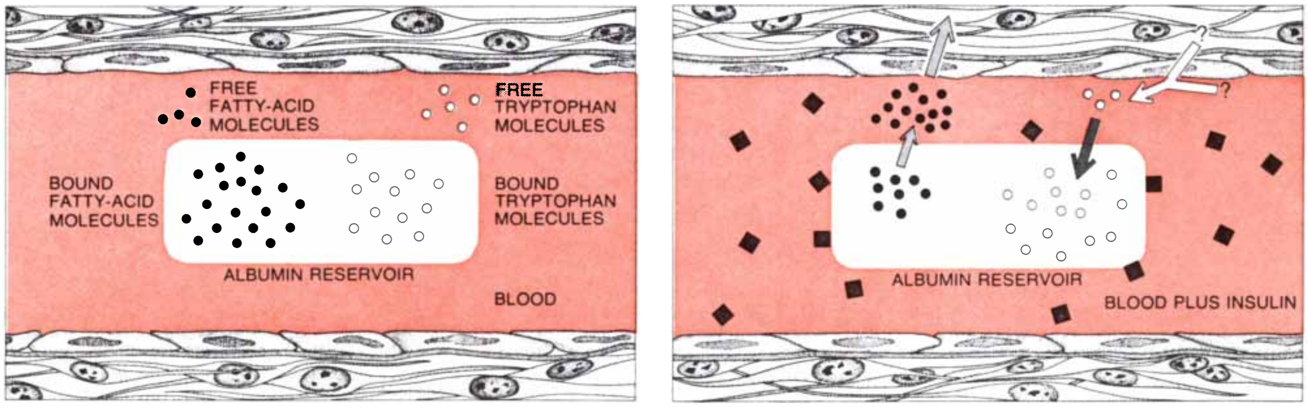
COMPETITION BETWEEN TRYPTOPHAN and five other amino acids for "transportation" to the brain accounts for the unchanged level of tryptophan in the brain after ingestion of a tryptophan-rich diet. Two groups of rats were fed contrasting rations; in one (*black*) all the amino acids found in dietary protein (including tryptophan and its five competitors) were present. The five competing amino acids were omitted from the other ration (*dark color*) so that the level of the five competitors would be lowered in the blood of rats fed it. Eating either ration raised the rats' level of blood tryptophan above the level of a fasting control group (*left, light color*). Brain tryptophan (*center*) and serotonin (*right*) levels rose far more, however, in the rats fed no competing amino acids.

muscle. Evidently in rats the affinity of albumin for tryptophan in the presence of insulin is so high that the total concentration of tryptophan in the bloodstream rises. In humans, however, the affinity appears to be lower. As a result the increase in the amount of albumin-bound tryptophan does not depress the level of free tryptophan in the blood to the point where additional tryptophan molecules diffuse from the tissue into the blood. Thus in humans insulin does not give rise to a net increase in the blood-tryptophan level.

The existence of such an albumin-bound and insulin-immune reservoir of tryptophan in the bloodstream gives this amino acid a competitive advantage over other neutral amino acids with respect to transport from the bloodstream into the brain. Our hypothesis awaits testing at least in one regard. It would be a useful experiment to feed subjects a high-fat diet in order to see whether an overabundance of free fatty acids locks up enough extra albumin in the bloodstream to limit or suppress the binding of tryptophan to albumin and thus shrink the size of the insulin-immune reservoir of bound tryptophan.

We are satisfied that the consumption of foods low in protein produces an increase in the synthesis of serotonin in rats' brains, although the reason has proved to be a quite complex one. The same assay methods cannot be used with human volunteers, so that we have no tangible proof that the same is true of man. Nonetheless, by feeding our volunteers protein-free and protein-rich substances and then measuring the amount of tryptophan and other neutral amino acids in their blood, we have established that humans respond exactly as rats do to the ingestion of the same diets. It seems logical to assume that the changes in the level of human brain tryptophan are similarly reflected in the level of brain serotonin.

We have not yet found direct evidence that the increased concentration of this particular neurotransmitter in the neurons of the raphe nuclei necessarily induces changes in the functional activity of the neurons. Intuition, however, tempts us to this conclusion



FURTHER EXPLANATION of the insulin paradox involves competition between tryptophan molecules and molecules of free fatty acid in the bloodstream for binding to available molecules of albumin. Some 80 percent of all the tryptophan in the blood is linked to albumin in this way; in the illustration this is schematically represented by the areas labeled "Albumin reservoir." When insulin enters the bloodstream (*right, black rectangles*), the bound fatty acids are reduced in number and many of them are driven from the blood (*gray arrows*). This allows more free tryptophan molecules to bind to the released albumin molecules (*black arrow*); the shift into the insulin-immune reservoir helps to keep the total blood-tryptophan level in rough equilibrium. In the rat still other tryptophan molecules migrate from red blood cells or adjacent tissue (*double arrow at right*), thus raising the tryptophan level.

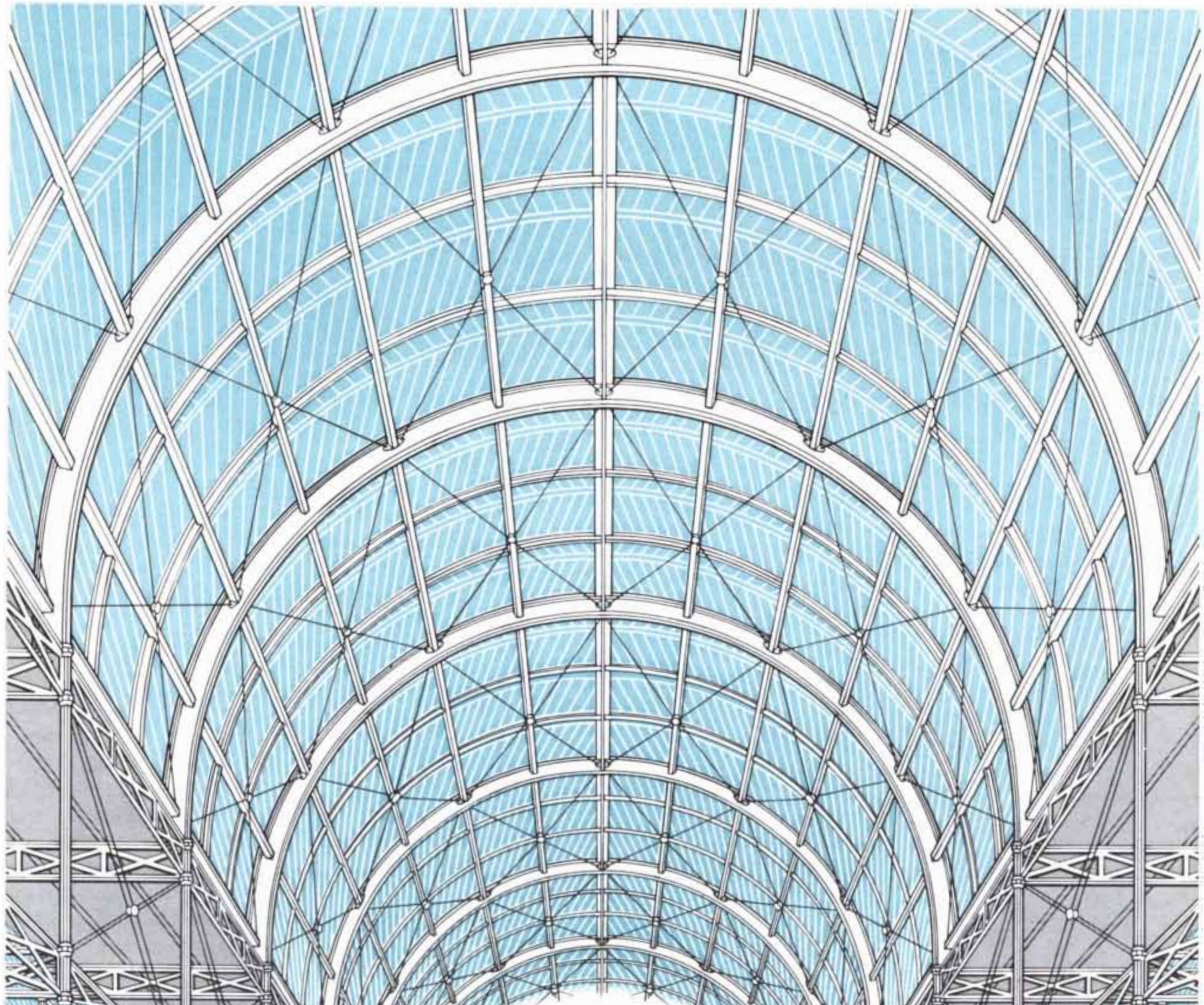
and leads us to suspect that the susceptibility of brain-serotonin concentrations to diet-induced changes probably provides mammals, man included, with a useful transduction mechanism.

It seems reasonable to speculate that diet-induced changes in the concentration of serotonin in the raphe-nuclei neurons would be paralleled by changes in the amount of serotonin released into the synapses whenever the neurons elect to transmit signals or are forced to do so through chemical intervention. If this is the case, then the serotoninergic neurons would function as sensors, converting a signal that is circulating in the bloodstream into a statement in brain language. A regular protein-rich diet would result in a minimum release of serotonin by the raphe-nuclei neurons, whereas even a brief period of a carbohydrate-rich diet would provide for a maximum release of the neurotransmitter.

It does not seem unreasonable that the brain could make good use of such responses to the protein-to-carbohydrate proportions of the diet, and perhaps to other nutritional and hormonal inputs that were sensed in the same way. The serotoninergic neurons may convey to the rest of the brain important data not only about the metabolic consequences of short-term and long-term changes in nutrition but also about growth and development processes and about stress.

The serotoninergic neurons of some mammals are known to be associated with a number of neuroendocrine mechanisms. For example, cats deprived of brain serotonin become insomniac, and the reduction of the brain-serotonin level in male and female rats gives rise to exaggerated sexual activity. Conversely, an increase in the brain-serotonin level decreases a rat's sensitivity to pain. The secretion of a number of pituitary hormones is also correlated with the level of brain serotonin. Recent work we have done with our colleague Loy Lytle suggests that changes in the brain-serotonin level of rats can influence both motor activity and food consumption. The latter finding suggests a kind of closed circle, with food consumption affecting brain biochemistry and brain biochemistry in turn affecting food consumption.

We may eventually find that this particular closed circle is a spiral viewed end on, or even a curved line that leads nowhere. Yet it remains a fact that, at least in our mammalian relative the laboratory rat, diet does control the synthesis of a significant neurotransmitter. We are loath to dismiss such a finding as mere coincidence.



TRELLISWORK OF GLASS AND IRON that composed the Crystal Palace incorporated two forms of wind bracing later adopted for the construction of skyscrapers. The first type consisted of double diagonal rods installed in the arched roof and between pairs of columns. The second and more innovative form appeared where the columns met the trusses that supported the galleries at the sides of the hall. Collars fixed to the columns were bolted to the top and to the bottom of the trusses, the two points of attachment making the joint rigid. The Crystal Palace was built for the London Exhibition of 1851; here a transverse section of the main vault is shown.

The Wind Bracing of Buildings

The evolution of the skyscraper has inspired elaborate methods for providing resistance to the wind. In the largest buildings a technique invented more than 100 years ago has been revived

by Carl W. Condit

A building must ordinarily sustain three kinds of force. The first is the "dead load" of the structure itself and its fixed contents; the second is the "live load" of the building's occupants and of such movable components as elevators. Both of these forces are gravitational and must ultimately be directed along vertical lines through the foundation to the earth.

The third force is produced by the movement of air and acts in a direction perpendicular to the walls of the building. As a result its effects are different from those of the first two forces: gravity loads subject vertical bearing members such as columns or piers mainly to compression; wind loads, on the other hand, produce in the structural framework complex stresses characteristic of bending, overturning, twisting and shearing.

When large buildings were of ponderous masonry construction, the effects of wind were relatively unimportant in structural design. A building capable of bearing the enormous dead load of thick stone walls is intrinsically invulnerable to wind damage. When builders began erecting exceptionally tall buildings supported by a light iron or steel frame, however, the problem of providing adequate resistance to the wind proved difficult to solve. Earlier forms of construction gave only limited guidance. In fact, many of the devices and techniques adopted by the builders of skyscrapers are derived not from earlier building practice but from methods devised for the construction of long-span bridges.

The structural integrity of a building depends on two properties. If the structure as a whole remains immobile with respect to the earth under the loads for which it is designed, it is said to possess stability. If the various elements of which it is composed remain immobile with respect to one another, it is said to possess rigidity.

A simple rectangular frame is stable under a fixed vertical load (if the individual members are strong enough to bear the load). The frame is not rigid, however, when it is subjected to a horizontal force or to a force that can be resolved into vertical and horizontal components. When the rectangle is made to bear such a load, it is distorted, becomes a parallelogram and eventually is flattened entirely [*see top illustration on next page*].

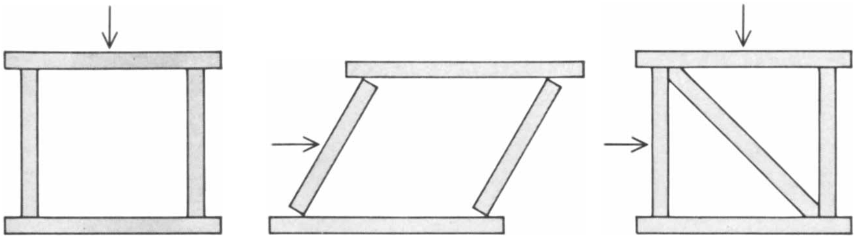
This weakness can be corrected by placing a brace diagonally across the rectangle, thereby transforming the frame into a rigid structure by creating two triangles. The triangle is a rigid figure, a figure that cannot be deformed without distorting one or more of its sides. With this brace the framework is both stable and rigid. If its members are strong enough and if its joints are reasonably secure, it will sustain both vertical and horizontal forces. Its only defect is that the diagonal member extending across the entire frame may render the space inside it useless. Other means of building rigid structures had to be found, and their invention constitutes a large part of modern structural history.

The medieval carpenter was well aware of the need to brace timber frames against wind loads. A pitched roof supported by a triangular roof truss is rigid in itself, but the truss and the columns that support it form a rectangle that is not. A diagonal cannot be placed across this rectangle because it would obstruct the entire bay. (A bay is the space enclosed by a single unit of the frame as defined by any pair of adjacent columns or posts.) The simplest solution to this problem was to introduce a small diagonal member called a knee brace between the truss and the column. The brace made the connection rigid and obstructed only the space near the roof. A stronger and more elegant device was the curving brace fixed by mortise-and-tenon joints to the truss and the column. These braces were called ship's knees because they were adapted from the curving knee braces that connect the deck beams and hull ribs of sailing vessels. Similar braces were also installed parallel to the long axis of the building [*see bottom illustration on next page*].

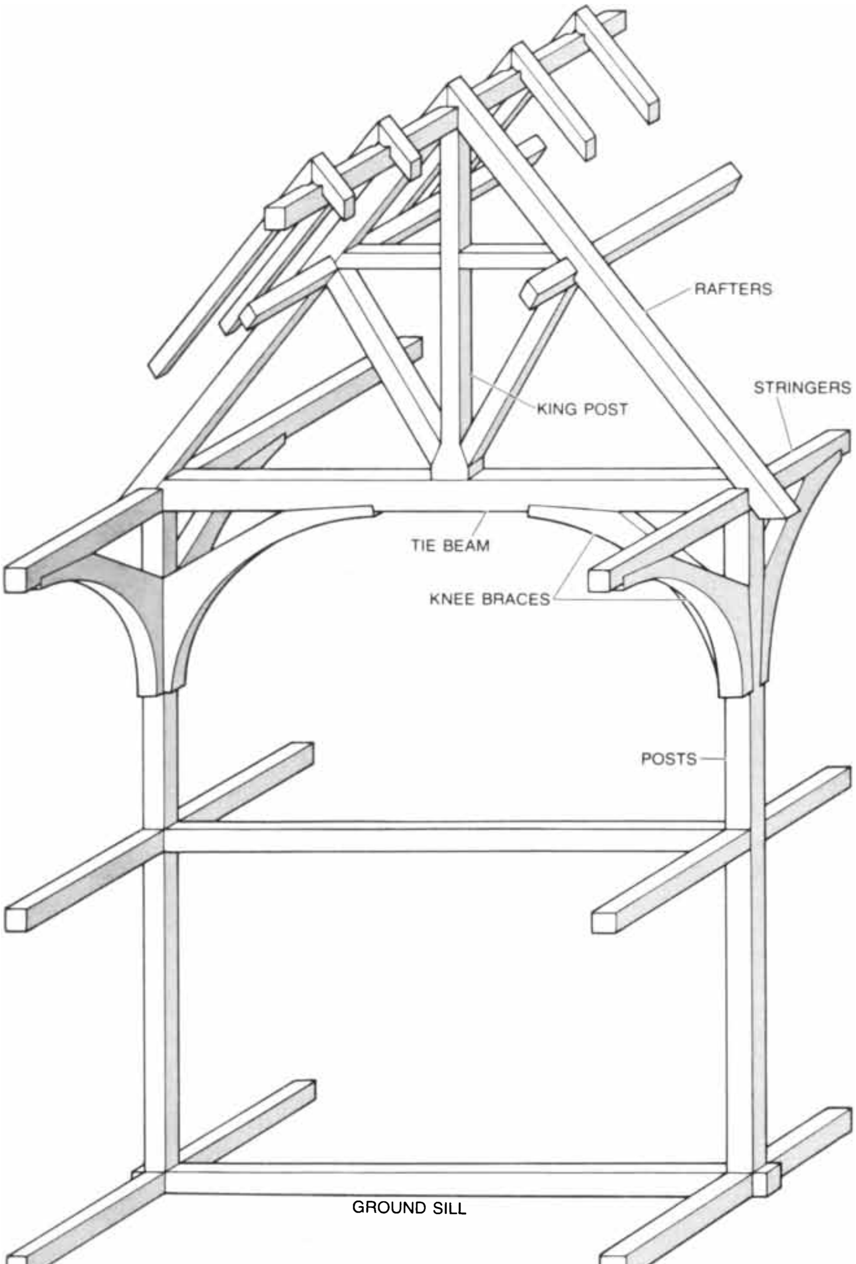
For large buildings made of masonry wind loads did not become a problem until the 12th century. In earlier periods great masses of stone had been used to sustain the dead load of the structure; against such thick and heavy walls wind is ineffectual. The Gothic builders, however, transformed the heavy Romanesque church into a highly attenuated structure of ribs, columns and buttresses that was almost a system of framing in stone. For this relatively insubstantial structure wind resistance was an important consideration [see "The Structural Analysis of Gothic Cathedrals," by Robert Mark; SCIENTIFIC AMERICAN, November, 1972].

During the Renaissance builders returned to the solid masonry wall of classical precedent for buildings of monumental scale, and the problems of providing wind resistance were once again irrelevant. For smaller buildings made of timber the solutions devised by medieval builders were satisfactory, and they remained standard until well into the 19th century.

With the Industrial Revolution of the 18th century a more innovative spirit arose, inspired by a complex of interrelated factors that accompanied the introduction of iron as a primary structural material. This new application of iron eventually led to two subsidiary developments in the building arts. One was growth in the size of framed structures,



PROPERTIES OF STABILITY AND RIGIDITY are demonstrated by a rectangular frame. Under a vertical load (*left*) the frame is stable; it remains immobile with respect to the ground. The frame is not rigid, however, when subjected to a horizontal force (*center*); its members do not remain immobile with respect to one another. Rigidity is achieved by bracing the frame diagonally (*right*), which transforms the rectangle into two triangles.



TIMBER FRAME OF THE MEDIEVAL HOUSE was among the earliest structures to require bracing against wind loads. The roof truss formed by the rafters and the tie beam is a triangle and is intrinsically rigid; the tie beam and the remainder of the frame form a rectangle, however, that is not rigid. Because a diagonal member placed across the entire frame would have obstructed the interior, smaller diagonal elements called knee braces were employed between the posts and the tie beam and between the posts and the stringers.

paralleled by a decrease in the cross-sectional area of supporting members. The other was the transformation of building from a pragmatic and empirical craft into a technology predicated on theoretical and experimental science.

The changes came slowly at first. The appearance of multistory iron-framed factories and warehouses as late as 1850, for example, was not appreciably different from that of the traditional timber-framed buildings that were their prototypes.

The earliest buildings in which iron was used for primary framing members were cotton mills erected to house the new power-driven textile machinery. Their designers followed traditional practice in supporting the structure with massive brick wall piers, but they replaced wood with iron in the interior columns and beams. Because of the heavy load of masonry, interior bracing was considered unnecessary, a view that persisted until the last decade of the 19th century.

It was the freestanding iron frame that brought the need for wind bracing to the attention of builders, since these structures lacked the protective masonry envelope. The first iron roof truss was introduced in 1810, and by 1820 builders had begun to replace the masonry wall with iron columns. An early example was a small market in Paris called the Madeleine, built in 1824, for which the frame of columns and roof trusses was entirely unbraced. A more advanced work is the Hungerford Fish Market in London, designed by the architect Charles Fowler and erected in 1835. The roof frame is supported by cylindrical cast-iron columns. Bracing is provided by iron ship's knees curving outward from collars fixed to the columns.

The application of science to problems of structural design, more accurate methods of analyzing stresses in framed structures and mills capable of rolling wrought-iron beams up to seven inches deep in cross section were the basis for rapid advances in iron construction during the remainder of the 19th century. The most prophetic work was the Crystal Palace, erected in 1851 for the London Exhibition [*see illustration on page 92*]. It was the creation primarily of Sir Joseph Paxton, a horticulturist and builder of iron-framed conservatories.

The Crystal Palace was the first great iron-framed building; its ground plan measured 408 by 1,848 feet. It was also the first building enclosed by curtain walls of glass, the first for which the structural units were prefabricated and the first in which a light frame was made rigid against wind loads by the technique that came to be known as portal bracing.

Two kinds of bracing were employed in the Crystal Palace. Double diagonals consisting of wrought-iron rods arranged in a series of *X*'s connected the major bearing members of the central vault. The second form of bracing appeared where the cast-iron trusses that supported the galleries met the octagonal columns. Flanged collars fixed to the columns were bolted to both the top and the bottom of the trusses, the two points of attachment eliminating the need for diagonal members. Both of these techniques evolved from methods used in the construction of truss bridges. Double diagonals traditionally brace the top and bottom frames of these bridges. The second type of brace used in the Crystal Palace is found in the end frames of the bridges. These are the portal (or entrance) frames, and the bracing is called portal bracing [*see illustration below*].

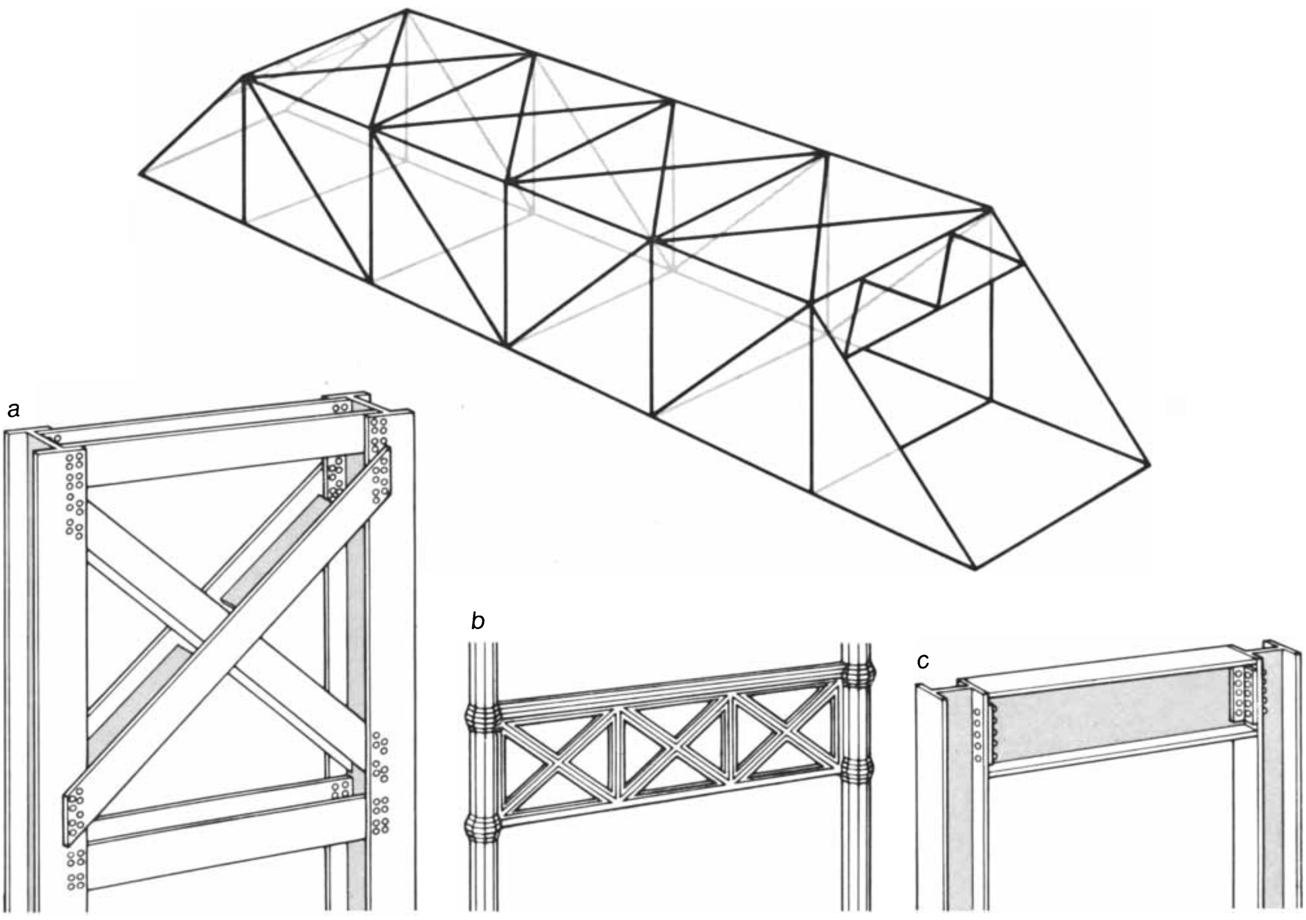
The principle that underlies this technique is at least as old as hammer-and-nails carpentry. If two boards are joined with a single nail, they will easily pivot around the nail; if two nails are used, the joint is made rigid. In the Crystal Palace the method was probably adopted for aesthetic reasons, since knee bracing would have seemed ungainly in this elegant frame.

The portal brace in what is essentially its modern form appeared a few years later. The earliest example is in the Royal Navy Boat Store at Sheerness in England, designed by Godfrey T. Greene and built between 1858 and 1860. Deep wrought-iron girders were riveted to cast-iron columns throughout the depth of the girders. Further developments of the technique were introduced with the skyscrapers that began to go up in New York and Chicago at the end of the century.

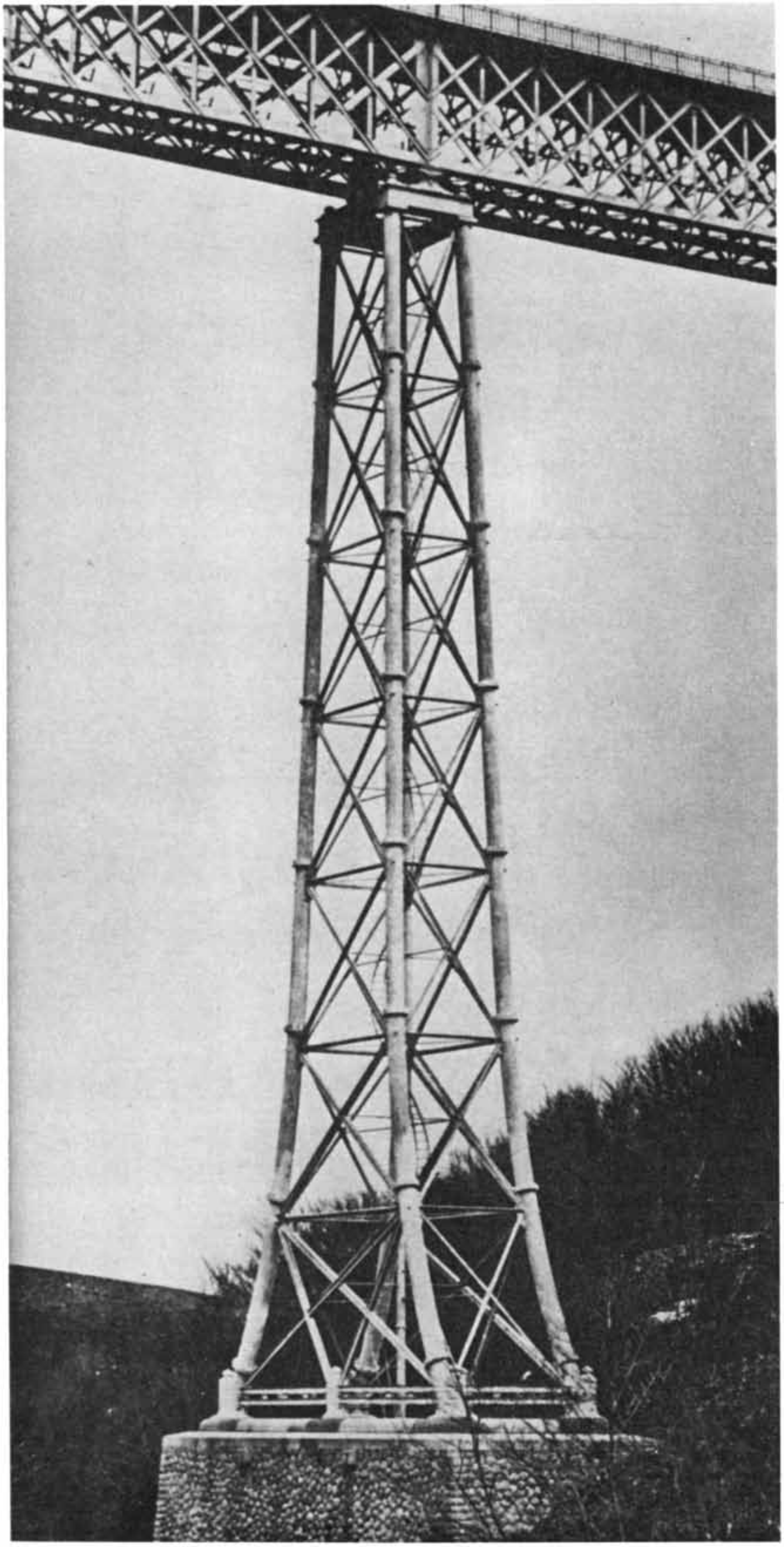
In the same period several builders introduced other fundamental techniques of modern high-rise construction. Although the developments seemed unrelated, a pattern of technological evolution began to emerge.

Several innovations first appeared in the construction of lighthouses. The form of maximum stability in the wind for such a structure is the "tree trunk" profile, a tapering cylinder that flares outward at the base. This form had been used by John Smeaton in 1759 for the masonry tower of the Eddystone lighthouse in England. It was adapted to iron construction by French engineers and came to maturity in the light at Roches-Douvres on the English Channel, built in the 1860's.

This lighthouse, 48 meters high, had an external covering of iron plates, supported by a wrought-iron frame of columns, peripheral beams and diagonal braces. Resistance to the wind was attained through three devices: the familiar diagonal members in the bottom three stages of the tower, where bending moments induced by the wind are greatest; gusset plates connecting column, beam



DERIVATION OF TWO BRACING TECHNIQUES is from the truss bridge. Double diagonal bracing, customary in the top and bottom panels of the bridge, has been adapted directly to buildings where its obstruction of interior space is not objectionable (*a*). Portal bracing derives from the end, or portal, frames of the bridge. In the earliest examples, such as the Crystal Palace, a truss was attached to a column at two points (*b*). In later buildings columns and girders are riveted throughout the depth of the girder (*c*).



OUTWARD-FLARING POSTS reduce stresses in a bridge "bent," or supporting tower, on the Bellon Viaduct of the Gannat-Commentry railroad in France. Bending and shearing forces are greatest at the base of the ironwork; the curvature of the posts tends to transform these forces into compression, which the cast-iron members are better able to resist. A consultant to the designers of the viaduct, built between 1868 and 1871, was Gustave Eiffel.

and diagonal in a riveted joint, and the wide base of the tree-trunk form.

In the same decade the engineers of the Gannat-Commentry railroad in France developed frames shaped like tree trunks for the iron "bents," or supporting towers, of the numerous bridges on the mountainous section of the railroad. The most advanced of these structures were designed for the Bellon Viaduct, built between 1868 and 1871 [*see illustration at left*].

The bents are made of four tubular cast-iron posts tied together at intervals by horizontal members and braced by double-diagonal rods in all the vertical and horizontal panels defined by the ties and posts. The unusual feature of the structure is the precisely calculated outward flare of the posts at their lower end. The flare not only provides the greatest depth of structure where the bending moments are greatest but also acts to transform shear forces at the base into compressive forces in the stout cast-iron tubes. The Gannat-Commentry bridges were designed by the engineers Delom, Geoffroy, Nordling and Thirion. A far more illustrious engineer was a consultant on the project: Gustave Eiffel. By 1870 Eiffel was the leading European authority on the aerodynamic stability of high structures.

The celebrated tower Eiffel designed for the Paris Exposition of 1889 is the supreme example of an edifice built strictly according to the dictates of structural and aerodynamic science. The principles of aerodynamic construction it embodied still guide the design of steel-framed skyscrapers.

This immense work of wrought iron, an extraordinary tour de force that exists solely as a monument to itself, was erected between 1887 and 1889 from designs prepared by Eiffel, assisted by the engineers Emile Nouguier and Maurice Koechlin and the architect Stephen Sauvestre. The tower was to be 300 meters high, divided into three stages, the first to house a restaurant and the second and third to serve as observation galleries.

The construction of long-span bridges by the 1880's had become a predictable science, so that the calculation of gravitational forces rested on extensive precedents. Eiffel knew that wind would be the crucial factor, on the resistance to which his tower would stand or fall.

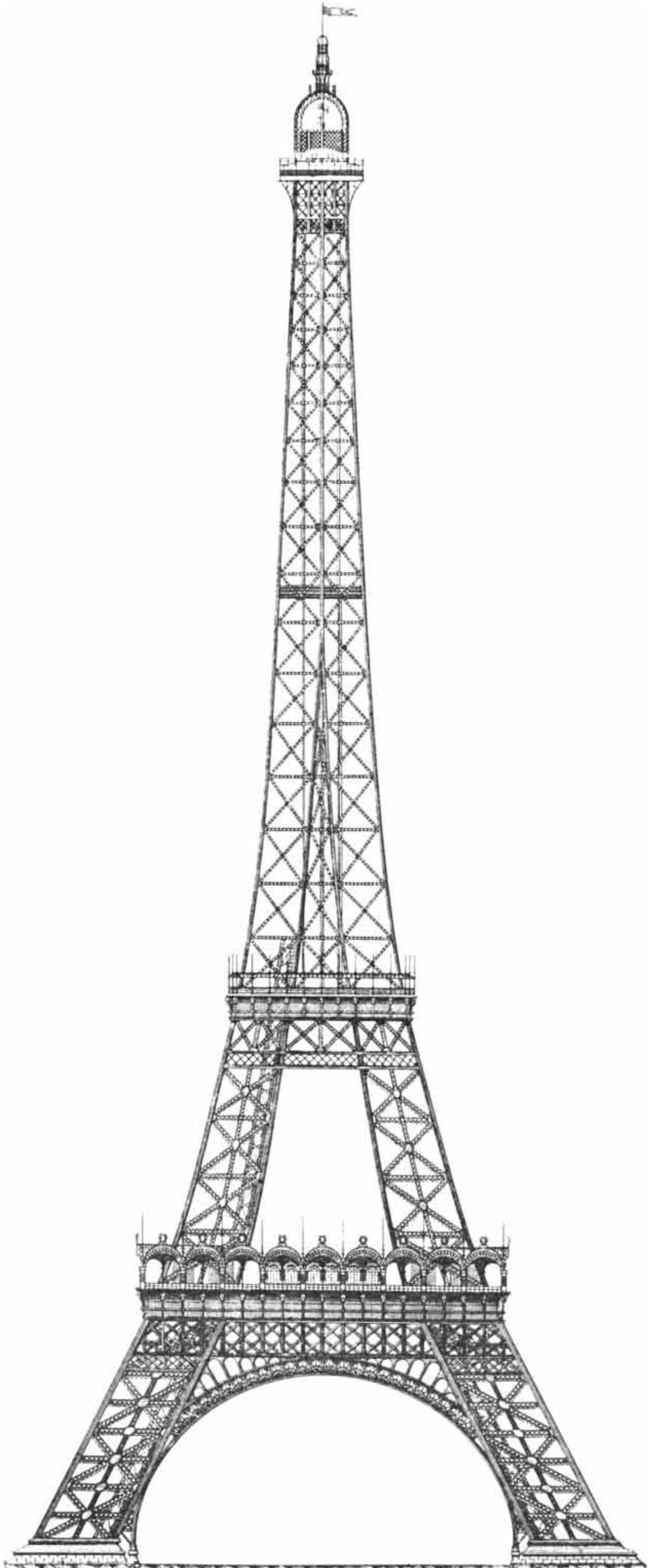
His initial decision was to treat the entire structure as a vertical cantilever. (A cantilever is a beam or truss supported only at one end.) The cantilever consists of two parts: one is the tower proper, extending from the second stage to the floor of the present meteorological and television station at the top; the other is visually distinct and consists of the base framework, which Eiffel called the *montant*, or upright [*see illustration at right*]. The transitional device is a peripheral truss that acts as a huge tie immediately below the second stage.

The tower proper is what today would be called a "space frame," a three-dimensional framework enclosing a defined space, in which all the members are connected and act as a single entity. It lies within a curvilinear tapering envelope that gracefully stretches out the strong curves of the base. The individual bents in both tower and *montant* are derived directly from bridge supports: each consists of four posts connected by horizontal struts and double diagonal braces.

Eiffel considered two alternatives in the calculation of wind loads, one a uniform load from top to bottom of 300 kilograms per square meter, the other a uniformly decreasing load ranging from 400 kilograms per square meter at the top to 200 kilograms at grade.

The resultant of the uniform wind load (the single vector into which the wind forces can be resolved) acts near the midpoint of the structure and yields a maximum bending moment at the base of the tower proper of 50,923,218 kilogram-meters and a maximum shearing force of 66,960 kilograms. For the varying wind load the bending moment is 59,745,250 kilogram-meters and the shearing force is 73,050 kilograms. If the supports of the tower had extended straight to the ground, the bending moment at grade would have been about 225 million kilogram-meters. Eiffel drastically reduced this enormous overturning force by directing it through the extreme curve of the four bents that compose the base frame.

The striking form of the *montant*, the most novel idea in the design, was chosen after rigorous analysis. The ancestor was the outward-flaring form of the



A PRECURSOR of the modern skyscraper, the Eiffel Tower was considered by its designer to be a vertical cantilever, that is, a monolithic structure anchored at one end but free to sway at the other. Its form, that of a curved, tapering spire, reduces the surface area exposed to the wind near the top and converts bending and shearing forces into compression at the base. Built for the Paris Exposition of 1889, the tower is 300 meters high. The drawing is from Eiffel's treatise, "La Tour de Trois Cents Mètres."

lighthouse and the iron bridge bent. Eiffel and his assistants gave this form a scientific basis.

The curve of each of the four bents conforms as closely as rigid metal construction allows to the curve tracing the diminution of the bending moment with height. The curve of the bents approximates a segment of a parabola; it is the same curve as is obtained by plotting the inverse relation of the bending moment to elevation. The axis also corresponds closely to the pressure line along which the dead load of the structure is transmitted [*see illustration below*]. The practical consequence of this design is that the enormous bending and shearing forces in the *montant* are progressively transformed into compression, the kind of force the rigid, hollow tubular bents are ideally suited to resist. Finally, the form reduces oscillations, although they are still considerable in the slender tower.

A precedent for the Paris tower was the Statue of Liberty, erected between 1883 and 1886 on what was then Bedloe's Island in New York harbor. The statue is the work of Frédéric Auguste Bartholdi, a sculptor, but the internal steel frame that supports it was designed by Eiffel. A hollow copper envelope standing 155 feet five inches above its masonry pedestal, higher than any building of the time and exceeded in height only by a few bridge supports, the statue posed a special problem in wind bracing. The central structure consists of four posts tied by horizontal struts and rendered rigid by double diagonals on all four faces. This tower carries a surrounding space frame to which the copper panels are attached. The arrangement of the diagonals and the trusswork indicate that Eiffel thought of the main figure as a vertical cantilever and the upraised arm as an appended eccentric cantilever.

It was 1960 before Eiffel's approach to the design of tall structures was applied to the construction of skyscrapers. It was adopted then because the high, towerlike building proved more vulnerable to wind than any previous form, both because of its great height with respect to its horizontal dimensions and because it lacks the traditional bearing wall of masonry.

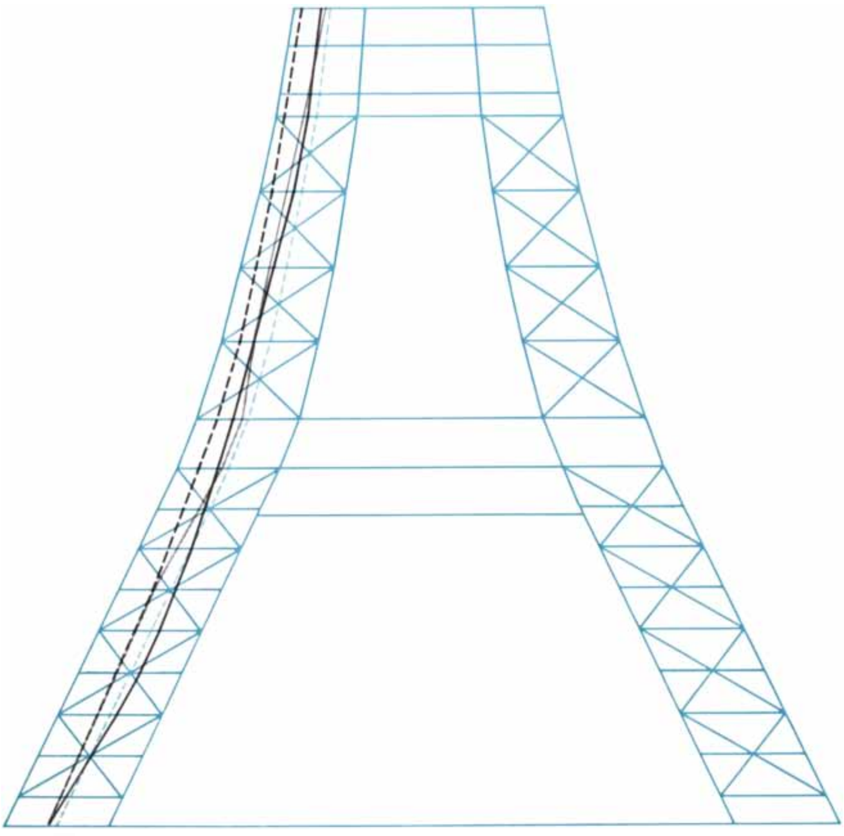
Among the pioneering skyscrapers were the Manhattan Building and the Monadnock in Chicago and the Havemeyer Building in New York. These relied for wind resistance on double diagonals and portal bracing; the last two also had thick masonry walls at the base.

A more sophisticated and more expensive type of bracing known as the portal arch was developed by Corydon T. Purdy for the high, narrow, slablike skyscrapers that became common in Chicago in the last decade of the 19th century. For the Old Colony Building, erected in 1893 and 1894, he designed an elaborate system, again derived from the portal frame of the truss bridge, in which steel plates were riveted together to form a solid arch that was in turn riveted throughout the depth and breadth of the bay to the columns and to the undersides of the floor girders [*see illustration on opposite page*].

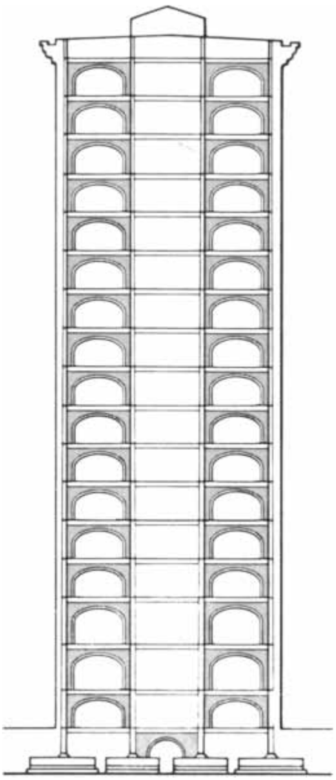
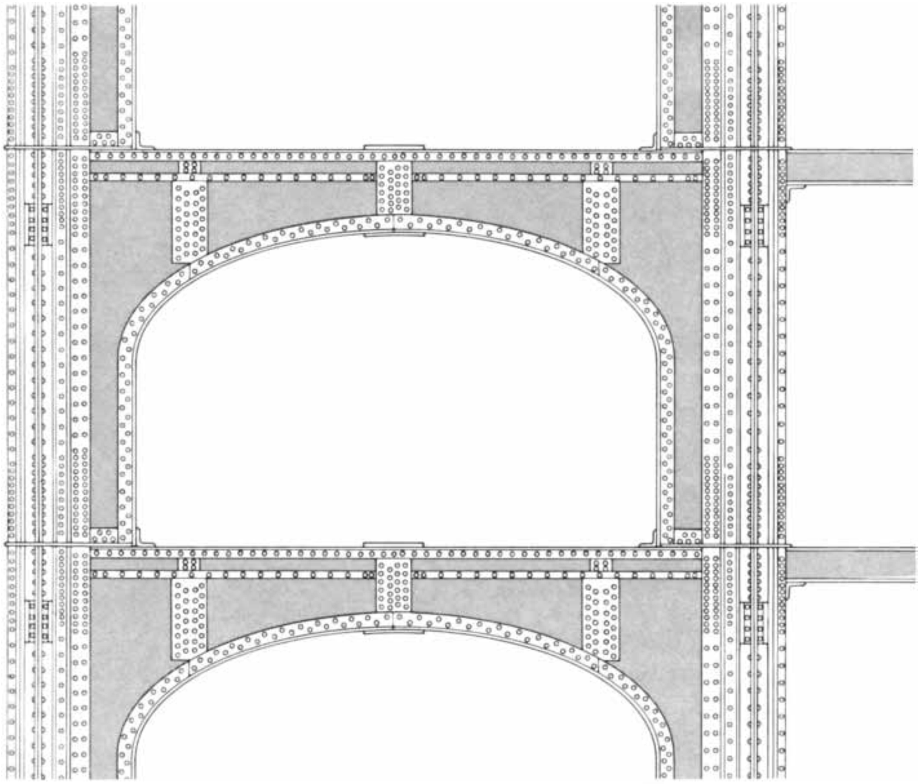
Purdy's system of portal arches was costly in materials and labor, but its strength led the engineers of the Gunvald Aus Company to adopt it for the Woolworth Building in New York. Built between 1911 and 1913, the Woolworth Building, at 760.5 feet, was the tallest in the world for almost 20 years. (The Eiffel Tower was still some 150 feet higher.) Portal arches brace the frame up to the 28th story; above this level conventional knee braces are riveted above and below the primary transverse girders where they meet the columns.

The continuing increase in the height of New York skyscrapers, from the 55 stories of the Woolworth Building to the 102 stories of the Empire State Building, accompanied by rising costs of materials and labor, soon made the portal arch prohibitively expensive. Not only was its fabrication expensive; it also demanded a great sacrifice of space that could otherwise be profitably employed. As a result older and lighter, although frequently no less awkward, forms of knee, diagonal and portal bracing came to predominate in tall buildings.

Certain characteristics of the structural and architectural design of the skyscraper in this period allowed relatively simple devices to serve as wind-bracing elements. Spans between columns seldom exceeded 30 feet and usually were closer to 20. Interior partitions were of heavy, hollow-tile construction and extended from floor to floor. Total window area was less than half the wall area, and the individual windows were approximately standard size. The curtain wall



BENDING MOMENTS in the base frame, or *montant*, of the Eiffel Tower correspond closely to the form of the tower itself. When graphed as a function of height, the moments describe a segment of a parabola; the four bents that compose the *montant* approximate the same curve. The bending moment for a uniform wind load of 300 kilograms per square meter (*solid black curve*) and the moment for a wind load that varies from 200 kilograms per square meter at the base to 400 at the pinnacle (*broken black curve*) diverge only slightly from the median line of the frame (*broken colored curve*). The pressure line (*gray curve*) along which the dead load of the structure is transmitted also conforms to the frame's shape.

PORTAL ARCH, an elaboration of portal bracing, enabled builders in the last decade of the 19th century to erect tall, slablike buildings without heavy bearing walls of masonry. The Old Colony Building in Chicago, built in 1893 and 1894, was braced by portal arches in all its 17 stories (*elevation drawing, left*). Although the center bays were left vacant, the vertical series of arches in the exterior bays formed rigid bents capable of supporting the structure. Each arch was fabricated from flat iron plates and riveted to the flanged cylindrical columns (*detail at right*). The system was expensive in materials and labor, and it obstructed interior space.

between vertical rows of windows was finished in a limestone veneer, and the spandrel panels (which vertically separate the windows) were frequently stainless steel. The structural consequences of these features were a high dead load per unit of floor area and a dense array of columns among which the bending and shearing forces induced by the wind could be distributed. The skyscraper thus possessed a certain built-in resistance to wind, analogous to that of the Romanesque church, and the engineer was as much concerned with gravity loads as with wind resistance.

When commercial building resumed in 1950, however, after the long hiatus of the depression, World War II and the postwar adjustment, a new architectural fashion emerged under the influence of the European architects Mies van der Rohe and Le Corbusier. The glass-walled skyscraper became the ruling mode, and the demand for unobstructed interior space led to the adoption of light, freestanding partitions and bay spans of 40 feet or more.

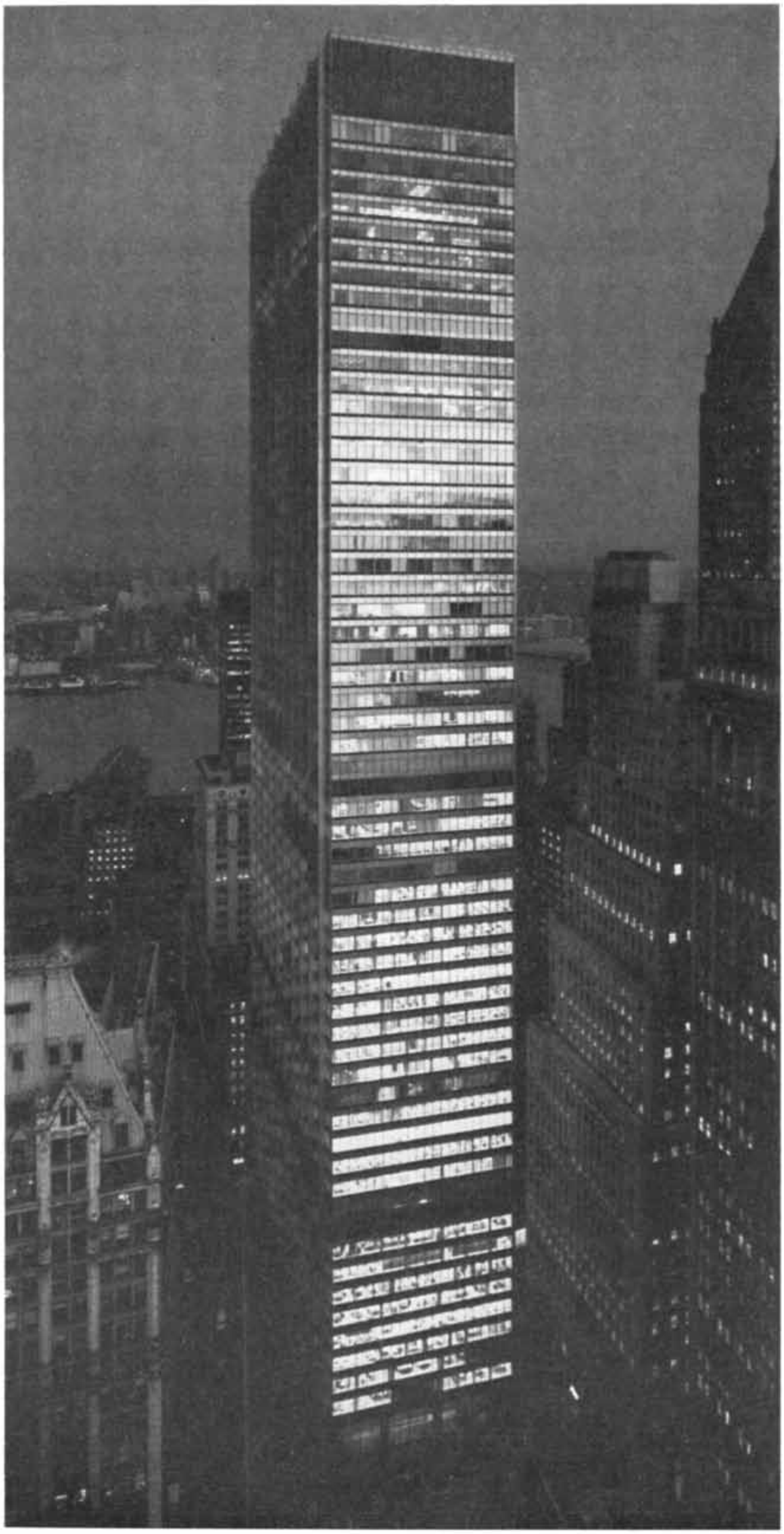
Engineers first approached the new architectural style through traditional techniques, whose main features, as has been shown, can be traced back to the mid-19th century. The work that represents the culmination of this century-long development, and at the same time indicates why new techniques had to be found, is the Chase Manhattan Bank building in New York [*see illustration on next page*].

Designed by the architectural and engineering firm of Skidmore, Owings & Merrill and the structural engineers Weiskopf & Pickworth, it was erected between 1957 and 1961 and was the first glass-walled building to reach a height of 800 feet. Of necessity it summed up the structural science of an age. It was the subject of a model structural analysis and exposition in H. Seymour Howard's *Structure: An Architect's Approach.*

The Chase Manhattan building is 60 stories high; its floors are divided into three bays transversely and nine longitudinally. The New York City building code specifies that buildings must be capable of withstanding a wind load of 20 pounds per square foot on all surfaces more than 100 feet above grade. (The frontal pressure of the wind on a building is proportional to the square of the wind velocity. The empirically derived formula adopted in the U.S. to relate these quantities is $P = .00256V^2$, where P is pressure in pounds per square foot and V is velocity in miles per hour. A load of 20 pounds per square foot is thus equal to a wind velocity of 88 miles per hour.)

For the Chase Manhattan building a wind load of 20 pounds per square foot yields a frontal pressure per bay of 422,000 pounds. The resultant of the distributed wind load acts at right angles to the plane of the wall at an elevation of 446 feet. Bending and shear forces are maximum at the plaza (ground floor) level. The bending moment for each bay is thus 422,000 $\times$ 446, or 188 million foot-pounds.

The action of the wind on a tall building in a densely built urban area is complex but must be predicted by the engineer before he can design the steel frame and calculate the size of its members. Eiffel's design of the Paris tower indicates that he understood the three primary effects: the bending of the entire structure before the horizontal force, the tendency of the structure to over-



TRADITIONAL BRACING METHODS CULMINATED in the design of the Chase Manhattan Bank building in New York City, completed in 1961. The 813-foot, 60-story tower was designed by the New York office of Skidmore, Owings & Merrill. It represents an approach to the limits of structural science for buildings supported by an interior frame of wind-braced steel beams and enclosed by a light "curtain" wall of glass and metal.

turn around the baseline of the leeward side and the shearing force on all columns. He also recognized that because iron is an elastic material wind moments must inevitably be translated into a measurable lateral drift at the top of the structure, which in turn stimulates oscillations that must not be allowed to reach a dangerous amplitude.

Wind-tunnel experiments have revealed still other difficulties. The proximity of many tall buildings causes turbulence in the airstream; moreover, the venturi effect in the narrow spaces between buildings can increase the velocity of the wind. High wind pressures accompanied by sudden changes of direction produce double flexures in the building frame. The formation and repeated shedding of vortexes, caused by the obstruction of the air current, leads to oscillations in the plane perpendicular to the direction of the wind. Finally, inequalities of air pressure, cantilevered floors and the placement of columns outside the plane of the wall subject the frame to torsion as well as to bending, and sometimes create negative pressures, that is, higher air pressures inside the building than outside it.

The engineer must design a frame and calculate the cross-sectional areas and shapes of framing members so that the structure will endure these forces (at least over its expected lifetime) and will not be subject to wind drift that is uncomfortable or alarming to occupants. The first step in the design process is to translate the various effects of the wind into quantitative mechanical actions on the frame. The standard practice today is a return to Eiffel's approach: a tall building is considered a vertical cantilever under wind loads.

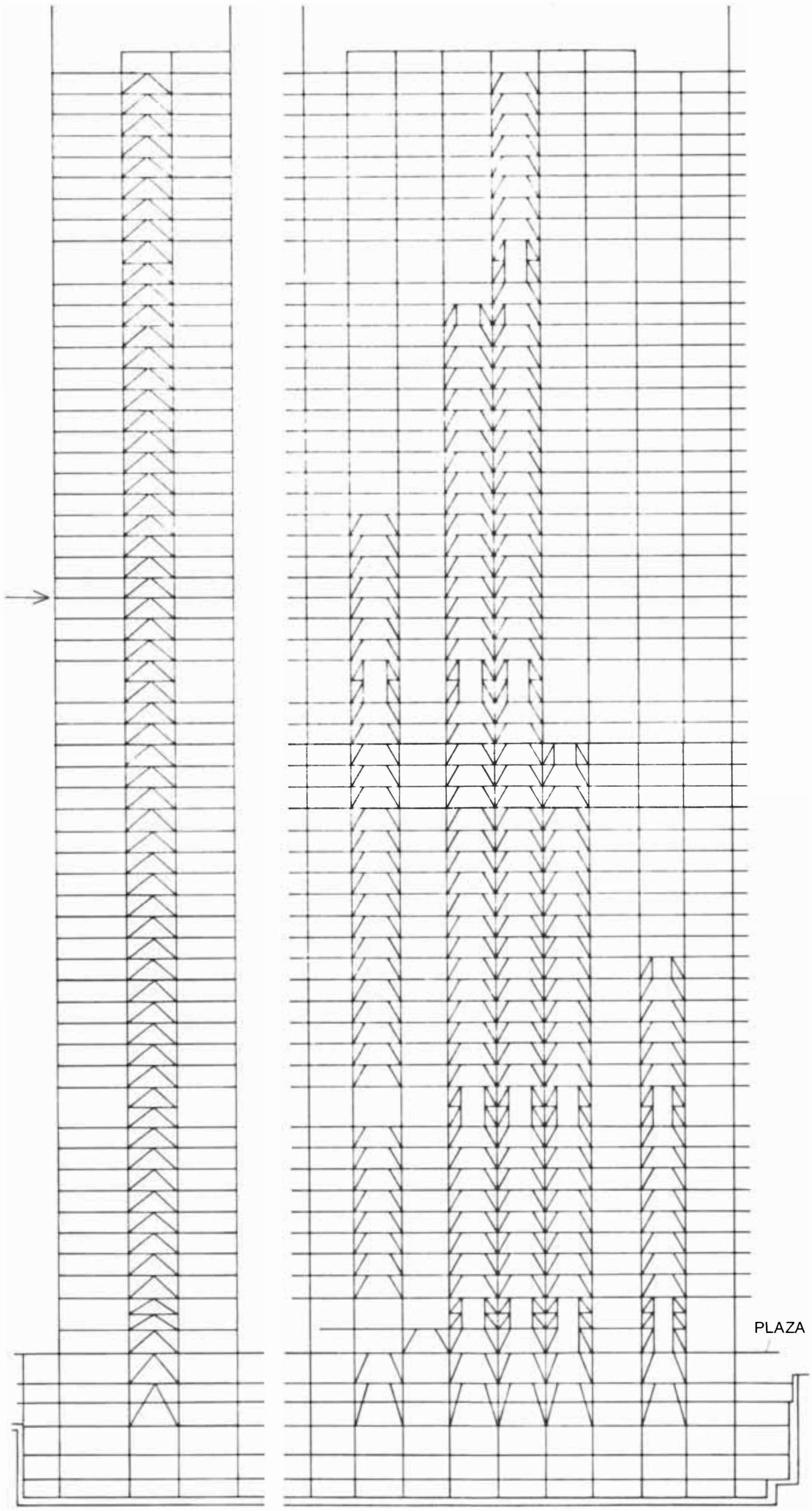
When a cantilever beam (or building) is deflected, the half on the side nearer the load is convex and the half on the opposite side is concave. The convex half is under tension, ranging, in theory, from zero at the neutral plane bisecting the beam to a maximum at the extreme fiber. The concave half is under compression, over the same range of values. The same generalizations can be made about a deflected simple beam (a beam supported at both ends) except that a simple beam is concave on the side nearer the applied force. The beam is also subject to axial shearing forces, which are maximum in the zone where tension and compression approach zero, and to transverse shearing forces, which are maximum at the supported end of the cantilever and close to each of the supports of the simple beam [*see illustration on next page*].

For a building treated as a cantilever, if wind action alone is considered, the columns on the windward side will be under tension and those on the leeward side under compression. Shearing forces will be present in the framing members parallel to the direction of the wind and in all columns. In the columns the greatest forces will be generated by winds blowing on the broader sides of the building, since in that situation there are fewer columns in depth to resist the shearing and bending forces.

In the case of the Chase Manhattan building a wind blowing directly on one of the broad faces is resisted by four columns in each bay. With a wind load of 20 pounds per square foot the windward columns will be under 1.25 million pounds tension, the leeward columns under 1.57 million pounds compression. The two interior columns will bear much smaller loads. The distribution of shearing forces is opposite in character, the interior columns bearing the major part of the load [*see illustration at right*].

Three types of brace ensure the rigidity of the structure under these loads. K trusses were inserted in the center bays of the transverse section on all floors. (The K truss in this case looks like a letter *K* turned on its side.) Floor-to-floor knee braces stiffen all bays in the longitudinal section where they do not interfere with interior planning or movement and portal bracing is employed throughout the building. In the center bays of both the transverse and the longitudinal sections diagonal bracing extends from the basement to the roof. There are two reasons for this concentration of bracing members: the central bays, like the median fibers of a deflected beam, lie along the axis where the frame is most likely to be pulled apart by the forces generated by the wind. In addition the openings in this region for elevator shafts are vulnerable because of complex stress concentrations at their edges.

There are two remaining wind effects the engineer must consider. First, he must be certain that the building cannot be overturned by the wind. This can be determined by a balancing of moments; if the building is to remain upright, the overturning moment of the wind about the baseline of the leeward side must be less than the righting moment of the dead load about the same line. For the Chase Manhattan building the wind moment per bay, as determined above, is 188 million foot-pounds. The dead load



STEEL SKELETON of the Chase Manhattan building is braced by K trusses in the transverse bays (*left*), by floor-to-floor knee braces in the longitudinal bays (*right*) and by portal bracing at all joints between column and girder. Wind blowing directly on one of the broad faces of the building acts at a point 446 feet above the plaza level (*arrow*). A wind of 88 miles per hour would produce in each transverse bay a pressure of 442,000 pounds.

of 15.3 million pounds per bay, acting over a distance of 56.2 feet from the center line of the building to the leeward baseline, yields a righting moment of 860 million foot-pounds. The building's stability under wind loads is ensured, since the counterbalancing moment is almost five times the overturning moment.

The final quantity the engineer must determine is wind drift, or oscillation. He specifies a limit for this deflection, then designs the frame to prevent distortions that would exceed the limit. For older buildings a factor of $.003h$, where h is the height of the building, was customary. For newer buildings of more than 800 feet this limit has been reduced to $.0015h$, which was the factor adopted for the Chase Manhattan building. The allowable drift is thus 14.4 inches, or 7.2 inches on each side of the central axis.

The structural system of the Chase Manhattan building would seem to be the last word in interior wind bracing, and in a sense it was. The number of bracing elements in the framework is immense: there are almost 3,000 diagonal members, many more thousands of gusset plates and angle irons, and millions of rivets. Tremendous weight and many man-hours of labor have been devoted entirely to maintaining the rigidity of the building under wind loads.

As the demand for even higher buildings continued, whatever the damage to the urban environment, it became clear that less costly methods of wind bracing would have to be devised. In the early 1960's several engineers simultaneously developed a new technique.

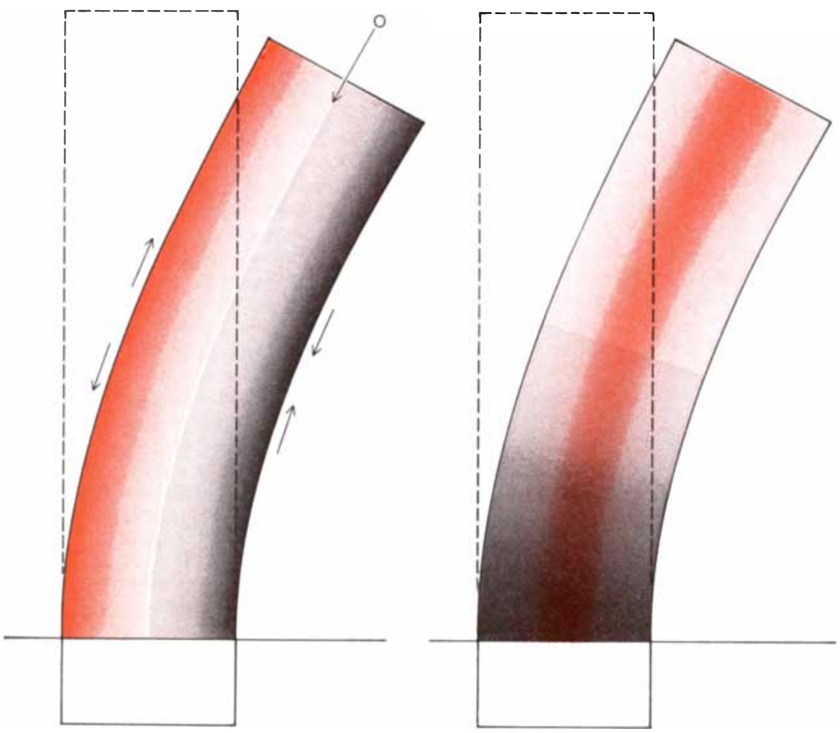
The underlying principle is a contemporary variation on concepts embodied in the design of the Eiffel Tower, although it is uncertain whether the designers knew this. The major innovators have been Fazlur R. Khan and Myron Goldsmith of the Chicago office of Skidmore, Owings & Merrill, and several engineers on the staff of the Seattle structural designing firm of Worthington, Skilling, Helle and Jackson.

These designers reasoned that if a building could be constructed as a hollow tube of solid walls, like a giant box column, or as a braced tubular framework, like the iron bents of early railroad viaducts, it would behave as a rigid cantilever with external walls sustaining all horizontal forces. In this way the proliferation of interior braces might be drastically reduced.

This end can be attained in several ways, and the earliest works demonstrate two of them. One is the International Business Machines Building in Pittsburgh, erected in 1962 and 1963. The architects were Curtis and Davis and the engineers were Worthington, Skilling, Helle and Jackson. The building is a rigid hollow box with walls of lattice trusses, a form invented in 1820 by Ithiel Town, abandoned long ago for bridges and never before used in a building. The second early example is the Chestnut-DeWitt Apartments in Chicago, constructed at the same time as the IBM building. It is the work of Skidmore, Owings & Merrill, for whom Khan was in charge of structural design. The building is a hollow tube of reinforced concrete in which the walls form a monolithic framework of closely ranked tapering columns and short beams.

The most spectacular and sophisticated work in this style is the 100-story John Hancock Center in Chicago, designed by Skidmore, Owings & Merrill, with Khan in charge of structure, and erected between 1965 and 1970 [*see illustration on page 105*]. A braced, tubular cantilever of tapering profile, it is an almost exact skyscraper counterpart of the ancestral bridge bent. Each of the four walls consists of columns and girders rigidly tied together by immense double diagonal braces that form X trusses over five panels that are each 18 stories high and a K truss over the top 10-story panel. The axes of column, girder and diagonal meet at a common point, where they are rigidly joined by a huge gusset plate.

The structural result of this impeccably designed "superframe" is that all bending forces arising from the wind and all axial loads caused by wind-induced tension and compression in columns are removed from the interior frame of the building and confined to the trussed walls. The diagonals also serve to distribute forces, so that any pattern of vertical loads on the column at a particular level is equally divided among all the columns at the base. Even more remarkable, since it is contrary to what would be expected in a truss, the diagonals are almost always under compression, regardless of the action of the wind; for this reason they are called "inclined columns." The economic consequence of this design is as spectacular as the appearance of the building: the



DEFLECTION OF A CANTILEVER induces at least four types of stress. Tension and compression (*left*) are greatest at the extreme fibers and approach zero at the median line. Tension (*color*) is induced on the convex side, compression (*gray*) on the concave side. Shearing forces (*right*) are also present. Longitudinal shear (*color*) is caused by the opposition of tensile and compressive forces and is greatest at the median line; transverse shear (*gray*) is maximum at the anchorage of the structure. In a building regarded as a cantilever deflected by wind loads, overturning and torsional forces may also be generated.

Read what Ed Kuen-Sang Shum says about the new Ampex portable instrumentation recorder:

"This is the first data recorder I've ever seen that lets you concentrate on the data and forget about the recorder. It was designed to be used by scientists and research engineers, not tape recorder mechanics.

"For example, while you're getting your experiment ready, you don't have to worry about whether or not the recorder is ready. It has an automatic diagnostic capability, built-in, that monitors internal voltages, capstan servo operation, recording bias, and power supply level. If there's a problem somewhere, even while you're recording, the system lets you know and tells you where it is so immediate action can be taken to get back on-line.

"Everything you need to get at is located on the front panel: FM calibration, adjustments, signal electronics modules, and all operating controls.

"And this machine has the lowest flutter and skew characteristics of any portable instrumentation recorder I know about. It's because somebody has really studied the tape-handling problem and come up with a unique tape path.

"When you want to take this recorder to a new location, or somewhere out in the field, you don't have to take along special power adapters and accessories to make it work. It runs on any AC power from 90 to 250 volts, from 47 to 420 Hz. There is also a unique AC/DC convertible power supply available.

"And it records data on as many as 32 separate tracks, from 15/16 to 60 inches per second, in either direction; covers IRIG bandwidths up to 300 kHz direct, 40 kHz FM.

"Maybe I'm biased, but this PR-2200 recorder is the nicest little package of mechanical and electronic innovations Ampex has ever put together. For my money, or anybody else's, there's nothing else around that even begins to touch it."

Of course Ed Shum is biased; he's our Section Manager for Longitudinal Recording, supervising the PR-2200 project. But even if you make allowances for his understandable enthusiasm, you have to admit he's got some good points. There are many others, detailed especially toward the recorder user's point of view, that we'd like to get into your hands. Please send for our new 8-page brochure today.

AMPEX

Ampex Corporation
Instrumentation Division
401 Broadway – M.S. 10-14
Redwood City, CA 94063

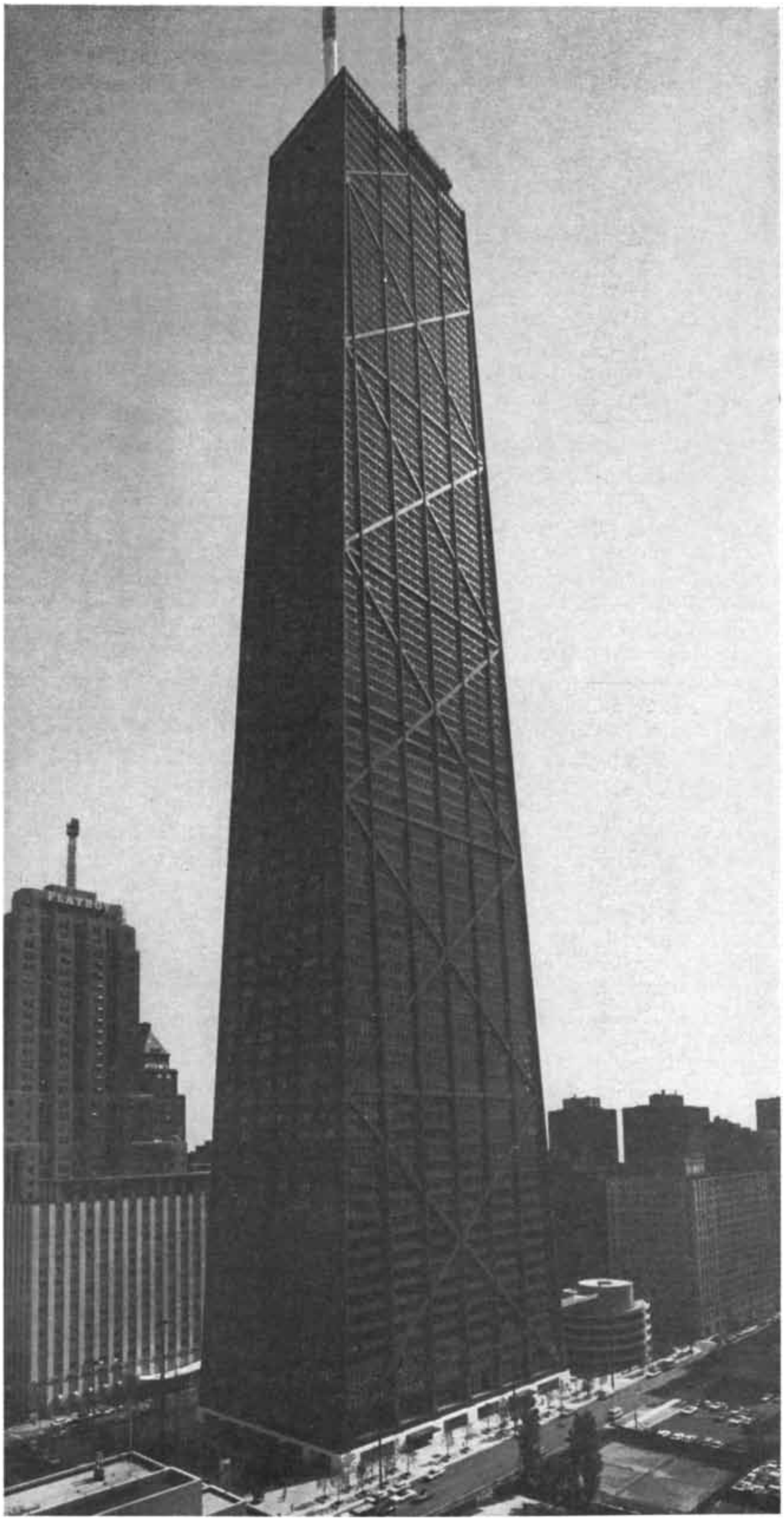
structure contains 29.7 pounds of steel per square foot of floor area, compared with the 40 to 50 pounds per square foot common in conventionally framed buildings.

There are other methods by which a rigid tubular cantilever can be constructed. The wall structure that most closely approximates a solid diaphragm is that of the Standard Oil Company Building in Chicago. Designed by Edward Durell Stone and the Perkins & Will Partnership and begun in 1970, the tower has 80 stories and stands 1,360 feet high. Small, widely spaced windows leave the major area of the wall a solid steel plate stiffened with buttresslike triangular prisms that, in addition to their structural role, house ducts, conduits and pipes.

Another variation is the World Trade Center in New York, which consists of two 110-story towers, each 1,350 feet high. Designed by Minoru Yamasaki, Emery Roth & Sons and Worthington, Skilling, Helle and Jackson, it was begun in 1968. The fine-grained surface pattern of the towers is formed by the closely spaced vertical elements of Vierendeel trusses. These trusses have no diagonal members, rigidity being secured by joints designed to resist bending and shear. A similar system is employed in the wall frames of the Sears, Roebuck and Company building in Chicago, at present the tallest building in the world at 1,450 feet. The architects and engineers were Skidmore, Owings & Merrill and construction was begun in 1970.

A third system was employed in the Boston Company building, erected in 1969 and 1970 in Boston. The architects were Pietro Belluschi and Emery Roth & Sons and the engineers were of the James Ruderman office. K trusses in the wall frames, supplemented by a ring truss (a large circumferential truss in the wall) at the second-floor level, transmit all wind loads and most gravity loads to four tremendous corner columns of cruciform section and tapering profile. This structural system has so far proved the most economical: the quantity of steel is reduced to 21 pounds per square foot of floor area.

All the ancestors of these framing systems may be found in the evolution of truss bridges over the past 150 years, and the "tubular" buildings may be regarded as cantilever bridges stood on end. When designs analogous to suspension-bridge construction are introduced, as they are beginning to be, the unity of steel-bridge and steel-framed skyscraper design will be nearly complete.



EXTERIOR FRAMEWORK of immense steel members sustains all wind loads in the 100-story John Hancock Center in Chicago. On each face of the building the braces form X trusses over five panels of 18 stories each and a K truss over the top 10 stories. The tower is a braced, tubular cantilever, a skyscraper counterpart of the 19th-century iron bridge bent. Bending forces induced by the wind are removed from the interior frame and confined to the trussed walls. In addition any pattern of vertical loads is distributed equally among the columns. Architects and engineers were the Chicago office of Skidmore, Owings & Merrill.

MATHEMATICAL GAMES

Cram, crosscram and quadraphage: new games having elusive winning strategies

by Martin Gardner

There are many simple two-person games, for example nim, for which perfect-play strategies are known. Other games, such as ticktacktoe and dots-and-squares, may appear just as simple but actually are so complex that no strategy has yet been found except when the game is played on special fields. In spite of recent progress in computer speeds and sophisticated programming, no one yet knows whether ticktacktoe on a 4-by-4-by-4 cubical matrix can always be won by the first or second player, or whether the game is a draw if played rationally. No one knows who has the win on a dots-and-squares field of six dots on the side.

This month we consider several elegant new games that have extremely simple rules and about which relatively little is known. Some may not have general strategies; if they do, perhaps a reader of this department will be the first to discover them.

Our first game, as far as I am aware, has not been described in print, although a few mathematicians have been involved with it for at least 20 years. I originally heard of it from Geoffrey Mott-Smith, the author of several books on games and puzzles who died in 1960. He told me it had been invented by a friend who called it "plugg." Since then I have received letters from a number of mathematicians who independently invented the same game. In 1966 John Horton Conway gave it some thought, and although he did not succeed in cracking it, he did formulate a partial strategy with which the final stages of a game could be analyzed by standard nim theory.

The game can be played in various ways, all isomorphic. If the "board" is a rectangular lattice of dots in unit square formation, the rules are as follows. Players alternately draw a line that connects two orthogonally adjacent dots. No line may touch a dot after it has been joined to another. In the standard form of the game the last player to connect two dots is the winner. (In *misère,* or reverse, play the last to move is the loser.) Let us call the game "dots-and-pairs." Clearly it is a graph-theory game.

If a supply of counters is handy, one can arrange the counters to form the lattice, then each move consists in removing two orthogonally adjacent counters. Another way to play dots-and-pairs is to sketch a checkered field on paper, or outline it on graph paper. A move consists in coloring two orthogonally adjacent cells or, more simply, drawing a line that eliminates the "domino."

Still another way to play the game—the most pleasant—is to place dominoes on a checkerboard. The markings on the dominoes are of course irrelevant. All that matters is that each domino must cover just two squares. Players alternately place one domino until no further play is possible. We shall call this version of the game "cram." It is the simplest nontrivial polyomino-placing game. (Readers will be interested to know that Solomon W. Golomb's polyomino-placing game, using the 12 pentominoes, has finally reached the marketplace. Issued by the Springbok Division of Hallmark Cards Inc., this masterpiece of packaging includes the 8-by-8 board, the pieces and the authorized instructions for both the game and various puzzles. The trade name is Pentominoes. It appears just 20 years after Golomb introduced polyominoes to a group of mathematicians in his memorable talk to the Harvard Mathematics Club.)

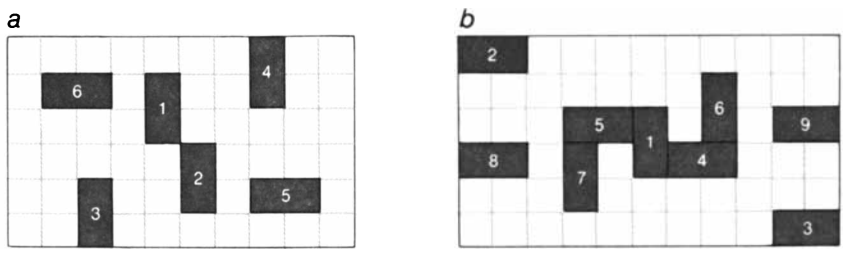
Winning strategies for cram are known for certain boards. If, for instance, the field is rectangular with two even sides and the game is standard (the last to play wins), the second player wins easily by symmetry play. He simply makes each move symmetrically opposite his opponent's last move [*see "a" in illustration on this page*]. To eliminate this strategy we can add a new rule: The second player's first move must not be symmetrical with respect to the first player's opening move. With this proviso the game can be played on a standard chessboard with 32 dominoes. It is not known which player has the win.

If standard cram is played on an even-by-odd rectangle, the first player wins by taking the two central cells, then playing symmetrically thereafter [*see "b" in illustration on this page*]. This strategy can be eliminated by denying the center to the first player on his opening move.

No general strategy is known for the reverse form of cram on even-even or even-odd fields, and no general strategy is known for standard or reverse play on odd-odd fields. Even when one of the odd sides in an odd-odd field is reduced to 1, the game is complex and still unsolved. In 1973 David Singmaster, then at the Istituto Matematico in Pisa, wrote a computer program for the 1-by-m field, standard game, with m less than 1,000. Assuming that the first player loses (because he cannot move) when m equals 0 or 1, Singmaster found 151 values of m that give the win to the second player. For m less than 100 the values are 0, 1, 5, 9, 15, 21, 25, 29, 35, 39, 43, 55, 59, 63, 73, 77, 89, 93, 97.

When m is even, the first player wins, of course, by taking the center and playing symmetrically. When m is odd, he wins for all values less than 100 that are not in the above set. I know of no computer analysis of the reverse game for 1-by-m fields.

Square fields of only the lowest orders have been investigated. The 3-by-3 game



Symmetrical winning strategy for cram on even-even fields (a) *and even-odd fields* (b) is trivial. It takes just a few minutes of analysis to see that the second player wins the standard game and loses the reverse game. The 4-by-4, with symmetry play denied the second player in standard play, takes considerably more work. Conway found it to be a second-player win in both standard and reverse play.

What about the 5-by-5? Because it is odd-by-odd, symmetry strategy is ruled out and no special rules are needed. Who has the win in standard play? In *misère*? As far as I know, neither question has yet been answered.

Cram is an "impartial" game. This means that any possible move can be made by either player. "Partial" games, or what Conway prefers to call "unimpartial" games, are those in which moves open to one player are denied to the other. Chess and checkers, for example, are partial games because each player can move only the pieces of his color. We can convert cram to a partial game by a rule proposed by Göran Andersson, who wrote to me about it in 1973.

The rule is delightfully simple. One player must make all his moves horizontally, the other must make all his moves vertically. Call it "crosscram." This, of course, instantly eliminates symmetry play from all rectangular boards. The 3-by-3, as before, can be quickly disposed of. The first player wins the standard game provided that his first move does not include a corner cell. The second player wins the reverse game.

Crosscram in the 4-by-4 form is sufficiently complicated to make a good pencil-paper game [*see top illustration on this page*]. I am too lazy to analyze it but would welcome hearing from readers who do, and I shall report on results in a later column.

Both cram and crosscram can be regarded as special cases of more general games. Cram is Piet Hein's Tac Tix, now more commonly called nimbi. (See Chapter 15 of *The Scientific American Book of Mathematical Puzzles & Diversions*.) Nimbi is usually played with a square array of counters. A move consists in removing as many orthogonally adjacent counters as desired, not just a pair, from any orthogonal row. Crosscram is a special case of partial nimbi in which the rules restrict one player to horizontal rows and the other to vertical.

Another interesting form of partial nimbi was invented in 1972 by James Bynum of Tacoma, Wash., who has permitted me to describe it here. It is the same as partial nimbi except that each play must have a maximum length; that is, the orthogonally adjacent cells must be bounded on each end by either the field's border or an opponent's perpendicular move. The game's first move must necessarily be an entire row or column [*see second illustration from top at right*].

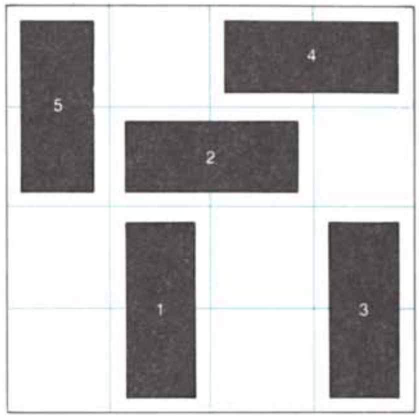
Bynum's game was solved in 1973 by Conway. *Misère* play is almost trivial. The second player wins on all square boards, and if the board is a nonsquare rectangle, the player whose moves parallel the shorter dimension wins regardless of whether he goes first or second. The winning rule is: Pick one of the two sides that parallel your moves, then always play as close to that side as possible. Standard play is more interesting. The first player has the win on all square boards as well as on all nonsquare rectangles with sides of the same parity (even-even or odd-odd). On even-odd rectangles the player whose moves parallel the even dimension wins regardless of who plays first.

Interested readers may enjoy seeing if they can develop a set of strategic rules that will ensure a win for the player who has the win. The game yielded readily to Conway's unpublished theory of unimpartial games. I shall say no more about his analysis because it will be included in a book on unimpartial game theory that Conway is reported to be writing.

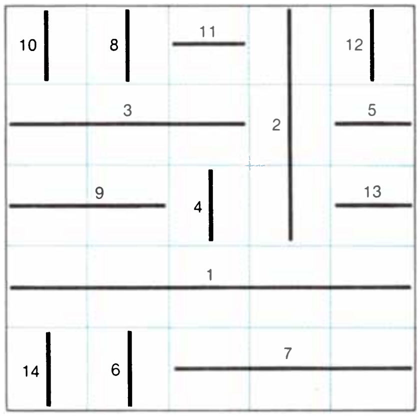
Quadraphage (square-eater) is a partly explored family of games invented and named by David L. Silverman in the late 1940's. He suggested the basic idea to Richard A. Epstein, who mentions one version briefly on page 406 of his *Theory of Gambling and Statistical Logic* (Academic Press, 1967). Silverman discusses two other versions on page 186 in his book of game puzzles, *Your Move* (McGraw-Hill, 1971). Elwyn Berlekamp has done considerable work on quadraphage, which he will summarize in a book on games that he is preparing in collaboration with Conway and Richard K. Guy. Here I shall introduce only some of the game's simpler aspects.

Quadraphage games are played on a chessboard of side n, usually square. Pieces consist of one chess piece, usually a king, and a supply of counters. The counters are the quadraphages, which I shall call quads for short. Each quad "eats" the cell on which it is placed, thus preventing the king from moving to that cell. The game starts with the king on the central square if the board is odd-odd, or on one of the four central squares if the board is even-even. The rest of the board is empty. One player moves the king in the usual manner. The other player places counters, q at a time, on Q cells. As in the game of go, counters do not move once they are placed. The object of the king is to get safely to the edge of the board. The quads try to box in the king so that it cannot escape. The quad player conventionally goes first, then the players move alternately.

If q equals 4 (four squares eaten on



Standard crosscram with a first-player win



Bynum's game with a second-player win

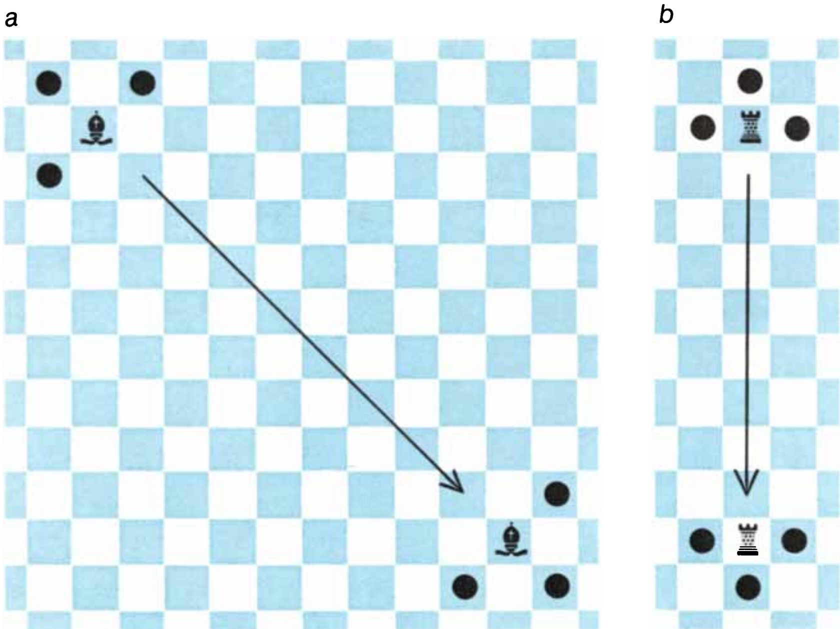

Two quads to trap king on an 8-by-8 board

each move), it is easy to show that the king can be captured in no more than three moves on all boards of side 5 or greater. (Of course, the king escapes immediately on a 4-board.) If q equals 3, it takes only a little more effort to find that the king can be trapped on boards of side 6 or greater.

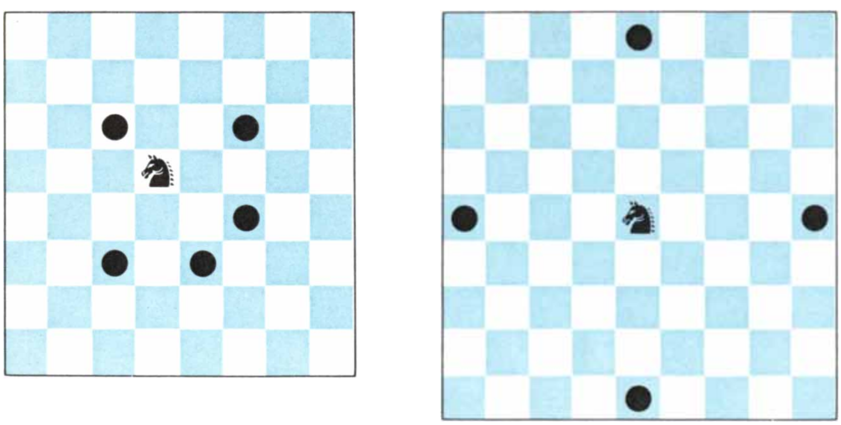
When q equals 2, the game starts to get interesting. Although I have not proved it, I believe the king can escape on the 7-board but can be trapped on the standard 8-board and all higher boards. The strategy is to move first as shown in the bottom illustration on the preceding page, then continue by placing quads two at a time on only the white border cells except when the king is next to the border and the quads must be placed side by side to prevent its escape.

What about $q = 1$? Can the king always escape no matter how large the board? Surprisingly, it cannot. Berlekamp has proved that on a board as small as 33-by-33 the king is lost. Unfortunately both proof and strategy are too complicated to give here. Golomb has shown that if the king's moves are limited to orthogonal moves and q equals 1, the king escapes on the 7-board but can be trapped on the 8-board.

Although the king escapes easily on a standard chessboard, a pleasant game (proposed by Silverman) results if the king tries to maximize its moves before reaching the border. Quads, placed one per move, try either to trap the king or to force it to the border as quickly as possible. If the king escapes, it is awarded one point for each quad on the board. No points are awarded if the king is trapped. The game is particularly enjoyable when it is played on the order-18 go board.



How three quads per move trap bishop (a) *and rook* (b) *on infinite boards*



Trapping a knight on order-8 board (left) *and order-9 board* (right)

The go board, with its large supply of stones, is a handy tool for working on the many unsolved quadraphage problems. What, for instance, happens when the king is replaced by a different chess piece? If the piece is a bishop, rook or queen, we must limit the length of its move to avoid triviality. Assume that the board is infinite but the piece cannot move a distance of more than, say, a billion squares. With this limitation a bishop is easily captured in a goose chase along a diagonal by placing three quads per move as shown in *a* in the upper illustration on this page. Once the ends are blocked the bishop can be confined to the diagonal. Three quads per move can similarly trap a rook in a goose chase along an orthogonal, as shown in *b*, and seven can trap a queen along either an orthogonal or a diagonal. Can the bishop or rook be trapped by two quads per move? By one? Can the queen be trapped by fewer than seven? I shall give answers to some of these questions next month.

If the piece is a knight, we must consider it free when it lands on a cell that is second from the border because on its next move it can leap over a border cell. On boards of sides 5, 6 and 7 the centered knight has eight moves to freedom, so that eight quads obviously are needed to trap it. Five quads per move will trap the knight on the order-8 board, and four per move are sufficient for the order-9. The lower illustration on this page shows one of several winning first moves for each board, although I must add that I am not certain all this information is correct.

Can three quads per move trap a knight on an infinite board? One quad per move surely is insufficient, although Berlekamp reports he has yet to see a rigorous proof.

Correction: The chart in last December's column, which listed simplest-known expressions by cubes for integers 1 through 99, had two errors. In the sum for 51 the sign for 602^3 should be positive and in the sum for 79 the sign for 33^3 should be negative.

HUMAN ECOLOGY

Problems and Solutions

PAUL R. EHRLICH and ANNE H. EHRLICH, Stanford University, and JOHN P. HOLDREN, California Institute of Technology

Here is a compellingly written and authoritative paperback that introduces the reader to the fundamentals of human ecology. Drawing partly on material from the Ehrlichs' distinguished work *Population, Resources, Environment,* the authors of this new book outline the basic problems of the human environment and analyze these problems in their global context: how do the actions and orientations of people in some areas affect the environmental future of other people? The authors then provide the reader with a general model for thinking about the environmental dilemma and proceed to discuss the nature of possible solutions. How must population growth be limited? How must all of us change our attitudes toward the world around us? The result is a broad, well-documented initiation into this field.

CONTENTS: Population, Resources, Environment—Is Mankind Really in Trouble? The Human Population. Carrying Capacity: Land, Energy, and Mineral Resources. Carrying Capacity: Food and Other Renewable Resources. Pollution: Direct Effects on Society. Disruption of Ecological Systems. Understanding the Web of Blame: The First Step to Solutions. Population Limitation. Changing Human Behavior: Toward the Environment and Toward Our Fellow Man. Synthesis and Recommendations.

1973, 304 pages, 77 illustrations, 26 tables, 6 boxes, paperbound, $4.95

george gerster, photofind-rapho

U. S. NUTRITION POLICIES IN THE SEVENTIES

Edited by JEAN MAYER, Harvard University

"In the past three years, food and nutrition policies have become a subject of great interest to many Americans, of almost obsessive concern to some. It is our collective hope that this book will make a contribution to the quality of the debates concerning national nutrition programs." —Jean Mayer

Based on the official reports of the White House Conference on Food, Nutrition and Health held in 1969, this book is a smaller, more simply written study of the basic policy recommendations made by the conference. Prepared by Jean Mayer, chairman of the conference, and other key participants, it is intended to be read and understood by the layman. Here is a concise, timely, and highly readable analysis of the main ideas on the causes, effects, and possible remedies for hunger and malnutrition in the United States.

1973, 256 pages, cloth $7.95, paper $3.95

SEED TO CIVILIZATION
The Story of Man's Food

CHARLES B. HEISER, JR., Indiana University

This book is a narrative on how man has learned to cultivate the plants and animals that stand between him and starvation. Beginning with a discussion of the origins of agriculture, it proceeds to describe the patterns of food use and distribution that have developed with the growth of civilization, the evolution and breeding of food sources, and some of mankind's problems relating to food supply and hunger.

1973, 243 pages, 84 illustrations, cloth $7.50, paper $3.50

ENVIRONMENTAL QUALITY AND WATER DEVELOPMENT

Edited by CHARLES R. GOLDMAN, JAMES McEVOY III, and PETER J. RICHERSON, University of California, Davis

A multidisciplinary approach to a critical subject by some of the world's leading specialists. This study's principal conclusions are that the public must be given an early and a greater voice in decisions concerning water development and that such development must be undertaken with its environmental impact as the foremost consideration. Included are recommendations for implementing change in water development policies.

1973, 510 pages, 51 illustrations, 30 tables, $17.50

TOWARD A STEADY-STATE ECONOMY

HERMAN E. DALY, Louisiana State University

Toward a Steady-State Economy is a collection of essays on the theme that economic growth for growth's sake is destructive and unsustainable. These essays include consideration of alternatives to "growthmania" as well as criticisms of it. Although topics range from physics to theology, the writers' main emphasis is on political economy. Among the contributors are Preston Cloud, Paul Ehrlich, Kenneth Boulding, Garrett Hardin, Nicholas Georgescu-Roegen, and C. S. Lewis.

1973, 332 pages, 18 illustrations, 3 tables, cloth $8.95, paper $3.95

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

THE AMATEUR SCIENTIST

The voiceprints of birdsongs and cockroaches in a maze

Conducted by C. L. Stong

It is well known that birds can be distinguished not only by their appearance and behavior but also by their song. Some songs, however, are so much alike that it takes a trained observer to tell them apart. Moreover, there may very well be differences in birdsong that cannot be detected even by a trained observer, because of inherent limitations in the human sense of hearing. Such differences may nonetheless be revealed by electronic devices. One of a number of possible observational approaches has been taken by Gary S. Thorpe of 6127 Balcom Avenue, Reseda, Calif. 91335. Thorpe, who is a graduate student in biology at California State University at Northridge, discusses the observation of birds and describes a few recording devices that he employs to aid in their identification.

"Essentially I collect animal sounds, including those of birds and insects, with a battery-powered tape recorder. I then feed the reproduced signals into the pen recorder of an electrocardiograph. The resulting graph displays a characteristic pattern of undulations that is unique for each individual and characteristic of each species even though it represents variations of sound amplitude at relatively low frequency rather than the fine details of the complex sound wave. Subtle differences in the sounds of birds that tend to elude the ear are often clearly recognizable when they are displayed as a graph. For this reason the field glasses, camera and notebook that were once considered to be ample aids to the bird watcher have been joined by the directional microphone, tape recorder and pen recorder.

"In my case the additional devices not only have helped me to identify specimens but also have opened a new dimension in the observation of bird behavior. As suggested by Nicholas E. Collias of the University of California at Los Angeles, sounds made by birds can be grouped in five categories. They are the noises made during flockings and group movements, the sounds made incidental to the procurement of food, warning cries emitted in response to predators and enemies, the reassuring clucks and chirps associated with parent-offspring relationships, and the impressive array of sounds characteristic of sexual behavior and aggression related to it.

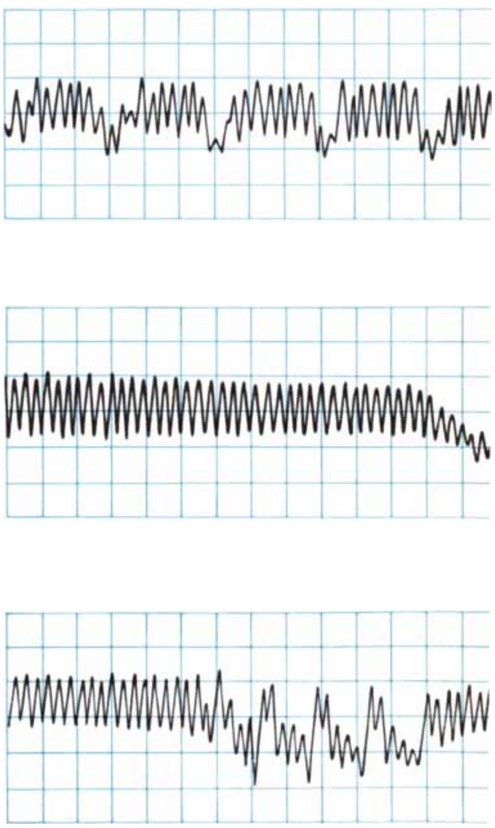
"Sounds in the first two categories are regarded as being incidental. In general they are of less interest to bird watchers than the primary sounds of the other three categories. Indeed, most laymen are inclined to regard as typical of animals (including birds, insects and amphibians) only the sounds associated with sexual behavior and aggression. Among such sounds are the more melodious songs of birds, the croaking of frogs and the characteristic buzzing of many insects, including the chirping of crickets.

"Traditionally the identification of animals in the wild, including birds, has depended on visual observation at relatively close range. The method leaves much to be desired. The presence of an observer may cause the animal to avoid the area or to behave atypically. A bird may switch from song to a screech of warning at the approach of a human being.

"Numerous stratagems have been devised to solve the problem. Perhaps the most popular solution is the backyard bird feeder. Relatively few species of birds venture close to houses, however, and they display patterns of behavior that suggest partial domestication. Much of our knowledge of birds and their behavior has been recorded by observers concealed in blinds that matched the appearance of the wild environment. Blinds, however, are difficult to build, uncomfortable for the bird watcher and restrict observations to a comparatively small area.

"A bird watcher who has a good pair of binoculars, a camera with a telephoto lens, a highly directional microphone and a battery-powered tape recorder can range widely. From a distance of 100 yards or so intimate observations can be made of birds that come to rest on a crag of a sheer cliff, across the rapids of a mountain stream or in the middle of an inaccessible swamp. Moreover, the apparatus can be fitted with an electronic trigger that responds to sound. By aiming the telephoto lens and the directional microphone at an area and setting the trigger the bird watcher can relax his vigilance while the birds automatically record observations of themselves.

"The tape recorder can be of any type that is capable of reproducing sound equal in quality to a reasonably good



radio set. It need not be of the high-fidelity type. My model has a tape cassette designed to record sound on four tracks. The machine is light, compact and battery-powered and thus is ideal for fieldwork. Unfortunately it is not weatherproof. I solve this problem by keeping the recorder and an electronic triggering device in a few nested plastic bags.

"The triggering mechanism is based on a voice-operated relay circuit that was developed originally for automatically switching on a radiotelephone when the operator speaks into a microphone. I modified the circuit for field operation with the tape recorder. The triggering unit consists essentially of a three-stage transistor amplifier that operates from a high-impedance microphone [*see bottom illustration on page 113*]. The output signal of the amplifier is converted to direct current by a diode. The direct current charges a capacitor that is coupled to the base terminal of the output transistor. The charged capacitor causes the output transistor to conduct. The interval of time during which charge remains on the capacitor, and hence the interval during which the output transistor conducts, is determined by the rate at which charge drains from the capacitor through an adjustable resistor.

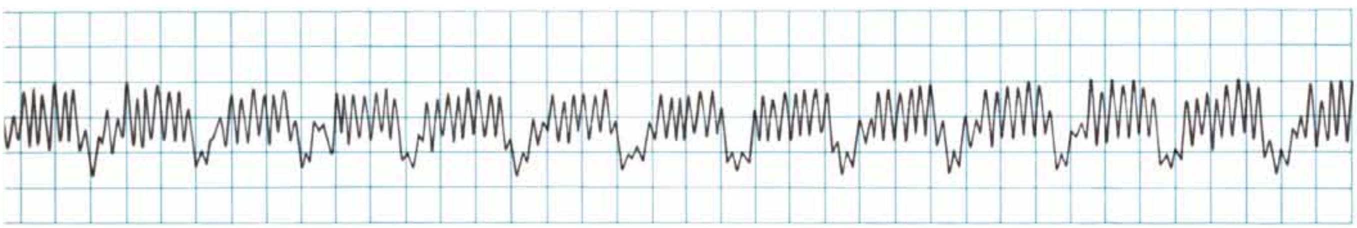
"The circuit of the output transistor includes the coil of an electromechanical relay that operates when an amplified signal from the microphone causes the transistor to conduct. The contacts of the relay are coupled to the on-off switch of the tape recorder. In effect, sounds picked up by the microphone automatically start the tape recorder.

"The recorder continues to operate until charge drains from the capacitor in the base circuit of the output transistor. The period of operation is about two seconds if the variable resistor is adjusted to 5,000 ohms. The time can be increased by substituting a larger variable resistor. For example, at 25,000 ohms the recorder will continue to operate during silent intervals of approximately five seconds. A two-second delay is usually adequate for recording most birdsongs without annoying interruption and for economizing on tape. The relay can also be designed to actuate the electrical shutter release of a camera. Incidentally, the circuit can be made with almost any small PNP transistor that is capable of dissipating .15 watt and has a beta (current gain) of 35 or more.

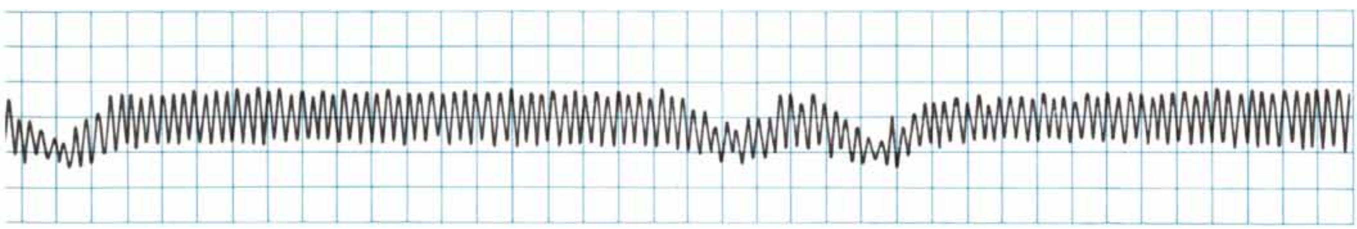
"To avoid disturbing birds I make some of my observations at a distance of 100 feet or more. To record intimate sounds at such distances I set up a directional microphone, which is the acoustical counterpart of a telephoto lens. I improvised the device by rigidly supporting an inexpensive crystal microphone at the focus of a paraboloidal dish of fiberglass about three feet in diameter. The dish is sold as a sled in toy stores.

"The amount of acoustic energy that a parabolic reflector can concentrate on a microphone varies with the pitch of the sound and the diameter of the dish. A dish three feet in diameter can in effect boost the output power of a conventional microphone about 10 times at a pitch of 180 hertz and higher. (This tone is roughly equivalent to *F* below middle *C*.) Paraboloids of larger diameter increase the response to tones of lower pitch but do not otherwise appreciably improve the performance or directivity of the microphone. Instructions for making parabolic reflectors of any desired size were given in this department in December, 1973. [The equation for z that appears in the illustration on page 127 of that issue should read $z = [l^2 + (y_1 - y)^2]^{1/2}$.]

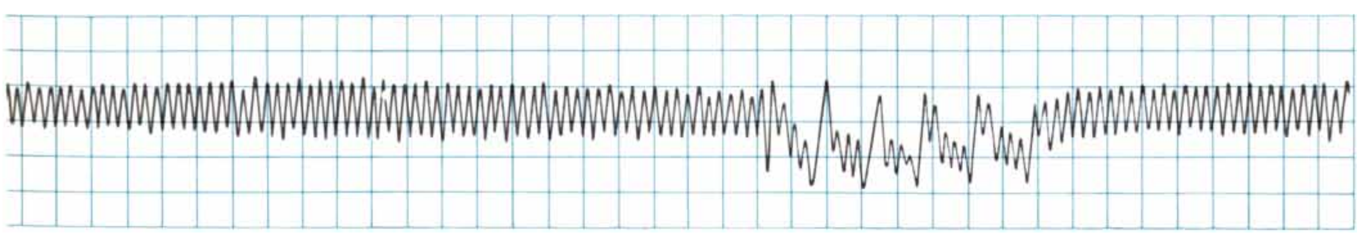
"The apparatus is exceptionally handy for ascertaining what species frequent selected ranges and also the seasons and the times of their visits. Essentially the



Graph of song of chickadee



Song of the northern Appalachian katydid



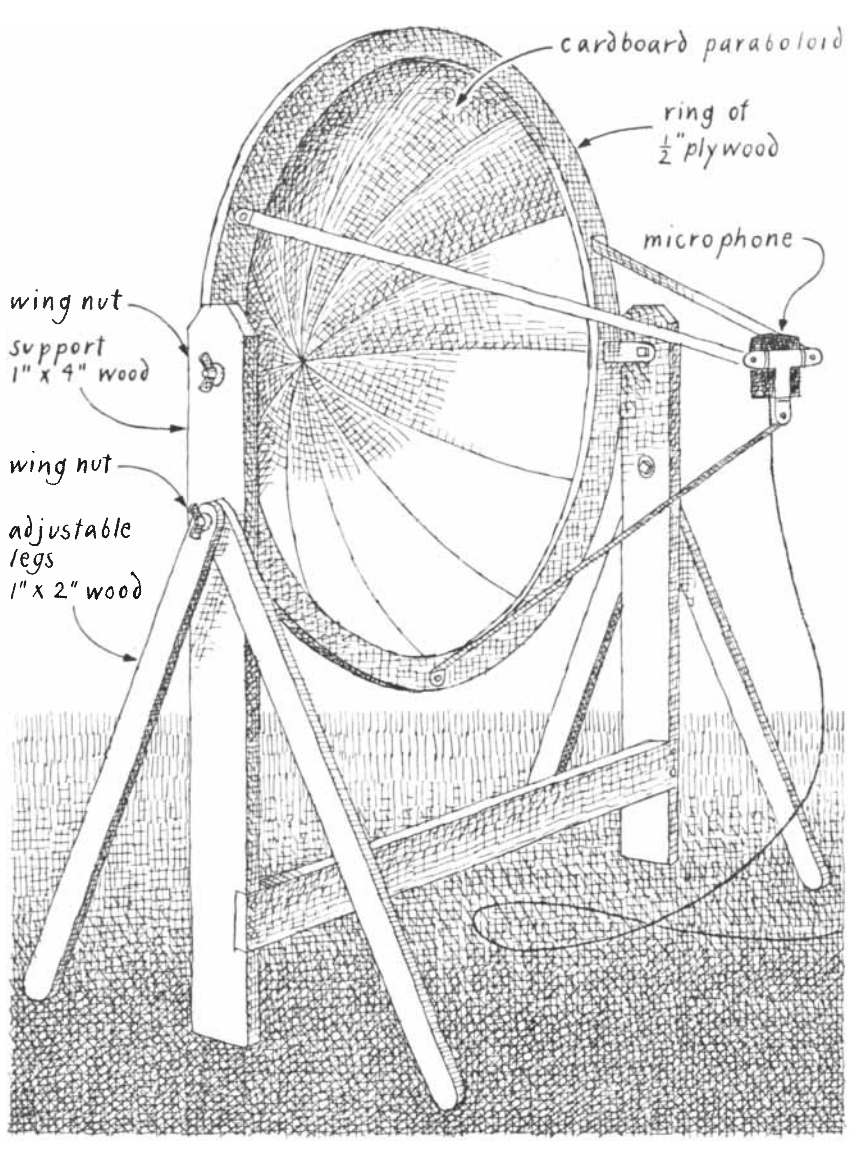
Song of the southern Appalachian katydid recorder and the camera act like a trap. One can often sequentially make close observations of 16 quarter-acre regions from a single fixed location.

"The primary advantage of the system, however, is that it enables one to intimately observe animals that become alarmed in the presence of a human being. (Have you ever tried to get close to a chirping cricket?) The complete system is battery-powered and can be set up quickly in places as different as bat caves, wilderness areas and the edges of swamps.

"Having recorded tapes, I transcribe them in the form of graphic patterns. This step can be accomplished in a variety of ways by means of sound spectrographs and related instruments. I am interested in the recordings primarily as aids in species identification. I learned by experiment that patterns of bird sounds made by pen recorders of the kind used in electrocardiographs are extremely useful for my purpose. Physicians and laboratories of physiology have the apparatus and will cooperate with the serious amateur. Most of my tapes have been transcribed with a Sandborn Visco EKG recorder.

"The output level of the reproduced signal that is fed to the pen recorder must be determined experimentally. To avoid damage to the instrument I set the volume control of the tape reproducer to zero and slowly increase the output until the recording pen swings to approximately 75 percent of its full excursion during passages of maximum amplitude. Any pen recorder of comparable performance will serve to transcribe the tapes. The paper speed should be about 50 centimeters per second. A homemade pen recorder of the type described in these columns by J. Barry Shackleford works nicely [see "The Amateur Scientist," SCIENTIFIC AMERICAN, March, 1972].

"The accompanying graphs [*preceding two pages*] depict the distinctive sound characteristics of a chickadee and two species of katydid. It is possible for the amateur to accumulate fairly rapidly a modest library of known patterns for identifying fauna of various kinds and classifying behavior. The world's largest collection of bird sounds (more than 15,000 recordings) is maintained in the Laboratory of Ornithology at Cornell University."



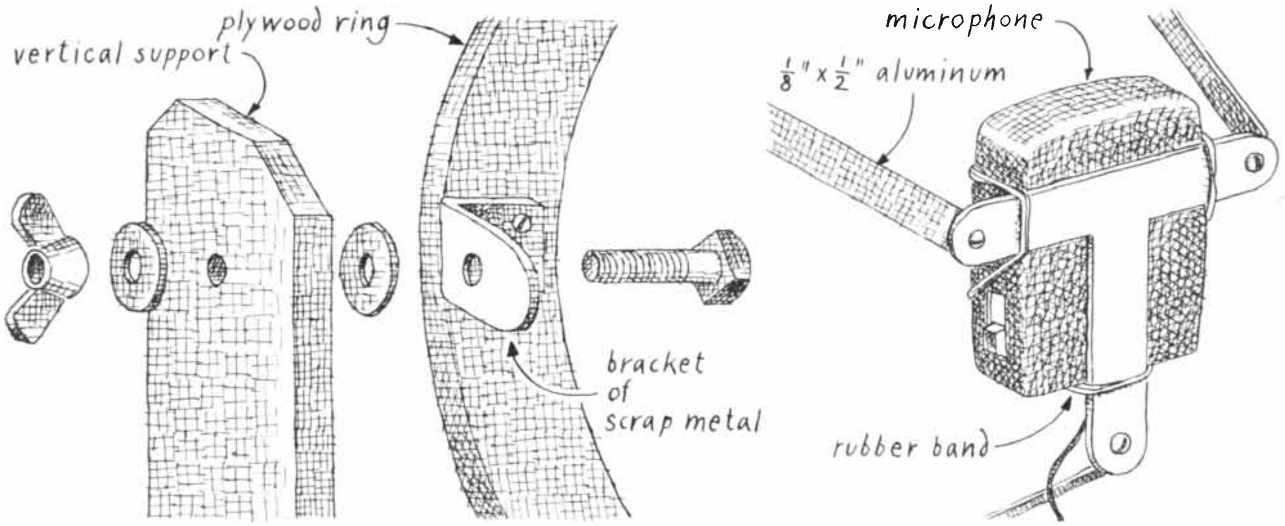
Gary S. Thorpe's directional microphone for recording birdsongs

Genie Floyd, a biology major at New College in Florida, has concluded from experiments at her home in Hartsville, S.C., that exposure to near-freezing temperatures significantly increases the retention by cockroaches of a learned routine. This came as something of a surprise, because one might have expected the chilling to have the opposite effect. "It is well known," she writes, "that the rate of a chemical reaction is accelerated by an increase in temperature and decelerated by a decrease in temperature. Biochemical reactions, and thus life processes, should be no exception, at least within reasonable limits.

"Relatively few multicellular organisms can survive appreciably elevated temperatures, but many can tolerate reductions that approach the freezing point of water. To check the supposition that exposure to low temperatures might dull an animal's memory I designed a set of experiments with cockroaches. I chose these insects because they can survive extremes of temperature and are incredibly adaptable.

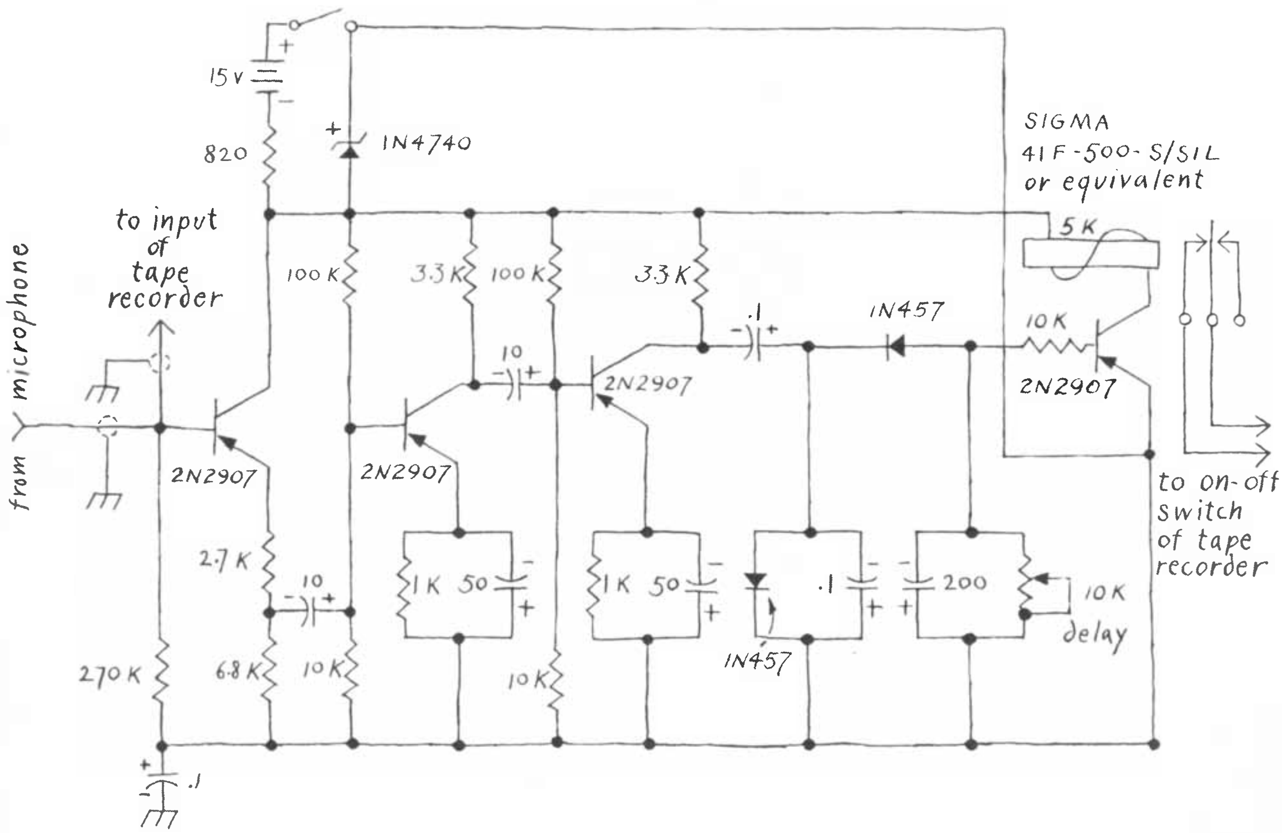
"Although cockroaches are commonly supposed to be widely available, I bought 20 specimens (*Periplaneta americana*) from the Carolina Biological Supply Company (Burlington, N.C. 27215). All were adults except for three nymphs that were near maturity. Until the experiments were made the roaches were kept in plastic cages in darkness, fed crumbled dog biscuit and watered with a tuft of wet cotton.

"The experiment consisted in training the roaches to run a maze according to a predetermined route. Before starting the experiment I separated the roaches into four groups of five individuals each. Each group was first trained to run the



Details of hinge for paraboloid and of microphone mounting

maze without error and then subjected to one of four conditioning programs. One group that served as an experimental control was allowed to rest in darkness for six hours at a temperature of 23 degrees Celsius (73 degrees Fahrenheit) immediately after it was trained. The second control group was allowed to rest in darkness at the same temperature for 12 hours. The two remaining groups—the experimental roaches—were subjected to identical conditioning except that after training one group rested in darkness for six hours at a temperature of six degrees C. (43 degrees F.) and the other rested at the same low temperature for 12 hours. Thereafter all the cockroaches remained undisturbed in their cages for one additional hour in darkness but at room temperature. This period of rest enabled the insects that had been refrigerated to recover. All the roaches were then made to traverse the maze again and were retrained as necessary



Circuitry of the triggering mechanism

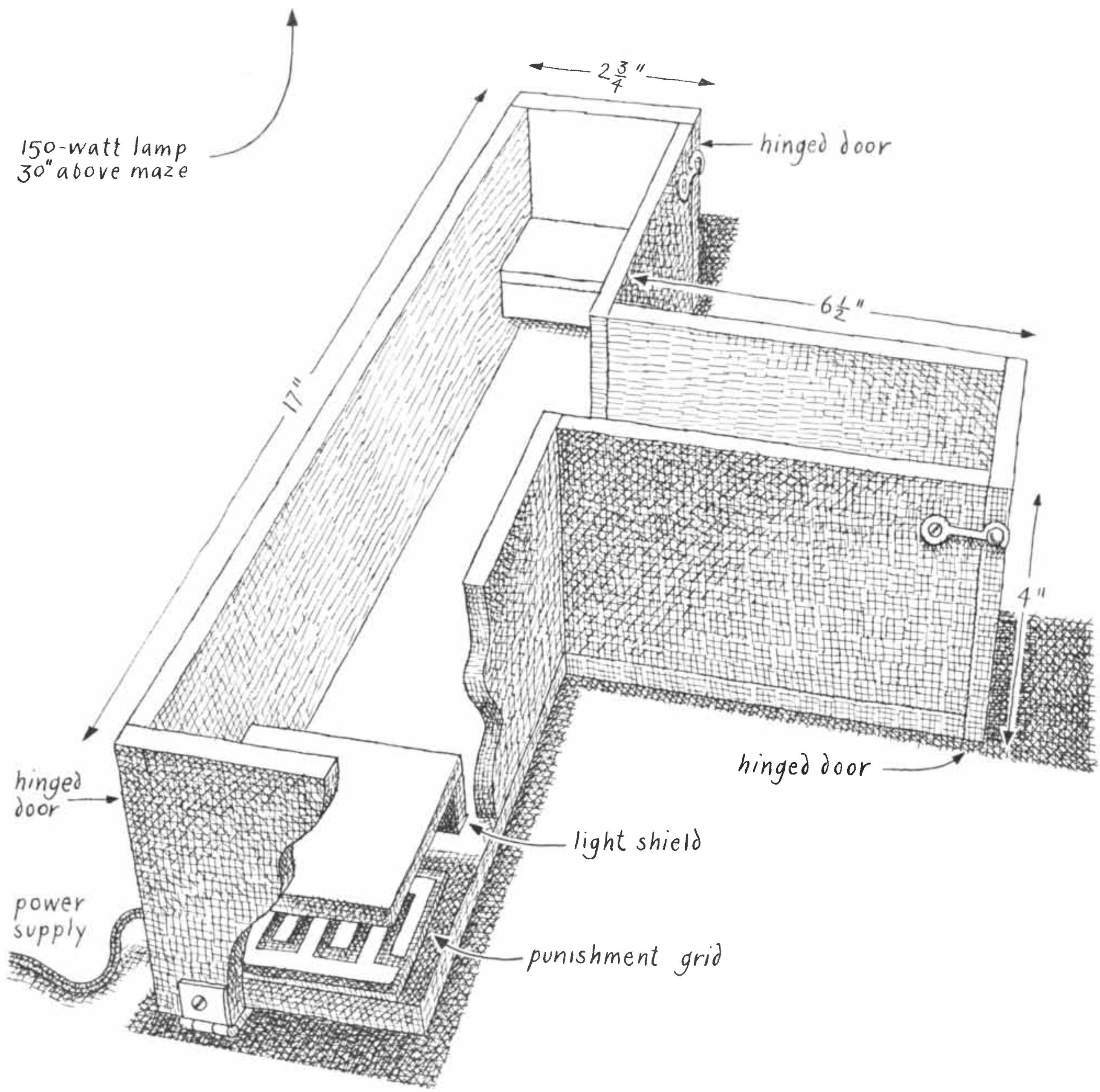
to make six successive runs without error.

"The maze consists of a small trough of plywood in the shape of a *T* [*see illustration below*]. The crossarm of the *T* contains shaded compartments at its ends, the inner surfaces of which are painted in flat black. A movable feature of the maze is a flat, rectangular grid of interlacing aluminum strips that can be electrified. The grid can be placed on the floor of either shaded compartment. A roach that steps on the metal strips receives a shock that is harmless but nonetheless punishing.

"The device consists simply of a rectangle of plywood to which I stuck a sheet of aluminum foil with doubly coated adhesive tape [*see top illustration on opposite page*]. The pattern of insulated but interlaced electrodes was cut with a safety razor blade. The plug of the electrical extension cord houses a pair of one-megohm carbon resistors. One resistor is connected in series with each conductor that energizes the metal strips. The resistors limit the maximum current that can be drawn by the apparatus, even on short-circuit, to about one ten-thousandth of an ampere.

"During experiments the maze is placed inside a cardboard box that confines the roaches even if they escape from the apparatus. Attempts to escape are discouraged by coating the inner walls of the maze and the upper walls of the box with Vaseline. The apparatus is lighted by a single 150-watt incandescent lamp suspended 75 centimeters (30 inches) above the center of the maze. Roaches can be handled without injury if the experimenter is reasonably gentle and deft.

"Roaches instinctively avoid lighted regions; as a biologist would put it, they are negatively phototaxic. I took advantage of this preference as the positive reward for desired behavior. Undesired



Genie Floyd's apparatus for experiments with cockroaches behavior was punished, or negatively rewarded, by means of the electric grid.

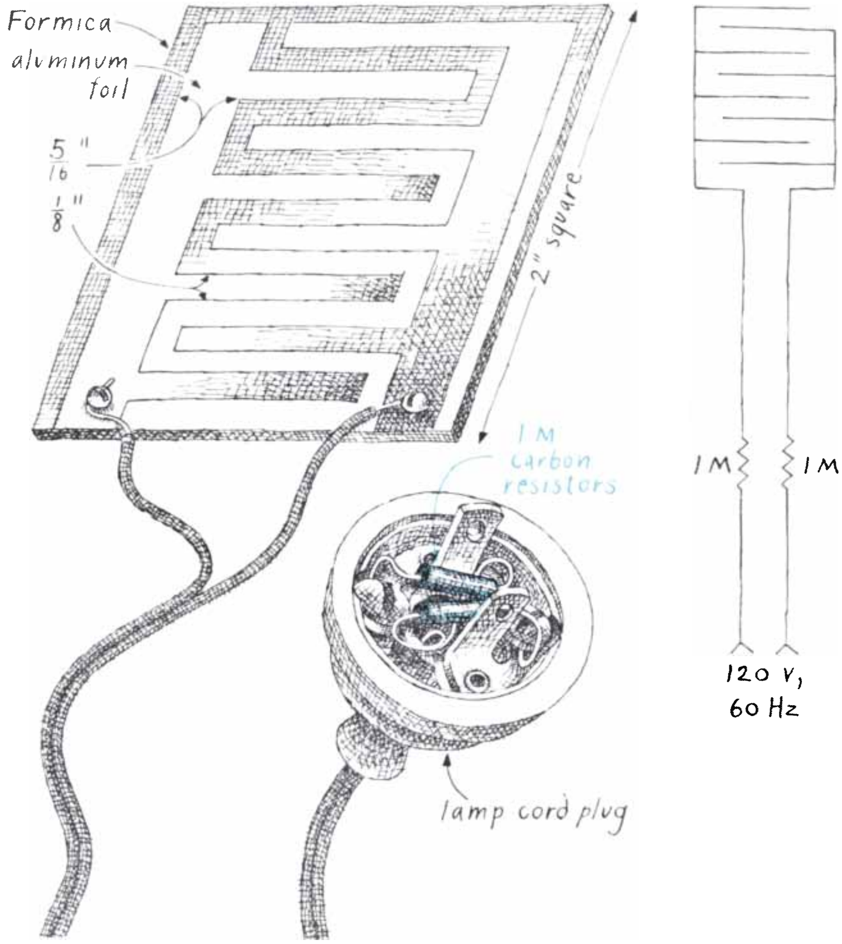
"Before I started experimenting or had installed the punishment grid I put each individual in the maze at the base of the leg of the *T* and allowed it to run to one of the dark compartments. The run was repeated three times at intervals of two minutes to determine if the roach preferred to turn right or left as it entered the crossarm. Most roaches turned left. For this reason I installed the punishment grid in the left-hand compartment.

"At this stage I separated the roaches randomly into four groups of five individuals. The individuals of each group were caged separately in small plastic bottles, where they remained in darkness without access to food for one hour. All were handled as gently as possible to minimize their excitement.

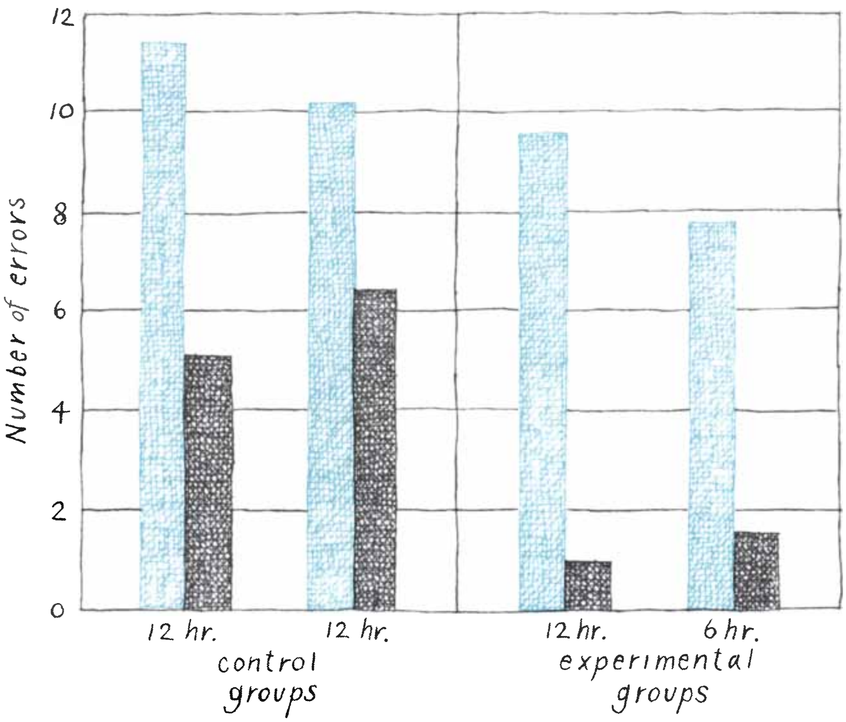
"After the punishment grid was installed the roaches were trained one at a time by placing each individual in the base of the leg of the *T*. To escape from the light the roach would scurry to the crossarm and turn left or right. Each run was tabulated. A left turn into the crossarm counted as an error. Invariably roaches that entered the left-hand compartment and received a shock promptly ran to the darkened right-hand compartment and stayed there. A roach that refused to move when placed in the maze or attempted to climb the walls was given a gentle push with a glass rod.

"A roach was considered to be fully trained when it ran without error from the base of the *T* to the right-hand compartment six times in succession. Individuals were allowed to relax in the refuge of the darkened compartment for two minutes between runs. The maze was wiped with alcohol after each roach completed its runs to remove spoor and other material that might influence the subsequent behavior of other roaches. Following this initial learning experience each individual was isolated in a plastic bottle with access to food and set aside for the interval of time at the temperatures I have mentioned and retrained as necessary. The tabulated results were then graphed.

"Contrary to my initial assumption roaches that were refrigerated after learning to run the maze made fewer errors than peers kept at room temperature. Moreover, roaches that were kept in the cold environment for 12 hours made fewer errors than those that were refrigerated for six hours. I hope other amateurs will repeat the experiments and even extend them to higher animals."



Arrangement and circuitry of the grid



Learning* (color) *and relearning* (black) *by cockroaches in the maze experiments

HOW MANY OF THESE COMPUTER BOOKS ARE IN YOUR PERSONAL LIBRARY?

Choose from books on Hardware Theory • Software Packages • Microprogramming
Data Security and Manipulation • Access Devices • Electro-optics • and much more.

PICK ANY 3 THAT YOU DON'T HAVE FOR ONLY $1.00 EACH

as a member of

THE LIBRARY OF COMPUTER AND INFORMATION SCIENCES

(You simply agree to accept 3 more books—at handsome discount—over the next 12 months.)

(retail prices shown)

39960. COMPUTER HARDWARE THEORY. *W. J. Poppelbaum.* Introduction to the theory used in designing computers and their input/output equipment. Covers advanced devices, electro-optics, noise theory, and non-linear circuits. **$16.95**

44030. ELECTRONIC COMPUTER MEMORY TECHNOLOGY. *Edited by Wallace B. Riley.* Complete analysis of the "state of the art," including evaluations of ferrite cores, thin films, MOS memories, much more. **$14.50**

50480. GPSS PRIMER. *Stanley Greenberg.* The first and only book on this exciting new simulation tool pioneered by IBM. An application-oriented (rather than machine-oriented) simulation language which features an elaborate software package. **$14.95**

40021. COMPUTERS: Auditing and Control. *Edited by Elise G. Jancura and Arnold H. Berger.* A big book that details every aspect of computer control, data security and the evaluation of the results of computerized activities. **$14.95**

42271. DIGITAL SYSTEMS: Hardware Organization and Design. *Frederick J. Hill and Gerald R. Peterson.* A refreshingly different analytical approach to computer organization and digital system design. Covers hardware, software, micro-programming, intersystem communications, more. **$17.95**

40100. COMPUTER SEMANTICS: Algorithms, Processors and Languages. *John A. N. Lee.* Brilliant extension of computer semantics to cover not only computer language study but the total description of the digital computer as well. **$16.95**

38160. CIRCUIT DESIGN OF DIGITAL COMPUTERS. *Joseph K. Hawkins.* Develops, simply and logically, the principles, applications and limitations of circuit design for digital logic and memory elements. Demonstrates how circuit behavior constitutes a compromise between digital performance requirements and basic material properties. **$21.50**

39880. COMPUTER DICTIONARY AND HANDBOOK. *Charles J. Sippl and Charles P. Sippl.* More than 14,000 individual definitions and explanations of computer terms, acronyms, and abbreviations make this a must for any computer professional. *Plus* a handbook section with in-depth articles on every aspect of the computer and information sciences. **$16.95**

40061. COMPUTER SECURITY MANAGEMENT. *Dennis Van Tassel.* "Must" reading for all computer professionals. Pinpoints the responsibilities of computer center employees and details the how-to's of security implementation. **$10.50**

32270. ADVANCED PROGRAMMING. *Harry Katzan, Jr.* For those thoroughly familiar with computer-language concepts and logic theory. Summarizes a wide range of recent literature written by creators of basic software systems. **$13.95**

(values as high as $56.40)

33430. THE ANALYSIS OF INFORMATION SYSTEMS. Revised Edition. *Charles T. Meadow.* This updated edition by a distinguished computer authority offers an all-inclusive treatment of the four aspects of information processing and retrieval. **$14.95**

61360. MATHEMATICAL STATISTICS WITH APPLICATIONS. *William Mendenhall and Richard L. Scheaffer.* Stresses inference as the objective of statistics and emphasizes the role of probability in inference making. Covers statistical theory, hypothesis testing, linear models, and much, much more. **$12.95**

36080. THE BEST COMPUTER PAPERS OF 1972. *Edited by Isaac L. Auerbach.* These 14 outstanding articles and papers were selected from the thousands that were published last year by six of the most prestigious computer organizations in existence today. **$14.95**

The Library of Computer and Information Sciences 7-37F
Riverside, New Jersey 08075

Please accept my application for membership and send me the three volumes indicated, billing me only $1.00 each. I agree to purchase at least three additional Selections or Alternates during the first year I am a member, at special members' prices. Savings range up to 30% and occasionally even more. My membership is cancelable anytime after I buy these three books. A shipping and handling charge is added to all shipments.

3 books for $1.00 each

Indicate by number the 3 books you want.

[ ] [ ] [ ]

Send no money. Members are billed when books arrive.

Name____________________

Address____________________

City
State__________ Zip__________

Book selections purchased for professional purposes may be a tax-deductible expense.

48680. FLOWCHARTS. *Ned Chapin.* Comprehensive, practical study of the use of graphic techniques to describe and interpret the operation of computers. Loaded with examples from actual computer situations. **$9.95**

42261. DIGITAL COMPUTING AND NUMERICAL METHODS. *Brice Carnahan and James O. Wilkes.* The authors, best remembered for the highly popular *Applied Numerical Methods,* here offer a brilliant presentation of numerical methods with practical application to their implementation on digital computers. **$13.95**

41670. DATA PROCESSING SYSTEMS DESIGN. *H. D. Clifton.* By the author of the best-selling *Systems Analysis for Business Data Processing,* this well-researched new volume contains a series of case studies which will enable readers to put these examples of good systems design to practical use. **$9.95**

69401. PL/I PROGRAMMING. *Joan K. Hughes.* PL/I has synthesized almost all that has been known about languages useful for solving numerical scientific and business data processing problems or combinations of them. There *is* no better introduction to this unique programming language. **$12.95**

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, a comprehensive list of abbreviations, acronyms and symbols and over 100 pages of data tables. Filled with useful information discriminatingly selected and skillfully arranged and cross-referenced. **$14.50**

HOW THE CLUB OPERATES

• The Book Club **News,** describing the coming Main Selection and Alternate Selections, will be sent to you 15 times a year at three to four week intervals. • If you wish to purchase the Main Selection do nothing and it will be shipped to you automatically. • If you prefer one of the Alternates, or no book at all, simply indicate your decision on the reply form always enclosed with the **News** and mail it so we receive it by the date specified. • The **News** is mailed in time to allow you at least 10 days to decide if you want the coming Main Selection. If, because of late mail delivery of the **News,** you should ever receive a Main Selection without having had the 10-day consideration period, that Selection may be returned at Club expense. • After completing your trial membership, you will be entitled to take advantage of our fabulous bonus plan which offers savings up to 70% and more. (Offer good in Continental U.S. and Canada only. Prices slightly higher in Canada.)

BOOKS

Humboldt deserved having so many things named after him

by Philip Morrison

HUMBOLDT AND THE COSMOS, by Douglas Botting. Illustrated. Harper & Row, Publishers ($15). The modern reader will find the name of a Prussian baron affixed to three inches of entries in the gazetteer of any good atlas: towns, mountains and capes from New Zealand to Nevada. This extremely handsome and readable volume seeks to recapture for us what that name meant in the long years between the French Revolution and the year of *Origin of Species,* when Humboldt died one spring afternoon, heavy with years and fame, at peace with the world although impoverished. His last written words were the phrase from Genesis: "Thus the heavens and the earth were finished, and all the host of them." The fifth volume of his *Cosmos* had gone to the publishers three weeks earlier, the last of a score of volumes he wrote, seeking unity in a world he had seen in all its diverse aspects.

One two-page color plate sums up the image, although not the man. It is a painting by a German artist of "Humboldt and Bonpland in their hut in the forests of the Orinoco." Outside we see palms and strange distant peaks; within, the shadowy thatched walls are decked with hanging skins and specimens and the work table is cluttered with a microscope, a theodolite, a compass and a sextant. The botanist Aimé Bonpland is holding the hand glass he has used to examine the flower in his left hand; the young Humboldt sits with a map on his knee. Across the packing cases that furnish the hut jungle fruits and flowers are strewn, a bright parrot perches on the pack frame and the great brass magnetic dip circle stands solidly in the background. The meaning is clear: however remote and surrounded by the wonders of nature, these two are no wanderers, no mere explorers, but men who are engaged in consciously extending to all the world the network of order and understanding that was the content of natural science.

The painting does not mislead. Baron Friedrich Wilhelm Karl Heinrich Alexander von Humboldt discovered little (although he was probably the first to publish the fact that the earth's magnetic field decreases with latitude and reaches a minimum in the Tropics). It was his accomplishment, rather, to carry beyond far-off jungle rapids and to the tops of snowy peaks the instruments and the concerns of science. "I wish I could give an adequate idea of the exquisite beauty of the nights," he wrote from Venezuela on his first great voyage. "Quite often I have been able to read the vernier of my little sextant by the light of Venus." Everywhere he went his sextants read the latitude, his chronometer the longitude, his famous enclosed mercurial barometer the altitude, his heavy magnetic instruments the state of the earth's field and his thermometers the temperature of air, land and sea. His thermometer dipped into the cold current off the coast of Peru; every local fisherman and Spanish voyager knew the northward set of the cold sea-stream, but Humboldt first fixed it, recorded it—and so justly earned for it the name it bears today.

His career not only bound together the zones of the world, from nearly four miles high on Chimborazo to the gold and platinum placer mines of the Urals (where his prediction that diamonds would be found was confirmed in a matter of weeks), but also held together the history of Europe as well. Born in the same year as Bonaparte, he lived to be 89; he ventured forth first by stage and sail and later by steamboat and railway. When the American writer Bayard Taylor visited him in the last years of his long life, Taylor wrote: "I pressed the hand which had touched those of Frederick the Great, of Forster, of Klopstock and Schiller, of Pitt, Napoleon, Josephine, the Marshals of the Empire, Jefferson, Wieland, Herder, Goethe, Cuvier, Laplace, Gay-Lussac"—and Taylor omitted Gauss, Arago, Metternich and Bismarck.

To his encyclopedic knowledge of science he added a consistent and courageous devotion to the radical ideas of the French Revolution. As a youth in the "philistine and provincial" capital of reactionary Prussia he frequented the disenfranchised circle of the Jews Marcus Herz and Moses Mendelssohn, "with their emancipated women, modern science and avant-garde poetry." He spent a few days in revolutionary Paris in 1790, a young liberal of 20, with Georg Forster, the author and naturalist who later served in the provisional government of Republican France. Humboldt remained true to his views throughout his life. When he was the young Chief Inspector of Mines for Prussia, he refused an award, urging the state to set up pensions for injured miners instead. He established with his own funds the first training school for workers in Germany, the Free Royal Mining School in Steben. He loathed slavery and railed against it. In the Andes he refused to be carried on a chair strapped to the back of an Indian porter, the usual transport of the traveling priests of the region, and he rejected the paternalism with which they justified their mission. "The Indians of the Orinoco are not great children; they are as little so as the poor labourers in the east of Europe, whom the barbarism of our feudal institutions has held in the rudest state."

He lived and worked the decades of his greatest fame in Paris, first under the Directory and then under the Empire. Only after the second Peace of Paris, his fortune spent on his voyages and on their lavish publication (he subsidized scholars and artists to prepare the treatises that made collections and data a real part of science), did he return to Prussia to live the rest of his life on his earnings from books and on a pension from his king. "I live in . . . the romantic affection of a noble prince, yet in a moral solitude such as only the stunted intellects of this dour and divided country can provide." He served the king as a walking encyclopedia, took lunch with him, read to him daily, answered every question, corrected the king's letters—and fled each year to Paris. Too famous to dismiss, in extreme old age he could no longer hold his court listeners and "found himself reading his liberal and uplifting passages to an audience of precisely one": an intelligent if reactionary Junker named Bismarck.

Humboldt was a man of warmth and courtesy. Hear Dolly Madison: "All the ladies say they are in love with him, notwithstanding his want of personal charms. He is the most polite, modest, well-informed and interesting traveler we have ever met." It is true he was modest, but he was eager enough for praise and repute. No lady, it seems, was ever in love with him, not even the young, dark-eyed, witty, dazzling beauty and flirt Henriette Herz, Marcus' wife. In his youth she had welcomed him and his linguist brother to her circle, giving him "the first taste of female warmth and comfort he had ever experienced." (He got none from his aloof, self-sufficient widowed mother.) All his life he formed passionate attachments to young men of brilliance; it seemed, as his sister-in-law said, he could "never be inspired by anything that does not come through men." His sexual ambivalence is documented by some remarkable letters that give accounts of dreams and impossible attachments. It was François Arago, the brilliant physicist and a devoted family man, who remained the central passion of Humboldt's life for nearly 50 years. Arago was far less free to return Humboldt's affection than Humboldt was to proffer it, and the older man "had to make do with crumbs—crumbs which... became the main sustenance of his personal life."

In 1848 they buried in Berlin the 183 heroes of the barricades, slain by the king's troops in the hopeless failure of the uprising. The 78-year-old Humboldt, chamberlain to the king, walked bareheaded in the March wind at the head of the cortège past the royal palace. "I have saluted 1789," the old Jacobin scientist wrote, "and now I am sad to say... I am reduced to the banal hope that the noble and ardent desire for free institutions is maintained in the people and that, though... it may appear to sleep, it is as eternal as the electromagnetic storm which sparkles in the sun."

Botting is an English explorer and photographer. He has traveled the Orinoco in the footsteps of Humboldt, but his is not the scientist's interest. We can thank him for a biography both interesting and penetrating, particularly in its account of events and psychology. One misses a chapter or two on Humboldt's instruments and their results, more graphs, charts, theories, conclusions. It is nonetheless a book well worth reading, and made most attractive by the designer, the illustrator and the printers in Britain and Japan.

PARADISE LOST: THE DECLINE OF THE AUTO-INDUSTRIAL AGE, by Emma Rothschild. Random House ($6.95). A hundred million cars and trucks now on American roads celebrate two men: Henry Ford and Alfred Sloan. Yet "when I grew up," said one dealer in 1970, "the automobile was the one product that universally signified success in America. What has happened on the way to Paradise?"

The works of the prophets are still manifest. Fordism is the technique of Model T mass production scrupulously carried to its limit: all the skill reverts to the machines. The workers must ask no questions but follow the rhythm of the line; they need possess no skill and no talent except coming to work on time. Sloanism is based on ever elaborated marketing of new models, not Ford's "any color so long as it is black" but rather the "splash of jewel-like color in every parking lot" that Sloan himself described.

The new GM plant in Lordstown, Ohio, where they make Vegas, is the principal automotive specimen dissected in this brief, cogent and fascinating book. Obedient to Fordism, the automation placed in the modern plant served only to make notorious the speedup of the workers, whose resentment at the 36-second rhythm flared into repeated strikes. The clever body-welding automata are given less skilled human helpers, who monotonously prepare the workpieces for the robot welders. It is cheaper so. As for Sloanism, you can buy "an El Camino-type Vega, V8 Vega, sunroof Vega, super-plush Vega, turbo-Vega, convertible Vega, and a four-door Vega... all concocted to see if it will sell." Thus does GM today market its subcompact "cheap but nonascetic car."

Fordism and Sloanism have become complementary. The logistics of automobile assembly are justly famed: the blue fender always rides in from the side just as the blue body glides past. The bewildering combinatorics are planned, however, outside the line. The assembly itself can be managed by workers and inspectors without any choices of their own; they can practice only a go or no-go trade. The cultivation of market diversity at the same time inhibits two opposite remedies: replacing the post-World War I Ford style of line with higher technology—continuous flow and truly automatic methods—and restoring craft skills and responsibilities. Only in the engine factories, which cope with fewer consumer choices, is there high automation and "a totally different ambience" from the unskilled, repetitive work of the parts handlers. The dream of plastic bodies made in one part or only a few parts by an automatic plant resembling the continuous-flow processes of the chemical industry, tended by a few responsible experts, is blocked by the Sloanist excesses: too many add-ons, shapes, colors, styles.

The industry is baroque, internally and externally. It has filled the streets and laid down the highway net, the malls and the sprawling suburbs. It has reached saturation, yet it struggles to remain as it was when cars meant personal freedom. Its growth depended, like the railways of the 19th century, on "social and institutional partiality" that provided roads, favorable tax policies and the dispersal of jobs and residences. Only now do we begin to see true costs, the wastes, the lowered margin of social return, particularly in cities. More than 80 percent of all commuters now drive to work; will we ever see $1,000 optionless city cars, with simplified fiberglass bodies, regularly left for rented use by the next person who comes to the parking lot? How that flies in the face of the ancient teachings of Ford and Sloan!

The world is one Detroit. Even the clever computerized plants of Nagoya have quickly saturated their roads; even if they now avoid Fordism's extremes by providing better working conditions, they too begin to talk of trading up for bigger models. As time passes, however, the social bias for the automobile will everywhere erode, and we shall once again see "auto transport as one way among many, a particularly costly way, of spending and wasting resources."

This compact argument is brilliantly presented. Emma Rothschild, a young British journalist with American graduate training in economics, has managed to put before the reader the most penetrating account in years of a large social issue, and she has done it with grace and style, logic tight, vision wide. There are some flaws in the gem. A scientific reader will miss the maps, graphs and tables the subject cries for. There is not even one; the book is aimed at readers from the other culture. There is also a faint air of insecurity about technical issues. A footnote in the book asserts, for instance, that the reclamation of silver in photographic processing, which is now commonplace in the U.S., is an expression of unusual and rigorous Chinese frugality! Let such points pass; the argument remains, bright and straight as a headlight beam.

PALAEOETHNOBOTANY: THE PREHISTORIC FOOD PLANTS OF THE NEAR EAST AND EUROPE, by Jane M. Renfrew. Columbia University Press ($20). Once in old Syracuse, in about the fourth century B.C., three heavy ears of six-rowed barley were presented as a votive offering. They remain this day as golden as they were when the sacred altar first received them, since they are not grain at all but cunningly worked replicas in the noblest metal. The frontispiece of this treatise shows them in color, with thanks to the fortunate American collector who holds them today. It is only the most remarkable of the hundreds of striking images that enlarge this elegant book beyond the Cambridge thesis from which it began.

Dr. Renfrew means by her title "the study of the remains of plants cultivated or utilized by man in ancient times." Food plants and drugs, seasonings and the like are her topic; plants with other, less common uses lie outside the work. Her field is Europe and the Near East; she cites work done elsewhere, but she does not deal in any detail with rice, maize, sorghum or the roots that have fed mankind on other continents. The bulk of the work describes in text and drawing, and in detail adequate for botanical identification, the most important species: cereals, pulses, flax, fruits, nuts, opium and cannabis, and a wide variety of weeds, such as fat hen and gold of pleasure, whose seeds or pods are often found in prehistoric contexts.

That Midas barley is unique. The material the paleobotanist examines under her low-power microscope is normally preserved less artistically (if more precisely) by carbonization, by waterlogging in a peat bog, by detailed impressions left (usually by accident) in wet clay that was later baked or, more rarely, preserved as a silica skeleton of the plant wall or dried in the arid atmosphere of some desert tomb. In the Altai mountains a pile of hemp seeds was found associated with a Scythian copper censer, all under a miniature tent covered with a felt rug. It was the ice of those cold Siberian mountains that preserved this system, which was described in *Scientific American* by M. I. Artamonov in 1965. Herodotus certified the scheme for a high: he judged the cannabis-containing vapors were "unsurpassed by any vapor bath one could find in Greece. The Scythians enjoy it so much that they howl with pleasure." So pointed a reference is not frequent in this book, which deals primarily with the days before men knew how to write, when they first domesticated the plants their forebears had gathered.

Long, patient examination of varied samples is the taxonomist's talent. Few tools beyond experience and the dissecting microscope are needed. A single plate of scanning electron micrographs shows the distinctive fine texture of the skin coat of modern grains; someday this technique may become important. Carbonized grain endures, sometimes entire ears of it, or even a crab apple or two, preserving minute detail. Flotation of the dusty sample in water or a heavier fluid separates the grains. Some change in shape occurs on carbonization; there are histograms of measurements of modern grains before and after baking 12 hours in a gas oven at 200 degrees Celsius. The grains shrink in length about 10 percent; barley changes the least, since its grain is tightly wrapped by a leaflike structure. The seeds have been carbonized by some mishap; perhaps they were once spilled on the hearth, scorched in the parching pan or singed when the dwelling burned. Identification, grain by grain, and careful counting can trace out the development of agriculture. The most frequent crop seed is set against the number of other species. One can watch the dominance of a desirable species grow as confidence increases.

In the Cycladic Islands they still sow a mixed crop, maslin, in which a single field produces both cereals and pulses: "oats, peas, beans and barley grow," and wheat as well. The author saw such fields growing in the summer of 1965. This practice is plainly a fail-safe mixed strategy, to ensure some return whatever the weather. By the days of the early New Kingdom, God expected a less conservative husbandry from his people; he spake to Moses forbidding the sowing of mingled seed. The settlement of Tell Mureybit in northern Syria is "the only permanent settlement site so far found in the Near East which can be shown to have subsisted on wild cereals." Its grain staples were the diploid wild einkorn and the wild two-row barley; its radiocarbon age is almost 10,000 years.

Man does not live by bread alone. Lentils are widely found. The familiar species of the pistachio nut, which is among "the most drought-resistant fruit trees known," has been found only once, in a late Neolithic level in Thessaly, but a smaller species is found at many older sites, and is gathered still for sale in the bazaars. The wild grape appears to differ from the cultivated vine in having shorter pips; at one Macedonian site where Dr. Renfrew worked, the grape pips lengthened over a 3,000-year period. Sorrel and field cabbage are found as seeds but may have been used as salad greens, and mustard, anise and coriander suggest the seasoning of food at least as early as the late Bronze Age. Apple, olive, fig, raspberry, blackberry, sloe and wild plum are all found in Neolithic Europe or the Near East. We still eat, one concludes, largely in the Neolithic, even though we mix the Old and the New worlds.

Although most of the text is aimed at the meticulous identification of species, Dr. Renfrew has so well organized and summarized the methods, history and results in her field that the general reader will glean a great deal from the edges and corners of the work. It was the dry winters of 1853–1854 that revealed the submerged, waterlogged villages of the Swiss lake dwellers. Oswald Heer published his study of their crops in 1866; a page of his pioneer illustrations is reproduced. By now his science has become well established, and we can expect more in the future with new tools and new diggings. Dr. Renfrew herself foresees "setting the early beginnings of agriculture back well before 10,000 B.C." in the Near East. Southeast Asia cannot have been far behind.

CONTRACEPTION, edited by L. L. Langley. Dowden, Hutchinson & Ross, Inc. ($22). GEOCHRONOLOGY: RADIOMETRIC DATING OF ROCKS AND MINERALS, edited by C. T. Harper. Dowden, Hutchinson & Ross, Inc. ($24). The most logical of libraries finds no escape from storing on the same shelf works on cheese and treatises on chalk—provided that both are oversize folios. The physical embodiment of the symbol commands, and mere meaning usually must obey. The publishing of books is likewise a physicosocial process, overriding semantics. Both of these two excellent but disparate volumes belong to a new and ambitious series of anthologies, collected "Benchmark Papers" in 20 different branches of science, pure and applied. Each discipline will in some years possess a dozen to 20 volumes of its key papers arranged by subfield, praiseworthy in every way, except perhaps for price.

Anthologies do not mark the progress

of science; they look rather to its past, and about them hangs a faint odor of decline. A good collection—these are excellent—presents a wide range of selection by an editor with knowledge and taste who is willing, moreover, to offer a few pages of his own comments to surround with context every group of related papers. The reader cannot miss a sense of unraveling of issues and growing power in such a series of outstanding papers. The books present 40-odd excerpts each. The papers are generally in direct facsimile but most contributions that were not originally in English are given in translation. Some selections come from books. It is usually not easy for an outright beginner in a field to read its original literature, even at a late date. Such books as these should supplement an explicit text or add unique enrichment for a reader who has fair familiarity with the discipline. Libraries that lack long runs of old periodicals, and serious students who prefer a guided tour to a long pathless ramble, will get real value from them. There are indexes of both subjects and authors in these books and generally careful editing citations, although with a few apparent errors and misprints.

The volume on contraception is the second in a series on human physiology. This topic of such profound importance begins with Genesis, with the method of Onan, who declined to impregnate his brother's widow Tamar, and "spilled [his seed] on the ground." Malthus' famous essay appears, as does a rejoinder by Francis Place, a self-taught workingman who became the founder of the birth-control movement in England. Writing in 1822, he is a courageous optimist. Not celibacy, starvation or long-deferred marriage but legalized trade unions, free emigration, repeal of duties—particularly on grain—and contraception are the "means of preventing the numbers of mankind from increasing faster than food is provided." "If a hundredth, perhaps a thousandth, part of the pains were taken to teach these truths that are taken to teach dogmas, a great change for the better might... be expected," he wrote.

The next papers range from the Latin of Fallopius through the development of rubber latex. Most of the rest of the volume is more directly medical, dealing with the surgery of sterilization in men and women, the physiology and testing of a variety of barriers to the sperm and modern interference with the reproductive cycle, mechanically by the IUD, by way of the steroid hormones with "the pill" or—still tentatively—with the prostaglandins. The verified inability of spermatozoa to survive normal body temperature has never led to practice; consider the intolerably hot baths of Japanese tradition. The volume ends with a 1972 paper reporting a stratified-sample survey: in the U.S. in 1970 about a third of the couples of reproductive age were not using contraception. Of the two-thirds who were, some 80 percent were "highly protected from the risk of unintentional conception... both blacks and whites and fairly uniformly... couples of widely varying educational levels."

The second volume begins not with ancient scripture but with a portion of Ernest Rutherford's book of 1906. For two years before that Rutherford liked to walk about the McGill campus with a rock in his pocket, which he would produce with the remark that he knew how old it was! The root idea of dating minerals by measuring the accumulation of the products of radioactive decay was already his. The latest paper presents dates from 1971. It details the complex selenological history of the first Apollo lunar samples, unraveled with the full arsenal of the uranium-lead, thorium-lead and rubidium-strontium ratios and even with cosmic ray particle tracks. (The absence of a paper on the etched-track methods of recent years is the only gap one sees in the volume.) The key papers that first applied and explained all these radiogenic methods are presented, with the emphasis throughout on application and results. The book is centered on geochronology and not on counting and radiochemical techniques. Here is the 1968 test set for continental drift, the radiogenic ages mapped for the rocks of all continent edges to challenge the fit of the crustal jigsaw puzzle. The main themes of the papers are the ramifications of the simple idea of Rutherford's, the variety of nuclei now used, the use of isotope-spiking and mass spectrometry and ingenious ways of extrapolating discordant measurements to extract statistically ages no single sample can give. Once again a more didactic treatment would probably provide a better introduction, but some of these papers are very pleasing complements, and the entire book is a rich vein for readers in or close to the discipline. Those gneisses from Godthaab in West Greenland, reported in 1971 to be 4 ± .2 billion years old and therefore much the oldest rocks known, do not escape citation in one of the editor's helpful commentaries.

INDEX OF ADVERTISERS

FEBRUARY 1974

BIBLIOGRAPHY

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

SOYBEANS

SOYBEAN BLUE BOOK (ANNUAL). American Soybean Association, Hudson, Iowa.

SOYBEAN DIGEST (MONTHLY). American Soybean Association, Hudson, Iowa.

MODERN SOYBEAN PRODUCTION. Walter O. Scott and Samuel L. Aldrich. The Farm Quarterly, Cincinnati, 1970.

ECONOMIC PRODUCTION POSSIBILITIES OF SOYBEANS IN NORTHERN INDIA. Folke Dovring, J. R. Jindia and R. S. Misra. Office of International Programs and Studies, University of Illinois, 1972.

SOYBEANS AND THEIR PRODUCTS. James P. Houck, Mary E. Ryan and Abraham Subotnik. University of Minnesota Press, 1972.

CHARGE-COUPLED DEVICES

CHARGE COUPLED SEMICONDUCTOR DEVICES. W. S. Boyle and G. E. Smith in *The Bell System Technical Journal,* Vol. 49, No. 4, pages 587–593; April, 1970.

EXPERIMENTAL VERIFICATION OF THE CHARGE COUPLED DEVICE CONCEPT. G. F. Amelio, M. F. Tompsett and G. E. Smith in *The Bell System Technical Journal,* Vol. 49, No. 4, pages 593–600; April, 1970.

CHARGE COUPLED 8-BIT SHIFT REGISTER. M. F. Tompsett, G. F. Amelio and G. E. Smith in *Applied Physics Letters,* Vol. 17, No. 3, pages 111–115; August 1, 1970.

THE BURIED CHANNEL CHARGE COUPLED DEVICE. R. H. Walden, R. H. Krambeck, R. J. Strain, J. McKenna, N. L. Schryer and G. E. Smith in *The Bell System Technical Journal,* Vol. 51, No. 7, pages 1635–1640; September, 1972.

SLOW, INAPPARENT AND RECURRENT VIRUSES

MEASLES ANTIBODY IN MULTIPLE SCLEROSIS. J. Adams and D. Imagawa in *Proceedings of the Society for Experimental Biology and Medicine,* Vol. 111, pages 562–566; 1962.

PERSISTENT INFECTION OF CELLS IN CULTURE BY MEASLES VIRUS. Robert Rustigian in *Journal of Bacteriology,* Vol. 92, No. 6, pages 1792–1811; December, 1966.

INFECTION AS THE ETIOLOGY OF SPONGIFORM ENCEPHALOPATHY (CREUTZFELDT-JAKOB DISEASE). C. J. Gibbs, Jr., and D. C. Gajdusek in *Science,* Vol. 165, No. 3897, pages 1023–1025; September 5, 1969.

EFFECTS OF VIRUS INFECTIONS ON THE FUNCTION OF THE IMMUNE SYSTEM. Abner Louis Notkins, Stephan E. Mergenhagen and Richard J. Howard in *Annual Review of Microbiology,* Vol. 24, pages 525–536; 1970.

INFLUENZA VIRUS EFFECTS ON CELL MEMBRANE PROTEINS. John J. Holland and E. D. Kiehn in *Science,* Vol. 167, No. 3915, pages 202–205; January 9, 1970.

PERSISTENT AND SLOW VIRUS INFECTIONS. John Hotchin in *Monographs in Virology: Vol.* 3. S. Karger, 1971.

MEASLES ANTIBODIES IN THE CEREBROSPINAL FLUID OF PATIENTS WITH MULTIPLE SCLEROSIS. Paul Brown, Françoise Cathala, D. Carleton Gajdusek and Clarence J. Gibbs, Jr., in *Proceedings of the Society for Experimental Biology and Medicine,* Vol. 137, No. 3, pages 956–961; July, 1971.

LATENT HERPES SIMPLEX VIRUS IN SPINAL GANGLIA OF MICE. Jack G. Stevens and Margery L. Cook in *Science,* Vol. 173, No. 3999, pages 843–845; August 27, 1971.

CHRONIC VIRAL INFECTIONS OF THE CENTRAL NERVOUS SYSTEM. Luiz Horta-Barbosa, David A. Fuccillo and John L. Sever in *The Journal of the American Medical Association,* Vol. 218, No. 8, pages 1185–1188; November 22, 1971.

INHIBITION OF ANTIBODIES TO NUCLEAR ANTIGEN AND TO DNA IN NEW ZEALAND MICE INFECTED WITH LACTATE DEHYDROGENASE VIRUS. Michael B. A. Oldstone and Frank J. Dixon in *Science,* Vol. 175, No. 4023, pages 784–786; February 18, 1972.

VIRIONS FROM PROGRESSIVE MULTIFOCAL LEUKOENCEPHALOPATHY: RAPID SEROLOGICAL IDENTIFICATION BY ELECTRON MICROSCOPY. J. B. Penney, Jr., L. P. Weiner, R. M. Herndon, O. Narayan and R. T. Johnson in *Science,* Vol. 178, No. 4056, pages 60–61; October 6, 1972.

SCRAPIE AGENT AND NEURONES. H. K. Narang, B. Shenton, P. P. Giorgi and E. J. Field in *Nature,* Vol. 240, No. 5376, pages 106–107; November 10, 1972.

PARAMYXOVIRUS-LIKE PARTICLES ASSOCIATED WITH ACUTE DEMYELINATION IN CHRONIC RELAPSING MULTIPLE SCLEROSIS. John Prineas in *Science,* Vol. 178, No. 4062, pages 760–763; November 17, 1972.

PROPHYLAXIS AND IMMUNIZATION IN MICE BY USE OF VIRUS-FREE DEFECTIVE T PARTICLES TO PROTECT AGAINST INTRACEREBRAL INFECTION BY VESICULAR STOMATITIS VIRUS. Michael Doyle and John J. Holland in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 70, No. 7, pages 2105–2108; July, 1973.

THE NATURE OF COMETS

COMETS AND METEOR STREAMS. J. G. Porter. John Wiley & Sons, Inc., 1952.

SURVEY OF THE UNIVERSE. Donald H. Menzel, Fred L. Whipple and Gerard de Vaucouleurs. Prentice-Hall, Inc., 1971.

THE COOPERATIVE ACTION OF MUSCLE PROTEINS

MUSCLES, MOLECULES AND MOVEMENT. J. R. Bendall. American Elsevier Publishing Co., 1969.

CONTRACTILITY OF MUSCLE CELLS AND RELATED PROCESSES: A SYMPOSIUM HELD UNDER THE AUSPICES OF THE SOCIETY OF GENERAL PHYSIOLOGISTS. Edited by R. J. Podolsky. Prentice-Hall, Inc., 1971.

PROPOSED MECHANISM OF FORCE GENERATION IN STRIATED MUSCLE. A. F. Huxley and R. M. Simmons in *Nature,* Vol. 233, No. 5321, pages 533–538; October 22, 1971.

MUSCLE CONTRACTION AND CELL MOTILITY. H. E. Huxley in *Nature,* Vol. 243, No. 5408, pages 445–449; June 22, 1973.

MOLECULAR CONTROL MECHANISMS IN MUSCLE CONTRACTION. Annemarie Weber and John M. Murray in *Physiological Reviews,* Vol. 53, No. 3, pages 612–673; July, 1973.

THE BATAVIA ACCELERATOR

ACCELERATORS: MACHINES OF NUCLEAR PHYSICS. Raphael Littauer and Robert R. Wilson. Doubleday & Company, Inc., 1960.

PARTICLE ACCELERATORS: A BRIEF HISTORY. M. Stanley Livingston. Harvard University Press, 1969.

THE BIG MACHINE. Robert Jungk. Charles Scribner's Sons, 1970.

INTRODUCTION TO HIGH ENERGY PHYSICS. Donald H. Perkins. Addison-Wesley Publishing Co., 1972.

FAR EAST & AUSTRALIA

10 Express flights a week by SAS, Thai International and Aeroflot.

connections from all over Europe to Copenhagen/ Frankfurt	Monday SAS	Tuesday SAS	Wednesday THAI	Thursday Friday Saturday SAS	Saturday THAI	Sunday Aeroflot	TRANS-SIBERIAN EXPRESS Thursday Aeroflot	Saturday SAS
departures:								
LONDON	11.20	15.10	10.05	8.15	10.50	11.20	8.15	8.15
COPENHAGEN	15.00	16.25	12.05	12.05		15.00	12.05	12.00
FRANKFURT	12.50	18.30	14.05	9.30		12.50	9.30	9.30
	Tashkent	NONSTOP	NONSTOP	Tashkent	NONSTOP	Moscow	Moscow	Moscow
arrivals, following day:								
DELHI		7.00				5.05		
BANGKOK	9.20	12.50	6.55	6.25	6.00	11.30		
HONG KONG[1]	14.10	17.25	11.25	11.25	11.25	19.40		
KUALA LUMPUR		15.50						
SINGAPORE	12.55	17.10	10.35[3]	13.15/11.05/10.35	10.35[3]	15.00		
JAKARTA	14.25							
SYDNEY			22.20[3]	22.20[2]	22.20[3]			
TOKYO							9.35	9.10

[1]) Connecting flights from Bangkok. [2]) Thai International from Bangkok, Sunday
[3]) Connecting flight from Bangkok, through-going aircraft Europe–Sydney.

FIRST AND ONLY NONSTOP SERVICES
FRANKFURT—DELHI (SAS)
FRANKFURT–BANGKOK (THAI)
LONDON–BANGKOK (THAI)

ONLY SERVICES WITH NEXT-DAY ARRIVAL IN SYDNEY– JUST ONE NIGHT EN ROUTE!

SAS SCANDINAVIAN AIRLINES
Representing THAI International

NUTRITION AND THE BRAIN

BRAIN SEROTONIN CONTENT: INCREASE FOLLOWING INGESTION OF CARBOHYDRATE DIET. John D. Fernstrom and Richard J. Wurtman in *Science*, Vol. 174, No. 4013, pages 1023–1025; December 3, 1971.

L-TRYPTOPHAN, L-TYROSINE, AND THE CONTROL OF BRAIN MONOAMINE BIOSYNTHESIS. R. J. Wurtman and J. D. Fernstrom in *Perspectives in Neuropharmacology*, edited by Solomon H. Snyder. Oxford University Press, 1972.

BRAIN SEROTONIN CONTENT: PHYSIOLOGICAL REGULATION BY PLASMA NEUTRAL AMINO ACIDS. John D. Fernstrom and Richard J. Wurtman in *Science*, Vol. 178, No. 4059, pages 414–416; October 27, 1972.

CONTROL OF BRAIN 5-HT CONTENT BY DIETARY CARBOHYDRATES. J. D. Fernstrom and R. J. Wurtman in *Serotonin and Behavior*, edited by J. Barchas and E. Usdin. Academic Press, 1973.

EFFECT OF PERINATAL UNDERNUTRITION ON THE METABOLISM OF CATECHOLAMINES IN THE RAT BRAIN. W. J. Shoemaker and R. J. Wurtman in *The Journal of Nutrition*, Vol. 103, No. 11, pages 1537–1547; November, 1973.

THE WIND BRACING OF BUILDINGS

STRUCTURE IN ARCHITECTURE. Mario G. Salvadori and Robert Heller. Prentice-Hall, Inc., 1963.

STRUCTURE: AN ARCHITECT'S APPROACH. H. S. Howard, Jr. McGraw-Hill Book Company, 1966.

BUILDING AND CONSTRUCTION. Carl W. Condit in *Technology and Western Civilization: Vol. I*. Edited by Melvin Kranzberg and Carroll W. Pursell, Jr. Oxford University Press, 1967.

THE JOHN HANCOCK CENTER. Fazlur R. Khan in *Civil Engineering*, Vol. 37, No. 10, pages 38–42; October, 1967.

PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON PLANNING AND DESIGN OF TALL BUILDINGS. American Society of Civil Engineering, 1972.

THE AMATEUR SCIENTIST

FIELD BOOK OF WILD BIRDS AND THEIR MUSIC. F. S. Mathews. G. P. Putnam's Sons, 1921.

"REMINISCENCE" IN THE COLD FLOUR BEETLE (TENEBRIO MOLITOR). Thomas M. Alloway and Aryeh Routtenberg in *Science*, Vol. 158, No. 3804, pages 1066–1067; November 24, 1967.


SMIRNOFF
SILVER®
1877
1886
PURVEYORS TO
THE IMPERIAL R
Ste Pierre Smirnoff
90.4 PROOF

Our basic wagon idea: little cash and lots of carry.

Like the original station wagon Ford pioneered back in 1929, the 1974 Ford Pinto Wagon combines a durable, economical car with a lot of space in back. And what could be more basic than that?

The economy: First and foremost, the Pinto Wagon is an economy car—low in price, easy on gas, low on upkeep.

The space: Put a Pinto Wagon's rear seat down, and you've got over 60 cubic feet of cargo space. To maximize cargo space, the spare tire is carried under the floor.

Other basics: A rugged 2,000cc overhead cam engine is standard. It's been developed for good gas mileage. And for those of you who want even a bit more pep, there's an optional 2,300cc engine.

Front disc brakes are also standard—for efficient and fade-resistant braking and little pedal effort.

Further, the steering is rack-and-pinion, like in sports cars. Who'd have thought a station wagon could handle that easy? They're all good reasons why the closer you look, the better we look.

See the 1974 Pinto Wagon at your Ford Dealer.

When you get back to basics, you get back to Ford.

FORD PINTO

FORD DIVISION Ford

(Shown here: Pinto Wagon with Squire Option, optional whitewall tires, luggage rack, and deluxe bumper group.)