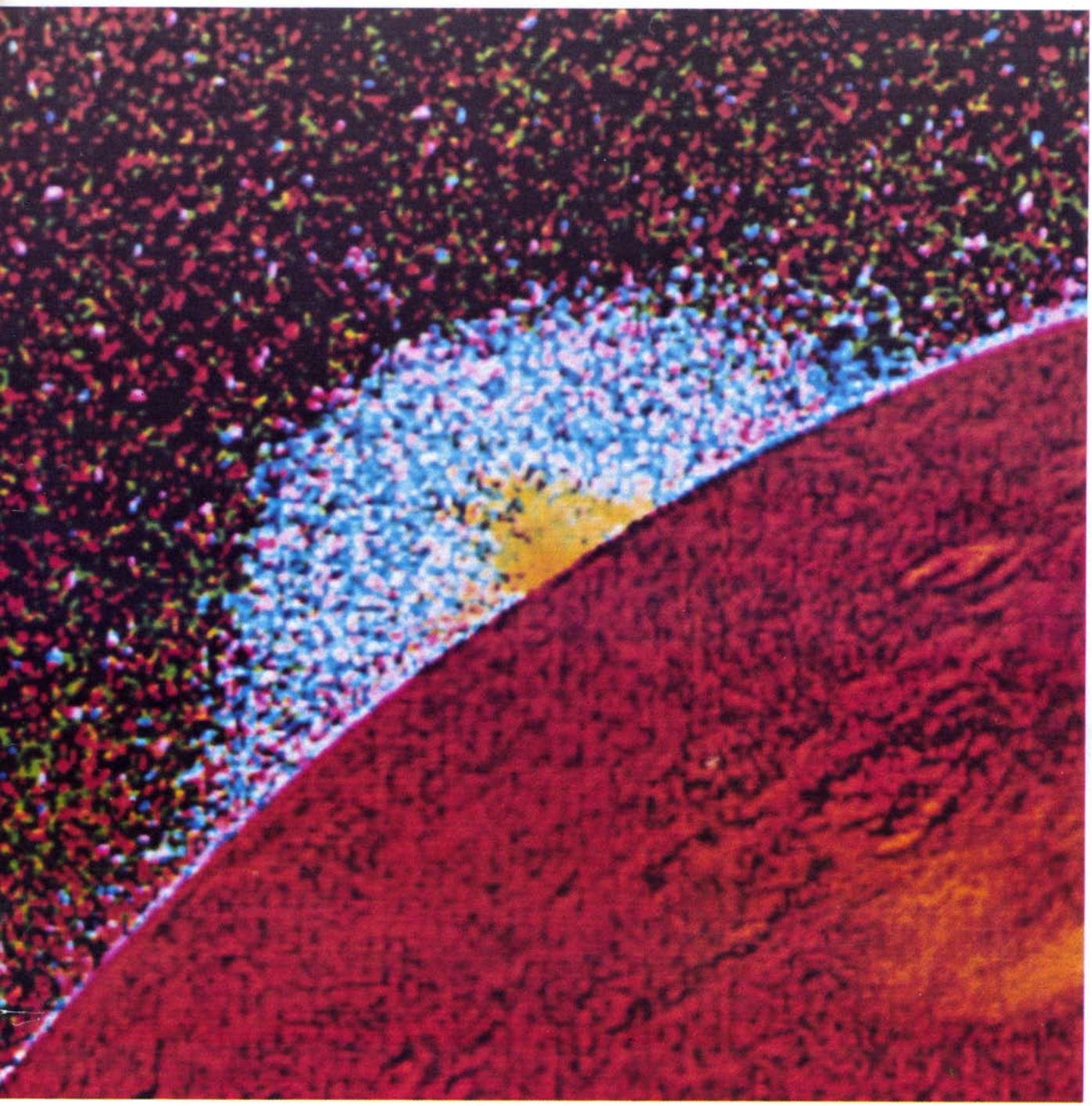
SCIENTIFIC AMERICAN



VOLCANOES OF IO

$2.00

0 01 73732 14109

January 1980

BMW. THE CAR COMPANY THAT HASN'T BEEN LEGISLATED INTO MEDIOCRITY.

The challenge that faces the automaker of the 1980's lies not in meeting tougher pollution standards, nor in coping with ever more restrictive safety regulations—not even in accommodating the petroleum shortages that are fast becoming a regular summer occurrence.

No indeed. The challenge lies in meeting the myriad of societal demands and still building automobiles that are worth driving.

Pessimists darkly predict that it cannot be done and point to the depressing number of uninspired automobiles on the market as incontrovertible proof.

Yet, fortunately, these gloomy predictions do not take into account the inventive genius and the obsessive determination of the engineers at BMW in Munich, Germany to build extraordinary automobiles.

Against all odds the BMW not only meets the demands of

society, it also provides the sort of exhilarating driving experience that automotive enthusiasts have all but given up for lost.

A SEEMINGLY INCONGRUOUS COMBINATION OF LUXURY, PERFORMANCE AND EFFICIENCY.

Conventional automotive wisdom has it that, inevitably, one must choose between luxury, performance and fuel efficiency. A sufficiency in any one necessitating a corresponding insufficiency in another.

And in the face of this, the BMW seems all the more remarkable for requiring no compromise whatsoever.

Round a particularly sticky curve and the legendary BMW suspension system—independent on all four wheels—flexes with an uncanny resiliency that makes one feel as if the car were built into the roadway. A sensation that results from highly advanced suspension technology and the virtually perfect integration of man and machine.

The BMW power plant—whether 4 or 6 cylinder—has long been a model for modern design. Press the accelerator and the fuel-injected, overhead cam engine responds in a manner that can only be described as exhilarating.

Yet, the 320i (with 5-speed standard transmission) delivers an impressive 25 EPA estimated mpg, 36 estimated highway mileage and, based on these figures, an estimated mpg range of 383 miles and a highway range of 551 miles.

(Naturally our fuel efficiency figures are for comparison purposes only. Your actual mileage and range may vary, depending on speed, weather and trip length. Your actual highway mileage and highway range will most likely be lower.)

And, while the BMW provides as long a list of luxury items as one could sanely require, its luxury is purposefully engineered to help prevent driver fatigue.

Vital controls are within easy reach; the tachometer, speedometer and ancillary instruments are well-marked and easy to read.

Its front seats are designed to hold their occupants firmly in place, and are so thoroughly adjustable that it is all but a mathematical impossibility not to find a comfortable seating position.

SAFETY BEYOND STURDY BUMPERS.

At BMW, automotive safety was of primary concern long before the first bumper law was written.

The passenger compartment is encased in a solid steel "cage." So strong is this construction that the car could be dropped on its nose from the fourth floor of a building without significant damage to the compartment itself.

Yet with all its sheer strength—its computer-programmed "crush zones," its double-braking systems—the BMW features a safety device even more estimable.

Its exceptional responsiveness and handling characteristics give it the ability to avoid accidents as well as merely survive them.

All in all, in a time of lowered automotive expectations, amidst increasing mediocrity, the engineers at BMW have actually improved the BMW.

If you'd care to judge for yourself, telephone your nearest BMW dealer to arrange a thorough test drive at your convenience.

THE ULTIMATE DRIVING MACHINE.

BMW, MUNICH, GERMANY

Hardly your everyday liqueur.

Some things in life are too precious for every day. Lochan Ora is one. An imported golden liqueur with a unique taste all its own, it's to be lingered over—sipped slowly, savored fully. You'll be proud to share it—but not with just anybody.

Lochan Ora. Imported Liqueur.

70 PROOF • IMPORTED BY GENERAL WINE & SPIRITS CO., N.Y., N.Y.

Established 1845 **SCIENTIFIC AMERICAN** January 1980 Volume 242 Number 1

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y 10017 CONTROLLED CIRCULATION POSTAGE PAID AT CHICAGO, ILL., AND AT NEW YORK, N.Y AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA, $22 PER YEAR. ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y 10017

executive health

the report that briefs you on what to watch

Volume XVI, Number 2 • November, 1979 • Pickfair Bldg., Rancho Santa Fe, Calif. 92067 • Area 714:756-2600

Hugh C. Trowell, M.D., F.R.C.P.:

A NEW DIETARY EXPLANATION for the cause of ESSENTIAL hypertension . . . the dramatic and significant African experience (1929-1958)

PUBLISHER'S NOTE: Dr. Hugh C. Trowell taught medicine in Kenya and Uganda medical schools for nearly 30 years, 1929 to 1958. During this critical period Kenya offered unique opportunities to make medical observations among the Africans, today called Kenyans. Never before, and never again, in the long history of medicine will highly trained doctors be able to observe three million men, women and children emerge from tribal life and undergo rapid Westernization of their diet and life-style. As a result Dr. Trowell offers here a new dietary explanation for the cause of essential hypertension and how to avoid or correct this high blood pressure by simple changes in what you eat.

—Richard Stanton

Standard teaching on high blood pressure is that a few persons develop high blood pressure due to kidney disease or some other disorder; this is secondary hypertension and treatment consists in treating the primary disorder. It is usually considered that high blood pressure has no obvious cause; it is then called essential or primary hypertension, and it is treated by reducing any obesity and by suitable drugs.

This report challenges such a point of view. It suggests that by reversing the high sodium to low potassium ratio of our Western diet, and of dietary fiber, over half of the people suffering from essential hypertension will achieve normal blood pressure within a few weeks.

First, let us survey the recent U.S. teaching concerning the cause of essential hypertension. The Annual Discourse to the Massachusetts Medical Society on May 24, 1978, was given by Dr. T. R. Dawber and entitled 'Unproved hypotheses.' He stated that "the hypothesis linking blood pressure to sodium intake, although sup-

How Long Will You Live?

Ignorance about how to take care of yourself can be the mistake that kills you in your 50's or 60's . . . "the dangerous years"!

Discover what research scientists now know about how to take the best care you can of the only body you will ever have.

Do as so many thousands of executives do. Subscribe to Executive Health Report. The members of our ***Editorial Board*** are among the world's most distinguished authorities on *preventive medicine*. Their wise advice can help you not only live longer but enjoy those extra years! (Note their high qualifications.)

Executive Health Report is not sold on newsstands but only by private subscription of $24 a year in the U.S.A. and its possessions, Canada and Mexico. Individual reports (back issues) $2.00 per copy. All other countries $26 by surface mail, $30 by air mail. (Only International Money Order or check cashable on U.S. bank will be acceptable.)

Subscribe now under this unusual introductory offer:

(1) Your choice of any three of the reports listed below ($2.00 each) FREE!

(2) If you are not satisfied with your first issue, your money will be *promptly* refunded.

Please study the reports listed here and circle your three choices:

New research findings On Walking . . . nature's own amazing "anti-age antibiotic"! Not running, not jogging, but walking is your most efficient form of exercise and the only one you can safely follow all the years of your life!

Linus Pauling, Ph.D.: For the best of health, how much vitamin C do you need? People who take the optimum amount of vitamin C may well have, at each age, only one quarter as much illness and chance of dying as those who do not take extra vitamin C.

New research findings On Zinc . . . the amazing metal so essential to your health.

On The Syndrome Of Longevity. Key factors dependably present ("constants") in the way of life of men who have outlived other men born when they were by a generation!

Roger J. Williams, Ph.D., D.Sc.: On Your Startling Biochemical Individuality. Some amazing facts about your body you need to know if you want to understand yourself (and other people) better.

Leonard Hayflick, Ph.D.: A New Theory On The Cause Of Heart Disease And Strokes . . . Vitamin B_6 may be the key. What new research has revealed that may downgrade cholesterol as the main predisposing factor . . . and why.

John K. Lattimer, M.D., Sc.D.: On that treacherous gland, your prostate — especially as you approach 50.

John K. Lattimer, M.D., Sc.D., F.A.C.S.: On The Improved New Blood Test For Cancer Of The Prostate — a giant step forward in the fight to outwit man's dread enemy.

Linus Pauling, Ph.D.: What About Vitamin E? Eminent investigators now suspect it may be one of the key factors to help resist disease and slow the aging process.

Hugh C. Trowell, M.D., F.R.C.P.: A New Dietary Explanation for the cause of Essential Hypertension . . . the dramatic and significant African experience (1929-1958).

Go easy gentlemen, too much social drinking damages your liver (even though you may feel well, eat well, and never get tight!).

On stress and the executive. Dr. Hans Selye, world-famous endocrinologist, explains why hard work is the secret of living a long, satisfying and successful life.

The truth about the right diet and exercise to lose weight safely and stay slender for life.

Alton Ochsner, M.D.: On "The Chair Disease" . . . Why blood clots in your veins are a little-realized hazard of "desk-bound" executives.

Stephen R. Elek, M.D. On "The Hurry-up Disease." Why it may be a key factor that triggers heart attacks as early as age 40.

If you sit at a desk all day . . . How to avoid a Pot-Belly and Double-Chin (or get rid of them for good!)

Jan Koch-Weser, M.D.: On Systolic Hypertension, more common than diastolic hypertension and as dangerous. What it is, what it does, and what to do about it.

Linus Pauling, Ph.D.: On Vitamin C and Cancer. Recent studies show that vitamin C has a large life-extending effect for patients with advanced cancer and suggest a similar large effect for earlier stages of the disease.

Leonard Hayflick, Ph.D.: On Getting A Good Night's Sleep! What research scientists now know about this mysterious process we so calmly call "sleep."

Albert Szent-Gyorgyi, M.D., Ph.D.: On A Substance That Can Make Us Sick (If We Do Not Eat It!) One of the world's most honored scientists, winner of the 1937 Nobel Prize for Physiology and Medicine explains the fascinating paradox of vitamins.

Denis P. Burkitt, M.D., D.Sc., F.R.C.S., F.R.S.: Is Dietary Fibre Protective Against Disease? The drop in our consumption of potatoes and whole grains may have something to do with the increase in intestinal cancer, diabetes and heart disease—to name a few!

Please use the coupon below under our special money-back guarantee.

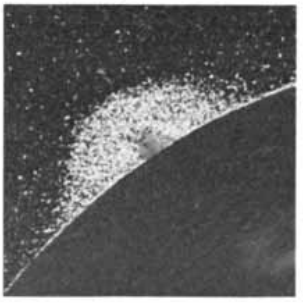
EXECUTIVE HEALTH, Pickfair Bldg., Rancho Santa Fe, CA 92067

Gentlemen: Enclosed is my check for $_____ for a year's subscription to Executive Health to start with this month's issue. I have circled the three $2.00 reports I am to receive free. It is understood that if I am not satisfied with my first issue, my money will be refunded. In addition, send me a list of some 100 other reports of yours still in print. There may be some from which I might greatly benefit and would otherwise miss.

Firm/Name (Please print clearly) _____

Street & No./Route _____ **Suite or Apt.** _____

City, State, Country _____ **Zip** _____ SA-4



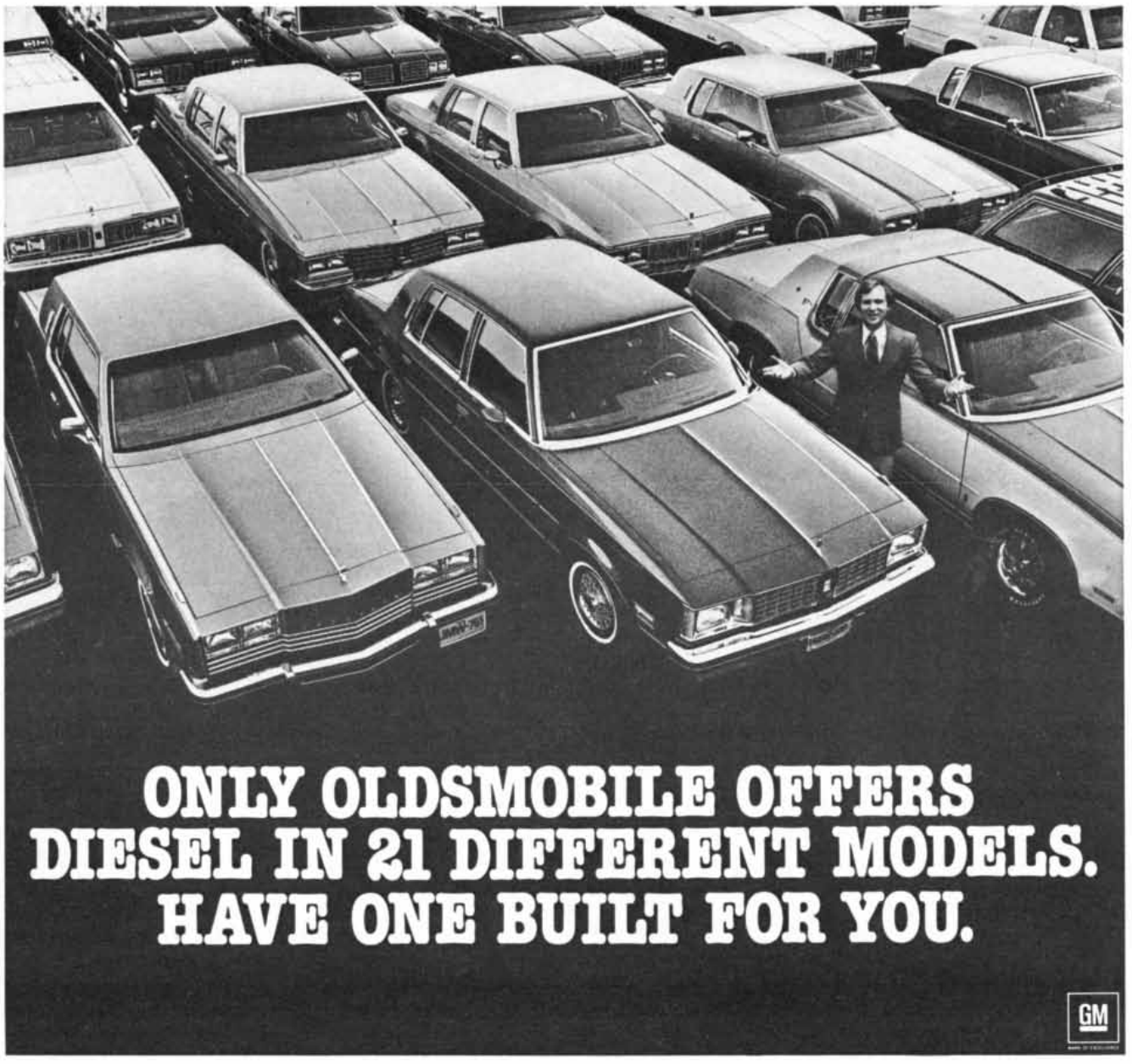
THE COVER

The picture on the cover shows Plume 2, the second of eight volcanic eruptions detected on the Jovian moon Io by the spacecraft *Voyager 1* in the course of its historic encounter with Jupiter on March 5, 1979. The symmetrical blue region, which rises some 200 kilometers above the surface of the satellite, is the outer envelope of the plume recorded through an ultraviolet filter. The image, enhanced by computer processing at the Jet Propulsion Laboratory of the California Institute of Technology, is a "ratio composite" of four images recorded sequentially through separate filters. The ultraviolet, violet and orange images have been divided by the green image and used as blue, green and red in the composite. The plume as it looks in visible light alone appears at the bottom left on page 94 in "The Galilean Moons of Jupiter," by Laurence A. Soderblom.

THE ILLUSTRATIONS

Cover photograph courtesy of Jet Propulsion Laboratory, California Institute of Technology

Page	Source	Page	Source
22–32	Ilil Arbel	92–99	Jet Propulsion Laboratory, California Institute of Technology
43	Fermi National Accelerator Laboratory	100	Ilil Arbel
44–57	George V. Kelvin	103–111	Jerome Kuhl
59	Leonard Hayflick, Children's Hospital Medical Center of Northern California	113–118	Tom Prentiss
60–65	Alan D. Iselin	121	Carey Calloway, Center for Disease Control, U.S. Public Health Service
67	EROS Data Center, National Aeronautics and Space Administration	122–123	Patricia J. Wynne
68	Stephen Deffeyes	126	Martin M. Kaplan and Hilary Koprowski, Wistar Institute of Anatomy and Biology
69–76	Andrew Tomko	128–134	Patricia J. Wynne
89	Jet Propulsion Laboratory, California Institute of Technology	139–148	Dan Todd
90–91	Ilil Arbel	155–159	Michael Goodman



Introducing the 1980 Olds diesels. The best-selling diesel cars in America.

The Olds diesels are a success. A runaway success. In fact, last year alone we sold over 100,000 of them.

Well, with diesel fuel *traditionally costing less* than gas, plus diesel availability in 21 different kinds of popularly priced Oldsmobiles, *and* the kind of mileage and range you see here, is there any wonder?

Remember: The circled EPA estimates are for comparison to other cars. Your mileage and range depend on your speed, weather and trip length; your actual highway mileage and range will probably be less than the highway estimates. Driving range estimates are obtained by multiplying the EPA and highway estimates by the standard fuel tank capacity ratings. Diesels are not currently available in California. Oldsmobiles are equipped with GM-built engines produced by various divisions. See your dealer for details.

Models	Fuel Tank Cap.	EPA Est. MPG	EPA Est. Range	Hwy. Est.	Hwy. Range
Cutlass	19.8	(22)	(435)	34	673
Cutlass Cruiser	18.2	(22)	(400)	34	618
Delta 88	27.0	(22)	(594)	34	918
Toronado	23.0	(22)	(506)	34	782
Ninety-Eight	27.0	(21)	(567)	31	837
Custom Cruiser	22.0	(21)	(462)	31	682

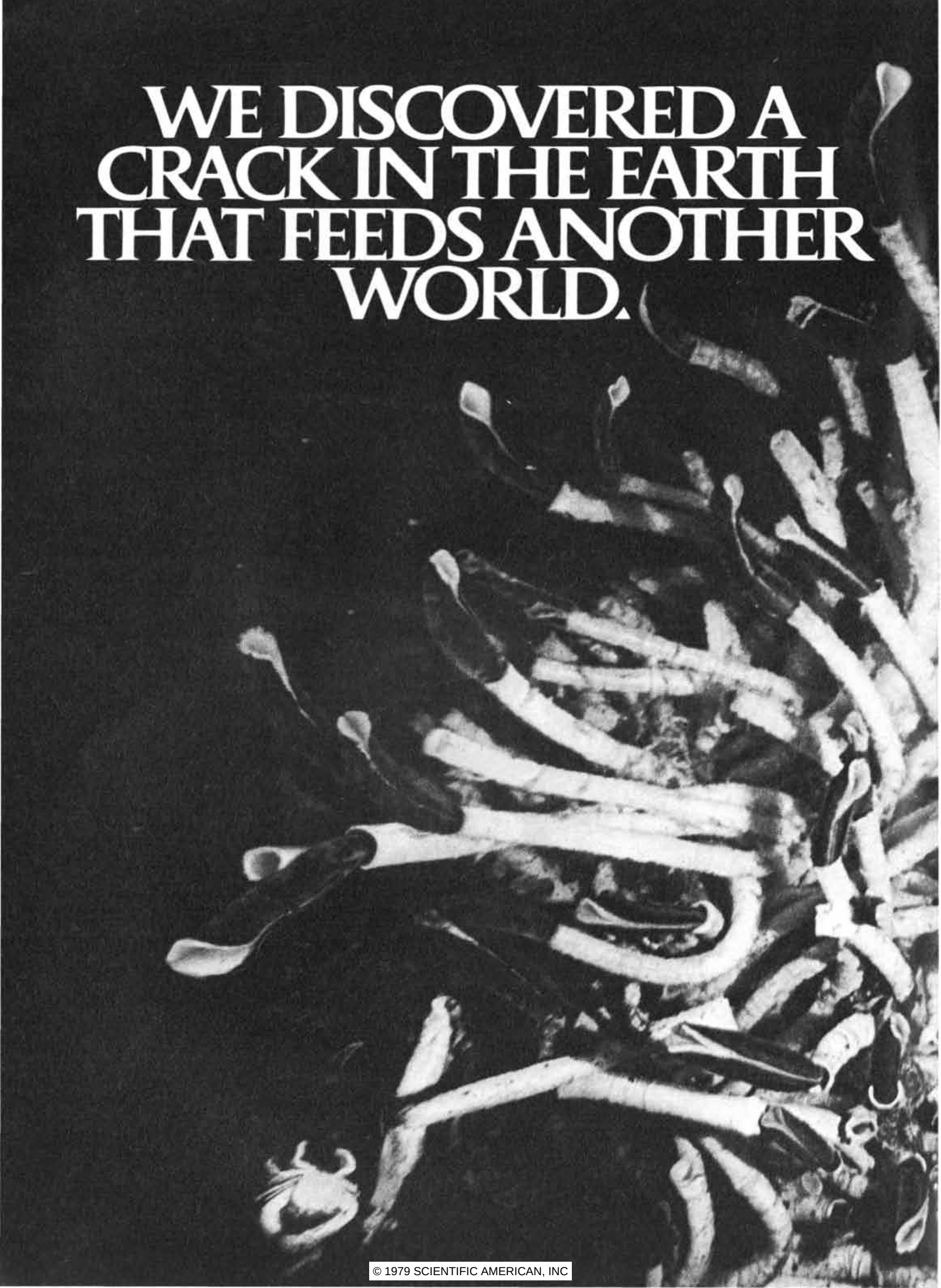
1980 Oldsmobile diesels. All the benefits of diesel. All the value of Oldsmobile.

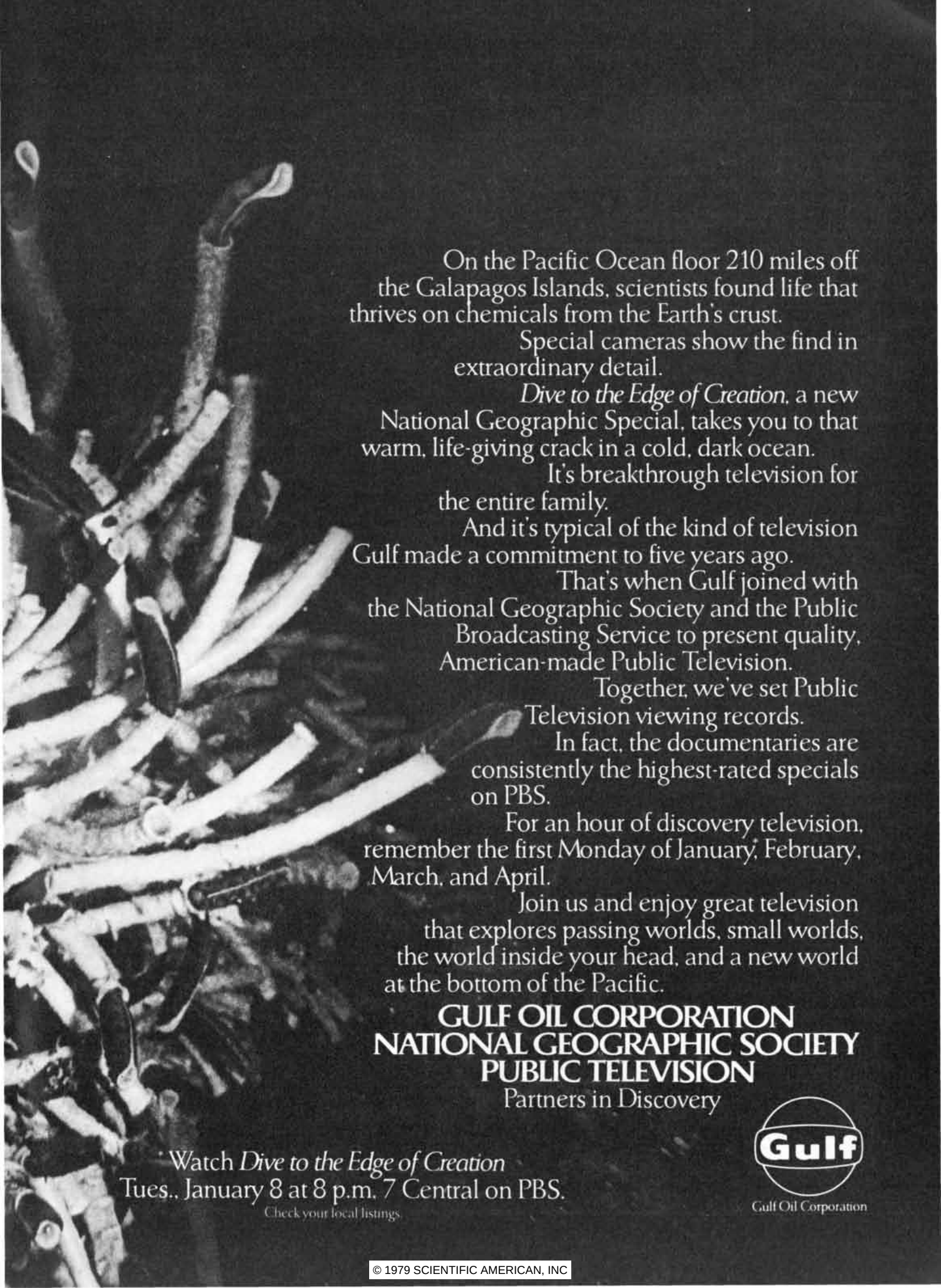
Buy or lease one of the new Olds diesels now.

Oldsmobile

HAVE ONE BUILT FOR YOU.

WE DISCOVERED A CRACK IN THE EARTH THAT FEEDS ANOTHER WORLD.





On the Pacific Ocean floor 210 miles off the Galapagos Islands, scientists found life that thrives on chemicals from the Earth's crust.

Special cameras show the find in extraordinary detail.

Dive to the Edge of Creation, a new National Geographic Special, takes you to that warm, life-giving crack in a cold, dark ocean.

It's breakthrough television for the entire family.

And it's typical of the kind of television Gulf made a commitment to five years ago.

That's when Gulf joined with the National Geographic Society and the Public Broadcasting Service to present quality, American-made Public Television.

Together, we've set Public Television viewing records.

In fact, the documentaries are consistently the highest-rated specials on PBS.

For an hour of discovery television, remember the first Monday of January, February, March, and April.

Join us and enjoy great television that explores passing worlds, small worlds, the world inside your head, and a new world at the bottom of the Pacific.

GULF OIL CORPORATION
NATIONAL GEOGRAPHIC SOCIETY
PUBLIC TELEVISION
Partners in Discovery

Watch *Dive to the Edge of Creation*
Tues., January 8 at 8 p.m. 7 Central on PBS.
Check your local listings

Gulf
Gulf Oil Corporation

LETTERS

Sirs:

"Acid Rain," by Gene E. Likens, Richard F. Wright, James N. Galloway and Thomas J. Butler [SCIENTIFIC AMERICAN, October], is by far the best article on the subject that has come to my attention. It nonetheless neglects to consider some factors that are of importance in studying the problem.

First there is a need to differentiate between sulfuric acid and nitric acid in rain. The sulfuric acid is largely manmade, although some of it undoubtedly results from volcanic action and the atmospheric oxidation of hydrogen sulfide from swamps. The principal sources for the eastern U.S. are the smelting of sulfide ores in Ontario, Missouri and Tennessee.

Nitric acid, however, is significantly natural in its occurrence. At Paris in 1852, 63.6 kilograms of acid, presumably mostly nitric acid, fell on each hectare, corresponding to a *p*H of 4.0. In rural atmospheres Bunyan in 1854 found acidities corresponding to *p*H 4.5, 4.8 and 5.3 respectively in summer, spring-fall and winter. This points up the significance of thunderstorms, which have long been recognized as a major source of soil nitrogen in a natural environment. The lightning flash is an almost ideal apparatus for achieving substantial yields in the endothermic reaction $N_2 + O_2 \rightarrow 2NO$ and then "freezing" the mixture once the flash is over. About 230 kilograms of nitric acid can be produced in a single flash. Two such flashes on a one-square-kilometer watershed of a mountain lake during precipitation of two centimeters of water will yield a *p*H of 3.42. In eutrophic and calcareous environments this nitric acid acidity is soon neutralized. But in the siliceous rock, thin acid soils and oligotrophic lakes of the Adirondacks and the White Mountains it persists, so that fish life there has been sparse at best. The larger lakes at the lower elevations are where severe effects should be studied, and then only when the fishing pressure is a constant.

Sulfuric acid is a different matter. The authors of the article rightly point out the significance of the Ontario sulfide-ore smelting at Sudbury, which has vastly increased as a result of the recent discoveries at Timmins. To these emissions must be added the American sources in Tennessee and Missouri. All have characteristics in common: sulfur dioxide emissions are in parts per thousand rather than parts per million, most have adopted high stacks as a means of reducing ground-level acidity and all lie to the west of the affected area.

The authors suggest that the combustion of coal, largely for power generation, is the major contributor of sulfuric acid. Some 600 million tons of coal is burned in the U.S. annually; it averages about 2 percent sulfur and will yield about 1.15 pounds of nitrogen oxides per million B.t.u. of heat released (from 87 pounds of coal). Potentially 37 million tons of sulfuric acid and 16 million tons of nitric acid can be produced annually; if these amounts were trapped in precipitation (1.05 meters annually) over the eastern half of the U.S., they could yield a *p*H of 3.5 or contribute 311 microequivalents of hydrogen ion per liter. Most of this coal, however, is burned east of the Appalachians, and westerly winds will carry much of the acid out to sea. Only in the New York and Pennsylvania–West Virginia areas does the burning of coal, on the basis of the authors' maps, appear to be a significant cause of acid rain. Moreover, the consumption of coal has been essentially constant over the past 20 years, so that the burning of coal is not responsible for the observed increase.

In the sensitive mountain areas that have rightly given rise to concern the significant causes of acid rain are in my opinion nitric acid from thunderstorms and emissions from smelters. I do not believe the data presented agree with the identification of emissions from power plants as the major source.

WINSLOW H. HARTFORD

Belmont Abbey College
Belmont, N.C.

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 C. John Kirby, 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 40 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name _____

New Address _____

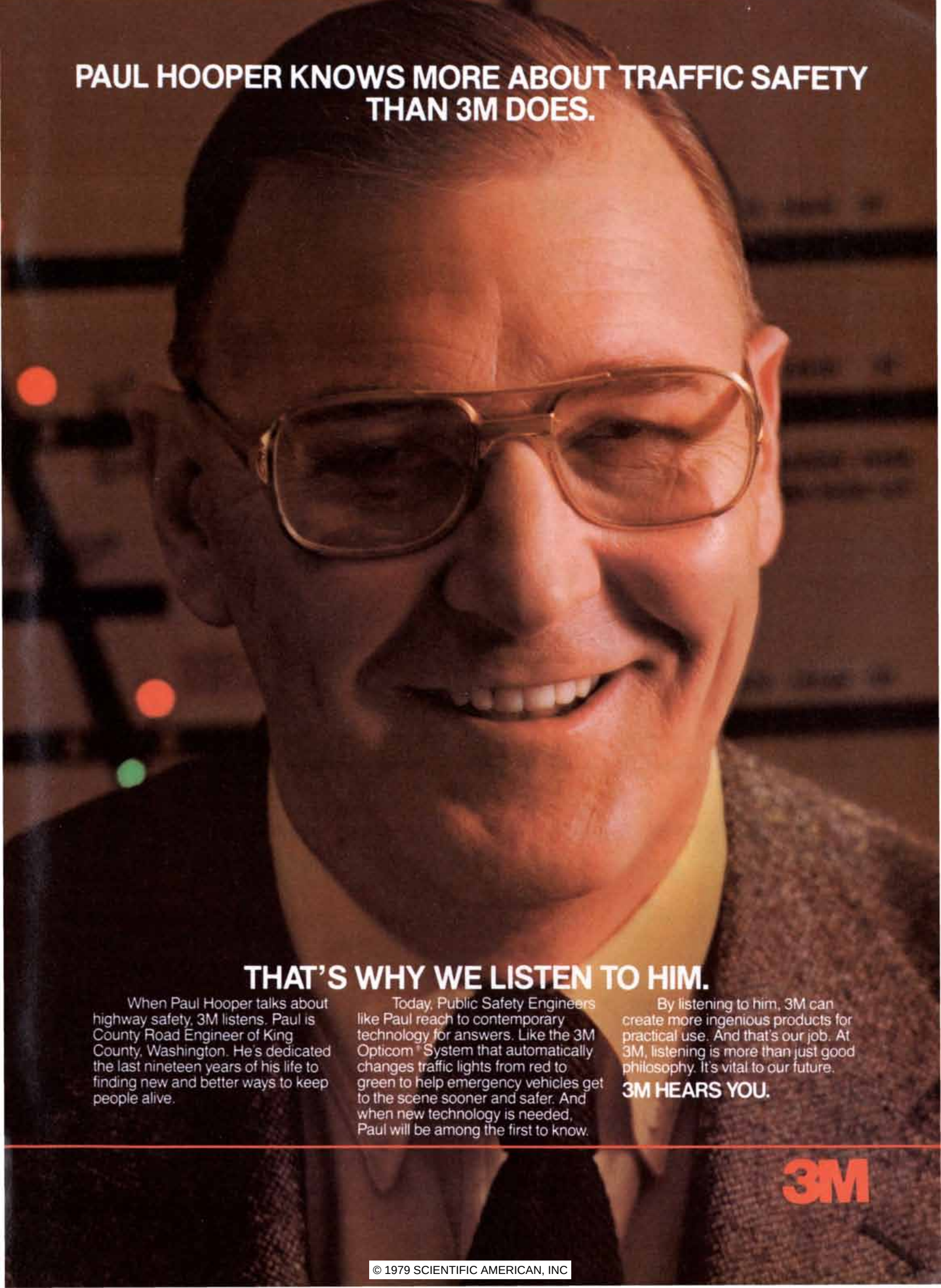
_____

_____

Old Address _____

_____

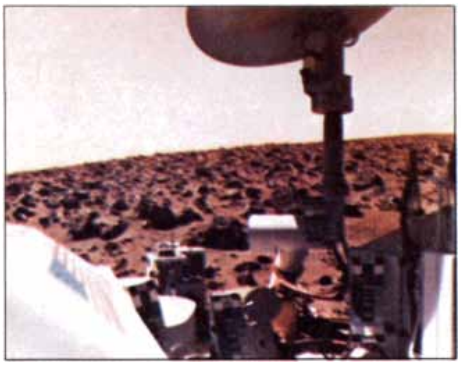
_____



Using space technology to save lives on Earth

Left on Mars by Viking I and II in 1976, Perkin-Elmer gas analyzers sampled the Martian soil for signs of life. In Skylab, when the astronauts exercised, the same basic instrument analyzed their respiratory condition. In every U.S. nuclear submarine, a similar Perkin-Elmer device stands watch over the purity of the air breathed by the crew.

In 1972, Perkin-Elmer put the technology developed for such space and defense applications to work in the health care field with the introduction of our first Medical Gas Analyzer. Since then, hundreds of these units have been used for pulmonary testing, physiological and anesthesia research, stress testing and other related applications.



The Medical Gas Analyzer, an exceptionally rugged and maintenance-free Perkin-Elmer mass spectrometer, is the heart of our new Respiratory Monitoring System (RMS III). By measuring and monitoring a patient's inspired and expired gases, the system provides the physician with information for improved respiratory and anesthetic management.

Up to 16 patients can be sequentially monitored with one RMS III. A unique rotary valve selects the patient line to be sampled and a small amount of gas is drawn into the mass spectrometer and analyzed. Computation and readout take place at a central monitor, located at the nurses' station, with remote data displays available for each operating room or bedside.

RMS III gives simultaneous inspired and expired values for up to six different gases, and compares them with pre-set levels entered via the system's keyboard. Audio and visual alarms are activated if any of these gases exceeds prescribed limits.

Each patient's breathing can be viewed as a moving wave form on the monitor screen. Eight- or twenty-four-hour trends of various physiological parameters as well as gas concentrations can also be displayed on the screen for each patient.

In many cases, developing problems can be anticipated before they would otherwise be apparent

Newest wave of powerful small computers for industry

From vans bouncing over desert tracks to the clanging production lines of heavy machinery factories, Perkin-Elmer 16-bit small computers have proven themselves a rugged, versatile breed.

Now we've brought out a new line of these popular computers, called the Series Sixteen. With Series Sixteen and new software—including FORTRAN—system builders can combine the most advanced computer control with economy and ruggedness to handle many jobs in manufacturing, automation, industrial design and data collection.

Drawing on 12 years' experience with industrial computers, Perkin-Elmer designed the new Series Sixteen for optimum price/performance, and packaged it for flexibility—it's a family of processors that answers a wide range of needs. Series Sixteen incorporates features that previously were extras, features like hardware multiply divide, power fail, single board memory expansion and expansion power

Designed to maximize printed circuit modularity, all units can be readily customized, or upgraded, or interfaced with larger computers.

That Series Sixteen's versatility and reliability provide an excellent solution for industrial applications is shown by some of the jobs earlier Perkin-Elmer small computers are doing all over the world.

They travel the wilderness in vans, collecting and taping seismic data from geological test blasting over possible oil formations.

They prepare and distribute the numerical control tapes which automate the fabrication of expensive, high-tolerance machine parts.

One of the big U.S. auto makers uses our small computers for carburetor testing and communicating results to a 32-bit computer, while a European car manufacturer uses them for control of robots in assembly-line automation.

And Perkin-Elmer 16-bit computers monitor the flow of petroleum through pipeline networks to and from tank farms on five continents.

With hardware proven for dependability and low cost, the new Series Sixteen stands out for state-of-the-art technology as well: high density LSI circuitry, fast Metal Oxide Semiconductor memory, dual bus architecture for high-speed and low-speed devices, no fewer than 255 vectored interrupts and battery backup for memory retention during power failures.

Software includes a comprehensive resource-sharing operating system, language processors for CAL Macro, BASIC and, especially, an enhanced scientific FORTRAN IV.

Sophisticated new system for chemical analysis

In determining the chemicals present in unknown substances, our new liquid chromatography Methods Processor routinely performs feats of quantitative and qualitative analysis once regarded as impractical, or even impossible. Use of the system—built around our powerful new LC-75 spectrophotometric detector with Autocontrol—is adding rapidly to basic chemical knowledge. It is also satisfying more practical needs in such applications as analysis of plastics and food, crime lab work, pharmaceutical development, dye and cosmetics manufacture, and monitoring environmental purity, to name just a few.

Liquid chromatography separates the molecular species in any substance which can be dissolved in an appropriate solvent. In essence, a sample is pumped under high pressure, at controlled temperatures, down a tube filled with small packed particles (such as silica gel). The molecules with most bonding affinity for the packing material are retained longest in the tube and emerge from it last. By flushing the packed particles with increasingly strong solvent and examining by spectrophotometry the effluents and the order in which they emerge, a chemist can identify ingredients and calculate how much there is of each in the sample

Obviously, the spectrophotometer is a key element in the process. And fittingly, the LC-75 with complete electronic digital control (Autocontrol) is the most revolutionary module in the system. It's the first spectrophotometric detector designed specifically for LC. It can shine through a dissolved sample an analytical wavelength precisely variable from 190 to 600 nanometers. Trace analysis is excellent—even with tiny samples. It takes the operator through even the most sophisticated chromatographic peak analysis step by step. And quantitative capability is superior—providing the widest range of usable sample concentrations plus reliably repeatable results. Autocontrol unlocks the tremendous spectroscopic power of the LC-75.

Combine the LC-75 with the four other modules in the Methods Processor system (Series 3B pump, Model 420 Automatic Sampler/Injector, LC-100 column air-bath oven, and SIGMA 10 Chromatography Data Station) and you get infinite options for performing exactly the kind of analysis you need.

For more information on these products, please write: Corporate Communications, Perkin-Elmer Corporation, Main Ave., Norwalk, CT 06856.

PERKIN-ELMER

Responsive Technology

"Last year I switched to rum. This year I graduated to Myers's Rum."

White rums may be what you learn on. But Myers's dark rum will advance your education. It will teach you just how good tasting rum can be. Because with Myers's Rum you get a smoother, softer taste that comes from master-blending and longer aging.

What makes Myers's precious imported rum cost more, makes Myers's taste better. In cola, soda, fruit juice or any of your favorite mixers.

MYERS'S MAKES IT BETTER. Taste how Myers's improves on cola, soda, tonic, fruit juice.
Free Recipe Book: Myers's Rum, P.O. Box 1622, FDR Station, New York, N.Y. 10022. Offer expires December 31, 1980.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JANUARY, 1930: "When one contrasts the high-speed, mass-production methods of modern industrialized agriculture with the more leisurely, more enjoyable but obviously less economical practices still in vogue in most of our agricultural regions, it is not hard to realize that 30 percent of our farmers produce 80 percent of our crops and livestock. During the past 20 years about five million people have left the farms, but the total agricultural output has increased; by 1950 there will probably be about 150 million people in the United States, and they will be fed by fewer farmers than were required in 1850 to feed only 25 million people. Three men operating a tractor combine can do the work of 18 threshers, and it now requires only 10 minutes of human labor to produce one bushel of wheat whereas it formerly required about three hours."

"A helicopter is to be built with four lifting air screws, each on a separate axis. With four air screws it is possible to have a large lifting area without excessive propeller dimensions. Another variation of the helicopter is the two-screw machine. There have also been helicopters with two screws rotating on the same axis, with one screw mounted above the other. Still another line of approach consists in using a single rotating air screw and in counteracting the torque of rotation by disposing an auxiliary screw at the rear of the machine. Besides mechanical troubles and dubious stability the helicopter offers tremendous problems in control under all circumstances. Innumerable patents have been taken out that indicate plausible helicopters on paper. The jump from a patentable idea to a practical helicopter is stupendous. The amount of money, ingenuity and trouble spent on the helicopter without tangible results is extraordinary. It is one of the mysteries of aeronautics as to whether its success will ever be attained."

"In spite of the fact that 'Eskimo Pie' had achieved almost phenomenal success commercially, having a marked effect upon the ice cream and confectionery industries, it has been judged unworthy of protection under its existing patent. The District Court for the Southern District of New York had ruled that the Burt ice cream lollipop (Patent No. 1470524) was valid and infringed. Various patents were cited in an attempt to show that the Burt patent was lacking in patentable novelty. One of these was a patent in which a stick was inserted into candy while the candy was in a warm, soft, sticky condition; upon cooling, the stick became rigidly fixed to the candy. The Court upheld the Burt patent, stating, 'It is one thing to insert a stick in a warm, soft, sticky substance that by cooling will adhere thereto, and quite another to effect a strong union between a stick and a frozen substance.'"

JANUARY, 1880: "To all appearances Mr. Edison has got the lamp he has so long been searching for, and curiously it is not at all what he thought it would be a few weeks ago. The light is generated in a strip of carbonized paper and not in a spiral of platinum or other refractory metal. The light produced is perfect; the lamp is inexpensive and apparently durable; the economy of the general system in which it is used is tolerably clear, and all its details seem to have been worked out with Mr. Edison's usual cleverness and practical skill. The only question that remains undetermined at this writing is whether the lamp will stand the test of time. It seems almost incredible that a slender thread of carbon can withstand the intense heat of the lamp, even in a perfect vacuum, without volatilization or fracture; but the lamps are stated to have stood action of the current, both in ordinary and in extraordinary strength, long enough already to upset all reasonable opinion as to the behavior of carbon under such conditions, and there is now nothing to be done but to wait for time to determine what the ultimate issue will be. The fact that all its predecessors in the field of incandescence have sooner or later come to grief is the chief if not the only one compelling a suspension of judgment in the present case."

"The most appalling of railway disasters occurred on the evening of December 28, at the bridge over the Firth of Tay, on the railroad between Edinburgh and Dundee in Scotland. At this point an iron bridge two miles long crosses the Firth on 85 spans, ranging from 18 to 88 feet above the water. On the evening of the disaster a train comprising locomotive and tender, four cars of the third class, one of the second class and one of the first class, together with a brakeman's van, entered upon the bridge near seven o'clock with a high wind blowing. In the bright moonlight the train was seen to reach the middle of the bridge over the navigable part of the Firth. Then suddenly with a flash of fire it disappeared. Subsequent examination found that a section of the bridge half a mile in length, comprising a dozen or more of the longer and higher spans, had fallen, and the train had been precipitated into the gulf. The first report of the managers of the railway said that there were nearly 300 passengers on the train besides the train-men. Not one survived. Later the authorities estimated the loss as low as 75. The exact number will probably never be known. It is impossible at this writing to obtain any clew to the cause of the disaster. The gale is said to have been the severest experienced in Scotland since 1868. It is most probable that the bridge was blown down."

"The transmission of power by electricity both for short and long distances is not only practicable but also economical; and the sanitary and other advantages of drawing power from a distance, for small manufacturing and for operating domestic machinery, are so enormous that the new system is sure to work great changes in all branches of industrial affairs. It is no stretch of the imagination to say that our children, if not ourselves, will see the small steam engine everywhere displaced by the electric motor, which will convert into motive power the subtle energy conveyed by wires from central sources of energy—huge furnaces constructed on the most approved scientific principles, out-of-the-way waterfalls, tidal currents, even the sun himself. And doubtless this cleanly and trusty servant will serve humanity in ways we are not able to dream of now, and at a cost that will be, by comparison with the present cost of light and heat and working energy, almost nominal."

"The news of Professor James Clerk Maxwell's death came with a sudden shock to many who are familiar with his name, for he was not an old man—he was in the prime of life, and great things were yet expected of him. Clerk Maxwell was at his strongest as a mathematician, or rather as a mathematical physicist. He belonged to the school of Sir William Thomson, together with the late Macquorn Rankine, Dr. Joule, Professor Tait and others. The principal labor of these physicists is to apply mathematical reasoning to experimental data in order to arrive at general laws, and their favorite study is molecular physics. Maxwell was one of the most promising and honored of these molecularists. His device of the 'sorting demon' will be remembered as long as the kinetic theory of gases is studied. By his colleagues he was deemed a star of the first magnitude, yet to shine forth in full splendor; but alas! for these hopes the star of the future has untimely faded into star mist. Much was expected from Clerk Maxwell that can never now be realized, and his premature death is a severe loss to the molecular science of the world."

THE AUTHORS

ROBERT R. WILSON ("The Next Generation of Particle Accelerators") has been a leading figure in the design and construction of machines for generating high-energy particle beams for more than three decades. Born in Frontier, Wyo., he went to the University of California at Berkeley, where he obtained his Ph.D. in physics in 1940. He then joined the faculty of Princeton University, where he collaborated with Enrico Fermi in making some early measurements of the neutron-absorbing properties of uranium 235; during that period he also invented a method of separating isotopes of uranium. In 1943 he moved to the Los Alamos Scientific Laboratory, where he was made director of the cyclotron group and later of the experimental-research division. He was director of the Laboratory of Nuclear Studies at Cornell University from 1947 to 1967, when he was appointed director of the Fermi National Accelerator Laboratory, a post he held until 1978. Wilson continues his involvement in accelerator planning as a member of the International Committee for Future Accelerators and as professor of physics at the University of Chicago's Enrico Fermi Institute of Nuclear Studies.

LEONARD HAYFLICK ("The Cell Biology of Human Aging") is a research biologist at the Children's Hospital Medical Center of Northern California. He received his Ph.D. in medical microbiology in 1956 from the University of Pennsylvania, where he worked for several years before joining the Wistar Institute of Anatomy and Biology. From 1968 to 1976 he was professor of medical microbiology at Stanford University. In the early 1960's, while at the Wistar Institute, he isolated and identified a new microorganism (*Mycoplasma pneumoniae*) as the cause of primary atypical pneumonia in man. Hayflick's cell-culture studies led him to the observation that normal human cells have a limited capacity for dividing, and his interpretation of this phenomenon as aging at the cellular level established the field of cytogerontology.

KENNETH S. DEFFEYES and IAN D. MACGREGOR ("World Uranium Resources") have collaborated on the subject of their article for a number of years. Deffeyes is professor of geology at Princeton University. A graduate of the Colorado School of Mines, he did his postgraduate work at Princeton, obtaining his Ph.D. in 1959. He then worked as a petroleum geologist for the Shell Development Corporation and taught at the University of Minnesota and Oregon State University before returning to Princeton as a member of the faculty in 1967. Over the years Deffeyes' work has ranged from sedimentary geochemistry and hydrometallurgy to chemical oceanography and the theory of plate tectonics. MacGregor is professor of geology at the University of California at Davis; he is currently away from Davis, working on a two-year appointment at the Department of Energy in Washington. A Canadian citizen, he was born in Calcutta, raised in South Africa, did his undergraduate work at the University of Aberdeen and received his M.Sc. from Queens University in Ontario before going to Princeton, where he took his Ph.D. in 1964. MacGregor was associate professor of high-pressure experimental petrology at the Southwest Center for Advanced Studies in Texas from 1965 to 1969, when he was appointed to his present post at Davis. Deffeyes and MacGregor want to acknowledge the assistance of James Kukula in writing the computer programs for the work reported in their article.

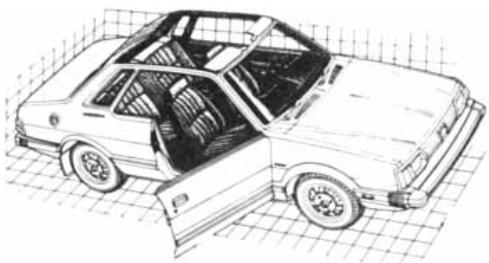
LAURENCE A. SODERBLOM ("The Galilean Moons of Jupiter") is chief of the branch of astrogeologic studies of the U.S. Geological Survey. He has undergraduate degrees in both geology and physics from the New Mexico Institute of Mining and Technology and a doctorate in planetary science and geophysics from the California Institute of Technology. Since joining the Geological Survey in 1970 he has participated in several of the National Aeronautics and Space Administration's unmanned planetary-exploration missions. Soderblom is currently deputy team leader for the Voyager Imaging Science Experiment.

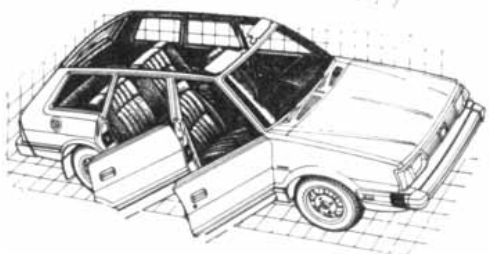
BARBARA GILLAM ("Geometrical Illusions") teaches and does research in visual perception and binocular vision at the College of Optometry of the State University of New York. Born in Australia, she studied psychology as an undergraduate at the University of Sydney and went on to obtain a Ph.D. for her work on binocular vision from the Australian National University. After a few years on the faculty of the University of Reading in England, she immigrated to the U.S. in 1967. Gillam writes: "I first became interested in illusions because I do not have a car and I frequently travel on the New York subway system and on buses. I find that doodling illusion figures with paper and pencil is much less attention-getting than trying out ideas in binocular vision, which involves frequently opening and closing one eye."

GEORGE W. BEADLE ("The Ancestry of Corn") got the time to resume

THE 1980 SUBARU. IT'S EVEN BEAUTIFUL IN PLACES YOU CAN'T SEE.

Your dollar may buy a lot less of everything else. But today, it buys a lot more Subaru. For 1980, we've made our cars larger, roomier, more comfortable and more elegant. While engineering them to pinch pennies.





A CAR THAT'S BEAUTIFUL CAN ALSO BE INTELLIGENT.

On the 1980 Subaru, the down-sloping hood and the air dam under its "chin" are designed for low wind resistance and high gas mileage. So our DL Hardtop 5-speed really delivers.

32 EST. MPG (CITY)* 46 EST. HWY.*

OUR ENGINE IS DESIGNED TO HELP OUR CAR HANDLE BETTER.

The horizontally opposed, aluminum Subaru engine lies low and flat with its weight positioned evenly over the front wheels. Coupled with Subaru front-wheel drive, this results in superior road holding on hills and curves. Even in snowy, slippery weather.

YOU CAN'T SEE SOME OF OUR BEST SELLING POINTS.

Many Subaru advantages are out of sight. We give you *two* separate brake line circuits. If one fails, Subaru still gives you a brake.

For 1980, Subaru introduces "zero scrub" suspension, which makes the car easier to turn and reduces steering wheel kick back, while increasing tire life. Rubber mountings create full-floating, four-wheel independent suspension that gives Subaru incredible road sense. And the new Subaru body styles have impact-absorbing structure, front and rear, to protect your valuable possessions. Like life. And limb.

4-WHEEL DRIVE FOR CROSS-TOWN AS WELL AS CROSS-COUNTRY.

Subaru four-wheel drive cars are comfortable, civilized and economical.

23 EST. MPG (CITY)* 33 EST. HWY.*

Subaru 4WD's handle nimbly and respond quickly. And shift from front-wheel to four-wheel drive at up to 50 mph.

So whether you choose four-wheel drive or front-wheel drive, your Subaru is engineered to carry you safely over the rough road ahead. Without putting you in the hole.

SUBARU®

INEXPENSIVE. AND BUILT TO STAY THAT WAY.

*1980 EPA ESTIMATES. USE ESTIMATED CITY MPG FOR COMPARISONS. YOUR MILEAGE MAY DIFFER DEPENDING ON DRIVING SPEED, WEATHER CONDITIONS AND TRIP LENGTH. ACTUAL HWY MILEAGE WILL PROBABLY BE LESS THAN HWY ESTIMATE. ACTUAL CITY MILEAGE WILL BE LESS IN HEAVY TRAFFIC. CALIFORNIA ESTIMATES LOWER. © SUBARU OF AMERICA, INC. 1979

Take any 3 books

values to $76.90

for only 99¢ each

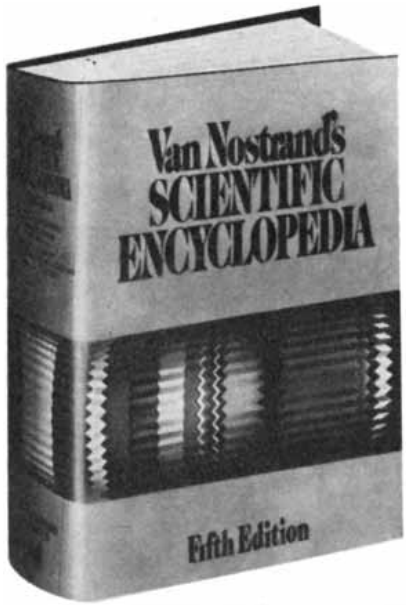
if you will join now for a trial period and agree to take 3 more books—at handsome discounts—over the next 12 months.

as a member of

Natural Science Book Club

EXTRAORDINARY VALUE!

The greatest single-volume source of scientific information available



00470-3. VAN NOSTRAND'S SCIENTIFIC ENCYCLOPEDIA. *Edited by Douglas M. Considine.* Nearly 200 experts have contributed to the thoroughly revised and expanded Fifth Edition of the most authoritative single-volume source of scientific information ever assembled. Over 2,500 illustrations. *Counts as 3 of your 3 books.* **$67.50**

(Publishers' Prices shown)

80250. STARS: Their Birth, Life, and Death. *Josif S. Shklovskii.* **$17.50**

55000. THE ILLUSTRATED ENCYCLOPEDIA OF ASTRONOMY AND SPACE. $16.95

41605. DARWIN'S FORGOTTEN WORLD. Spectacular photographic journey to the Galapagos Islands. More than 200 color photographs. **$19.95**

50800-2. GRZIMEK'S ENCYCLOPEDIA OF ETHOLOGY. Comprehensive exploration of animal behavior. 700 pages of scientific, lucid information. 774 photographs, drawings and diagrams. A landmark reference. *Counts as 2 of your 3 books.* **$39.50**

33005. THE AMERICAN WALK BOOK. *Jean Craighead George.* An illustrated guide to the country's major historic and natural walking trails from New England to the Pacific Coast. **$13.95**

62915. MONUMENTS OF CIVILIZATION: Ancient Cambodia. *Mazzeo and Antonini.* Portrays the magnificence of ancient Khmer civilization—over 100 color photographs. Oversized volume. **$25.00**

63845. THE NATURAL HISTORY OF THE WHALE. *L. Harrison Matthews.* **$19.95**

87750. THE WORLD OF DINOSAURS. *Michael Tweedie.* **$14.95**

69328. PLANETS AND MOONS/STARS AND NEBULAS. *William J. Kaufmann III. The two count as one book.* **$26.00**

Discover the secrets of the natural world

40695. COSMOS, EARTH AND MAN: A Short History of the Universe. *Preston Cloud* **$14.95**

84246. THREE ADVENTURES/LIFE AND DEATH IN A CORAL SEA/DIVING COMPANIONS. *Cousteau and Diole. The three count as one book.* **$23.85**

36395. BLACK HOLES, QUASARS AND THE UNIVERSE. *Henry L. Shipman.* An absorbing exploration of the frontiers of astronomy. **$15.50**

74300-2. ROME AND HER EMPIRE. *Barry Cunliffe.* A magnificently illustrated volume that makes the history of one of the world's greatest civilizations come alive. *Counts as 2 of your 3 books.* **$50.00**

63217. MURMURS OF EARTH. *Carl Sagan et al.* The story of the recorded messages carried by the *Voyager* space probe, man's first attempt at interstellar communication. **$15.00**

54995. THE ILLUSTRATED ENCYCLOPEDIA OF ARCHAEOLOGY. *Glyn Daniel.* **$17.95**

32900-2. AMERICAN RIVERS. *Bill Thomas.* An exciting textual and magnificent photographical exploration. *Counts as 2 of your 3 books.* **$30.00**

61610. MAYA: The Riddle and Rediscovery of a Lost Civilization. Charles Gallenkamp. **$15.95**

82771. TEMPLES, TOMBS AND HIEROGLYPHS/RED LAND, BLACK LAND. 2 excellent introductions to the complex and fascinating subject of Egyptian history. *The two count as one book.* **$25.90**

73939. THE RESTLESS UNIVERSE/WHITNEY'S STAR FINDER. *The two count as one book.* **$21.90**

41510-2. THE DANCE, ART, AND RITUAL OF AFRICA. *Michel Huet, et al.* A spectacular, outsize volume which records the pageants and rituals of Africa in more than 250 magnificent photographs. *Counts as 2 of your 3 books.* **$35.00**

34210. THE ARCHAEOLOGY OF NORTH AMERICA. *Dean Snow.* Prehistoric Indian cultures, from the first crossings into this continent to the 20th century. 195 photos, charts and maps. **$18.95**

35506. BACKPACKING EQUIPMENT BUYER'S GUIDE/WALKING SOFTLY IN THE WILDERNESS: The Sierra Club Guide to Backpacking/SIMPLE FOODS FOR THE PACK. *Softcover. The 3 count as one book.* **$19.85**

56208. AN INTRODUCTION TO ANCIENT EGYPT. *T.G.H. James.* Explores art, history, religion and culture of ancient Egypt—from pre-dynastic Badarian farmers to the deaths of Antony and Cleopatra in 30 B.C. 100 black-and-white photos, 20 full-color plates. **$20.00**

79975. THE SPOTTER'S GUIDES TO BIRDS, TREES, WILDFLOWERS, ROCKS & MINERALS, SHELLS, AND THE NIGHT SKY. Six compact field guides, each illustrated in full color. **$23.70**

63350. MYSTERIES OF THE MEXICAN PYRAMIDS. *Peter Tompkins.* Stunning array of 500 photos, drawings and maps illustrates the fascinating Mesoamerican pyramids which have mystified archaeologists and laymen for centuries. **$20.00**

44930-2. THE ENCYCLOPAEDIA OF IGNORANCE. Fascinating encyclopedic exploration of the frontiers of scientific research. 58 internationally renowned scientists highlight topics ranging from evolution to black holes. *Counts as 2 of your 3 books.* **$30.00**

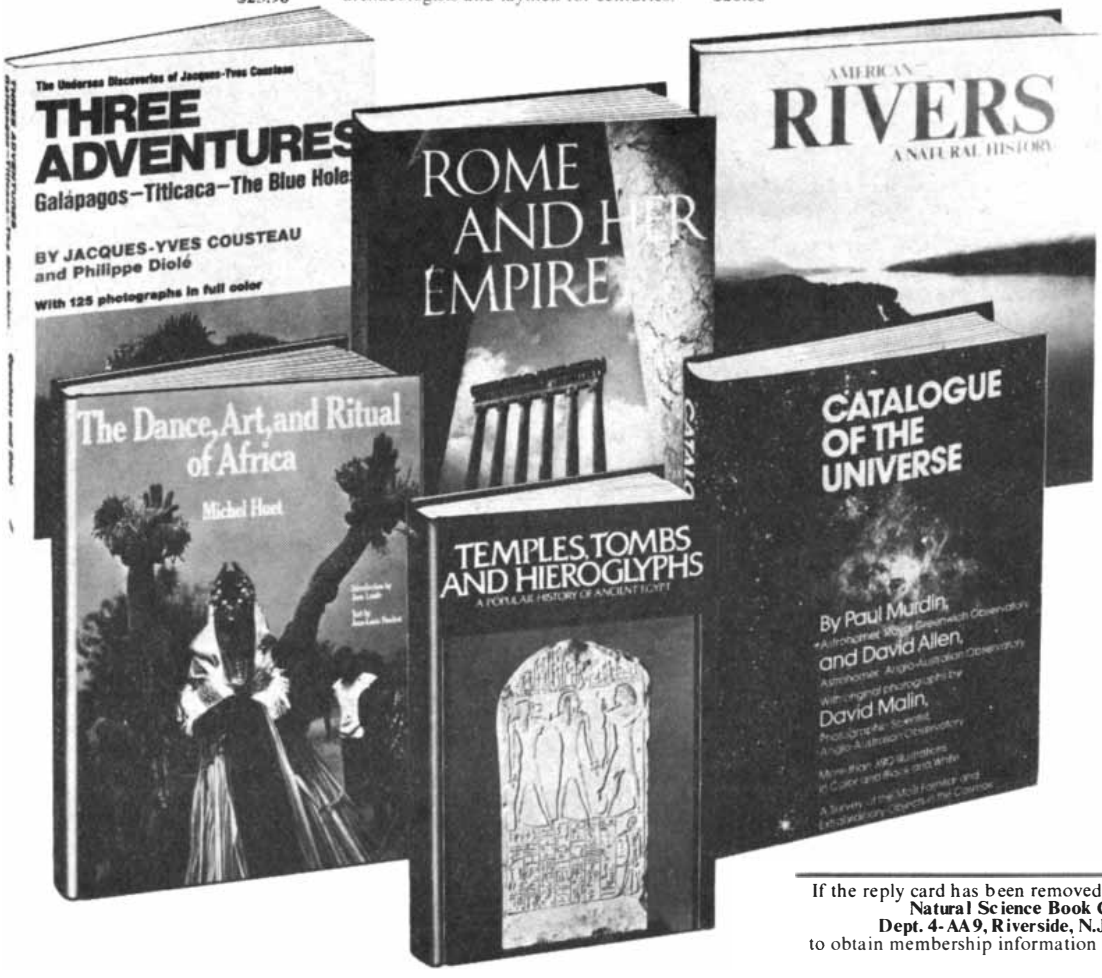
77570. THE SECRET OF ATLANTIS. *Otto Muck.* A dramatic re-creation of the life and death of Atlantis, which as Muck postulates, was destroyed in the course of a single day by a cosmic explosion of unparalleled proportions. **$12.50**

77579-2. SECRETS OF THE ICE AGE/THE CAVE ARTISTS. Vividly re-creates the life, environment and habits of Upper Paleolithic man. *Plus* explores the craft, styles and meaning of prehistoric cave art. Both books beautifully illustrated. *Counts as 2 of your 3 books.* **$31.90**

87890-2. WORLD PREHISTORY IN NEW PERSPECTIVE. *Grahame Clark.* Enormous, global survey of *all* civilizations from the first palaeolithic gatherings to the Aztecs. *Counts as 2 of your 3 books.* **$29.95**

37353. CATALOGUE OF THE UNIVERSE. *Paul Murdin, et al.* Three volumes in one: an atlas, encyclopedia, and book of descriptive astronomy. **$15.95**

34730. THE ATLAS OF EARLY MAN. *Jacquetta Hawkes.* A panoramic overview of concurrent events in history through charts, graphs, and illustrations, from prehistory to the Roman Empire. **$20.00**



If the reply card has been removed, please write to
Natural Science Book Club
Dept. 4-AA9, Riverside, N.J. 08370
to obtain membership information and application.

his investigation of the origin of *Zea mays*—a long-standing research interest—by retiring in 1968 as president of the University of Chicago. He began work on this project almost 50 years ago as a graduate student at Cornell University, where he collaborated with his thesis adviser, the late plant geneticist R. A. Emerson, in determining that "cytologically and genetically the wild teosinte plant native to Mexico and Guatemala is the most likely direct ancestor of corn." After Beadle received his Ph.D. in genetics in 1931, he went to the California Institute of Technology, where he continued his research on corn but gradually shifted to studies of the fruit fly *Drosophila melanogaster*. Later, as professor of biology at Stanford University, he collaborated with Edward L. Tatum and others on a new approach to an understanding of the relation of genes to enzymes, working with the red bread mold *Neurospora crassa* as their experimental organism. (It was for this work that Beadle and Tatum shared the 1958 Nobel prize in medicine or physiology.) Beadle returned to Cal Tech in 1946 to become chairman of the division of biology, a position he left in 1961 to go to Chicago.

MARTIN M. KAPLAN and HILARY KOPROWSKI ("Rabies") have worked together on and off for almost 30 years on the study of rabies, mainly at the Wistar Institute of Anatomy and Biology, where Kaplan is a frequent visiting investigator and Koprowski is director and institute professor. Kaplan currently serves as director general of the Pugwash Conferences on Science and World Affairs, based in Geneva. His degrees are in veterinary medicine and public health, obtained respectively from the School of Veterinary Medicine and the Graduate School of Medicine of the University of Pennsylvania. He was a member of the staff of the World Health Organization (WHO) for 26 years, retiring in 1976. Koprowski, in addition to his duties at the Wistar Institute, is professor of research medicine at the University of Pennsylvania. He received his M.D. from the University of Warsaw in 1939 and came to the U.S. in 1944. After 12 years at the Lederle Laboratories of the American Cyanamid Company he joined the Wistar Institute in 1957.

NORMAN SMITH ("The Origins of the Water Turbine") is a lecturer on the history of technology at the Imperial College of Science and Technology in London. He is also coeditor (with A. R. Hall) of the annual publication *History of Technology*. He obtained his Ph.D. in civil engineering from the University of Bristol in 1962 and taught engineering at the University of Canterbury in New Zealand for a time before going to Imperial College in 1965.

If you can't pronounce "Entschuldigen Sie" your computer can.

From Texas Instruments: The World's FIRST two-way language translator to pronounce and display foreign words and phrases.

Texas Instruments is America's largest manufacturer of consumer electronics

Just imagine! Two people speaking different languages, ordinarily not able to understand each other, can now communicate. It's very easy to have a two-way conversation. Simply press the buttons on the TI Language Translator and spell out a word or phrase. The computer will immediately translate visually and vocally. For example, press the keys that spell out the phrase, "Do you have any messages for me?" in English and the TI Translator will spell out "AVEZ-VOUS DE MESSAGES POUR MOI?" on the display. At the same time the translator will pronounce the phrase in French.

The TI Language Translator is the world's only hand-held, computerized translator that can actually SPEAK! and pronounce foreign words and phrases, using correct accents and idioms. At the same time, it displays in large, clear illuminated letters and proper spelling.

In the past other translators have made exaggerated claims about their ability to "talk" when, in fact, all they could do was display words. But the TI Language Translator literally TALKS in an articulate voice with the idioms associated with that language.

Because it's small enough to carry in your pocket or purse, you can take it anywhere. In restaurants, airports and hotels. In taxis, shops, or on a street corner. It's great for business meetings, travel or foreign language study. Let the translator help you learn a language by familiarizing you with the most commonly used words and over 1000 phrases. If you're already familiar with the language, use the translator to brush up on spelling and sharpen your speaking skills.

Every plug-in Language Module can translate interchangeably four languages: English, French, Spanish, German, and speaks the one language of your choice. And many more languages are on the way. Just plug in the module and the translator will speak that language. There are no tapes or records to wear out or fumble with. It's like having a private tutor you can take with you anywhere.

The TI Language Translator comes complete with carrying case, earphone, a.c. adapter (110V), and a word/phrase guide with instructions. Requires four AA batteries and language module (not included). A one year manufacturer's warranty is included.

Don't leave home without it!

Try the TI Language Translator for 14 days with no risk. If you are not completely satisfied your money will be completely refunded.

The TI Language Translator costs $300.
Language modules cost $60 each.

Please order the module(s) you wish.

☐ Spanish speaking only ☐ French speaking only
☐ English speaking only ☐ German speaking only

CREDIT CARD BUYERS: TO ORDER CALL TOLL FREE 24 HOURS A DAY
To order in California call toll free (800) 432-7451
(800) 854-3831

DWS marketing international, Dept. 21, 350-A Fischer Avenue
Costa Mesa, California 92626 • (714) 540-4444

Call our toll-free number for quickest service or send your check, Money Order, or credit card info to us. Please allow ample time for checks to clear. We accept Am. Ex., Carte Blanche, Diner's Club and all major credit cards. Add $2.50 for insured postage & handling. Calif. addresses add 6% sales tax.

MATHEMATICAL GAMES

Checkers, a game that can be more interesting than one might think

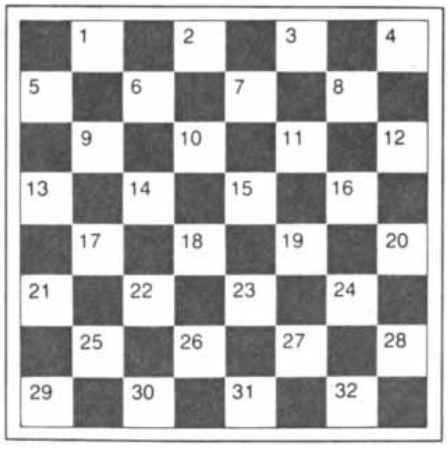
by Martin Gardner

"The game of draughts we know is peculiarly calculated to fix the attention without straining it. There is a composure and gravity in draughts which insensibly tranquillises the mind."

—JAMES BOSWELL,
The Life of Samuel Johnson

The quotation is from a section for the year 1756 in which Boswell writes about Johnson's preface to William Payne's *Introduction to the Game of Draughts,* published the same year in London. That book, by a mathematics teacher, was the first in English on the game that in the U.S. is known as checkers. Johnson seldom played the game after leaving college. Boswell expresses regret over it because he thinks checkers playing would have afforded his friend "innocent soothing relief" from periodic bouts of depression.

Nothing is known about the beginnings of checkers, although most game historians now think it originated in southern France sometime in the 12th century. In Britain and the U.S. it is surely the best-known of all board games when you consider the number of children who learn to play it and never forget its rules, even though checkers is far below chess in the size of its literature, in the number of adults who become top-level players and in the public excitement generated by contests for the world checkers championship. How many people can name a single checkers expert or tell you who the current world champion is? He is Dr. Marion F. Tinsley, a topologist in the department of mathematics at Florida A. and M. University and probably the greatest checkers player who ever lived.



Squares are numbered for checkers notation

Rules for chess are now standard throughout the Western world, but not so for checkers. Outside of English-speaking countries there are dozens of regional variations. The version most popular in Europe and the U.S.S.R., called Polish checkers (except in Poland, where it is called French checkers), is played on a 10-by-10 board, each side starting with 20 men. It is the standard French form of the game. In French Canada the board is even larger: 12-by-12, with 30 pieces to a side. Rules for checkers differ widely around the world. It is curious to note that in all European countries except Britain the pieces are called ladies; only here and in English-speaking countries are they men.

Several consequences follow from the fact that checkers is simpler than chess. One is that a grand-master checkers player is less likely than his chess counterpart to lose to an inferior by making an error. For checkers buffs this is one of the game's great attractions. They love to quote Edgar Allan Poe's discussion of the two games at the beginning of *The Murders in the Rue Morgue:*

"I will, therefore, take occasion to assert that the higher powers of the reflective intellect are more decidedly and more usefully tasked by the unostentatious game of draughts than by all the elaborate frivolity of chess. In this latter, where the pieces have different and *bizarre* motions, with various and variable values, what is only complex is mistaken (a not unusual error) for what is profound. The *attention* is here called powerfully into play. If it flag for an instant, an oversight is committed, resulting in injury or defeat. The possible moves being not only manifold but involute, the chances of such oversights are multiplied; and in nine cases out of ten it is the more concentrative rather than the more acute player who conquers. In draughts, on the contrary, where the moves are *unique* and have but little variation, the probabilities of inadvertence are diminished, and the mere attention being left comparatively unemployed, what advantages are obtained by either party are obtained by superior *acumen.*"

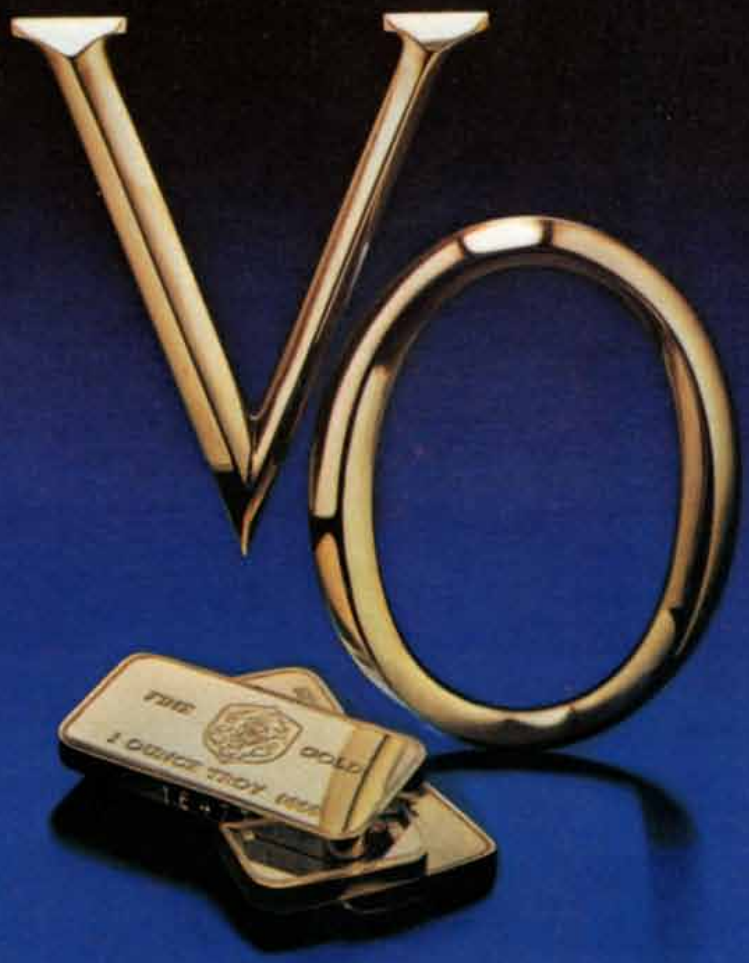
Tinsley has put it this way: "Playing chess is like looking out over a limitless ocean; playing checkers is like looking into a bottomless well."

Another consequence of the simplicity of checkers is that by 1900 the game's openings had been so completely analyzed that most tournaments ended in draws. To inject more drama into the play Britain introduced (in about 1900) the practice of putting on cards every pair combination of Black's first move and White's response. Before each match a card was chosen at random, and the game had to be played with the specified pair of opening moves. Since each side has a choice of seven moves, there are 49 possible pairs. Two of them (9–14, 21–17, and 10–14, 21–17) were ruled out because they give away a white piece. Later it was found that two more pairs (11–16, 23–19, and 12–16, 23–19) give Black such a strong advantage that they too were discarded, leaving 45 cards.

Standard checkers notation is based on the numbering of squares shown in the illustration on this page. For reasons of clarity it is customary in checkers diagrams to reverse the colors of squares and show the pieces on white cells instead of black. Actual play is always on black squares, with the "double corner" at each player's lower right. The players are customarily called Black and White even though the pieces are red and white. Tournament games are now played on green-and-buff boards; black-and-red boards are considered toy-store atrocities. Black always moves first, and games are recorded with Black starting on the low-numbered cells. If you work on any of the recreations in this column, it is a good idea to label the black squares of your board as shown.

Alas, as decades went by experts soon became so familiar with all variations that follow the two-move openings that "safe" play was adopted and the draws began to pile up again. The British "two-move restriction" was replaced in the U.S. in the mid-1930's by the "three-move restriction," a practice now followed in most checkers tournaments here and in Britain. There are 142 cards, each with a different triplet of the first three moves. Because many of these triplets give an advantage to one side (usually the second player) two games are played with each selection to allow each player the first move.

Some of the best things are measured by the ounce.



Seagram's V.O.
The symbol of imported luxury. Bottled in Canada.
Enjoy our quality in moderation.

Canadian whisky. A blend of Canada's finest whiskies. 6 years old. 86.8 Proof. Seagram Distillers Co., N.Y.C.

Ford Fiesta.
It received a seven-flag salute.

The car that wowed Europe is winning the hearts of America.

Ford Fiesta. The little front wheel drive car that comes from Germany. Applauded all over Europe by the experts for its engineering, design and overall performance. It was voted the most significant import of the year in 1978 by readers of Car and Driver. And when you understand how beautifully Fiesta performs, you'll understand why it continues to get rave notices.

Germany 1976
Car that Makes the Best Sense — *Mot*

Finland 1976
Car of the Year
— *Tuulilasi*

Front wheel drive.

Denmark 1976
Car of the Year
— *Morgen Posten*

Through rain, ice, sleet, hail and snow, Fiesta's remarkable front wheel drive traction will help you keep your appointed rounds. And its Michelin steel-belted radials will help you come to grips with all kinds of roads.

Italy 1977
Most Successful
Foreign Car — *Motor*

Best gas mileage of any German import.

Americans love Fiesta's manners. It prefers sipping to guzzling. 1979 EPA est. mpg:

28 EPA EST. MPG **39** HWY. EST. MPG

Compare this estimate to other cars. Diesels excluded. Actual mileage may differ depending on speed, weather and trip length. California estimates will be lower. Actual highway mileage will probably be lower.

Yugoslavia 1977
Car of the Year
— *Automotive Writers*

A masterpiece of European engineering.

Fiesta is assembled by Ford in Germany. And its European engineering makes it feel right at home on streets and highways of America. It's quick, nimble and maneuverable.

Ford Fiesta is sold and serviced by over 5,000 authorized Ford Dealers across America. There's even an Extended Service Plan available, providing longer protection over your car's basic warranty. So test-drive a Fiesta today. You'll discover why it's won international acclaim.

Spain 1977
Car of the Year
— *Criterium*

Great Britain 1978
Design Council Award

FORD FIESTA

FORD DIVISION Ford

Fiesta. Wundercar from Germany.

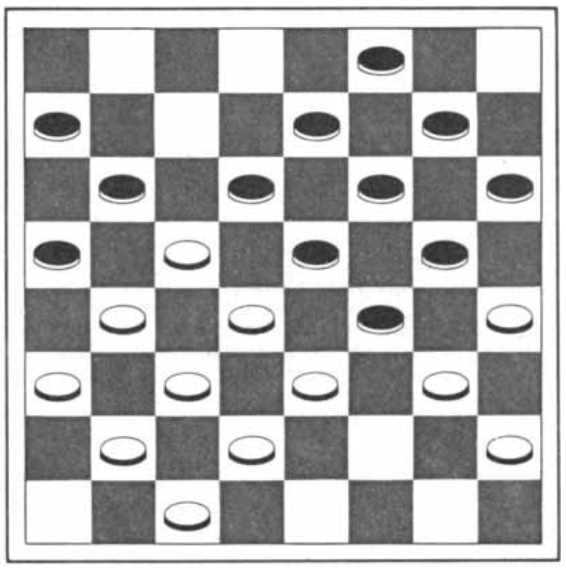
Fiesta 3-Door Sport

Without the opening-moves restrictions, a practice known as go-as-you-please play, experts would play nothing but draws. Even with the three-move restriction about 80 percent of all tournament games still end in draws. When an expert does win, it is usually because the loser made a blunder or because the winner managed to keep secret (sometimes for years) a "cook" he had discovered. As in chess, a cook is an improvement on standard "book play" that catches an opponent by surprise. Players have traditionally been allowed only five minutes to think before each move and one minute for a capture that can be made only one way. In recent years this practice has been replaced by the use of chess clocks, and players are allowed 30 moves in an hour. When someone springs a new cook, his victim simply does not have enough time to analyze it.

In 1967 the late Walter Hellman, a steelworker in Gary, Ind., who was then world champion, defended his title against the U.S. champion, Eugene Frazier. The contest went to 36 games, of which 31 were draws and five were wins by Hellman. Hellman's last win was on a cook. "I had used that cook once before," Hellman told a reporter, "but it had never been published. Frazier had one possible move to thwart the attack, and five minutes doesn't allow much time to figure it out."

A third consequence of the simplicity of checkers is that the best computer programs for checkers play a more formidable game against middle-level players than the best computer programs for chess. Until a few years ago the strongest checkers program was the work of Arthur L. Samuel, a learning program that improves as it plays. Since retiring as IBM's director of research, Samuel has continued to improve his program at Stanford University's Artificial Intelligence Laboratory. In recent years a powerful program of the nonlearning type has been developed by Eric C. Jensen and Tom R. Truscott, two graduate students at Duke University working under Alan W. Biermann, who teaches artificial intelligence. It is thought to be stronger than Samuel's program, although this has not yet been determined by a definitive match.

Checkers players are ranked on three levels: minor, major and master. Backers of the Duke program believe it plays initially on a master level. After playing against the program for a while, however, a grand master can discern its weaknesses and begin to exploit them. Its greatest weakness is that it plays without master plans. It does not even follow book moves in opening play, usually scattering its pieces over the board in patterns grand masters consider stupid. Its strength is the incredible speed with which it can analyze all possible moves to much greater depths than a human opponent, and within those depths it never makes a mistake. Chess programs may still be decades away from grand-master play, but the Duke program, Biermann believes, is already "knocking at the door" of the world championship. Jensen and Truscott regard it as now being about the 10th strongest player in the world.

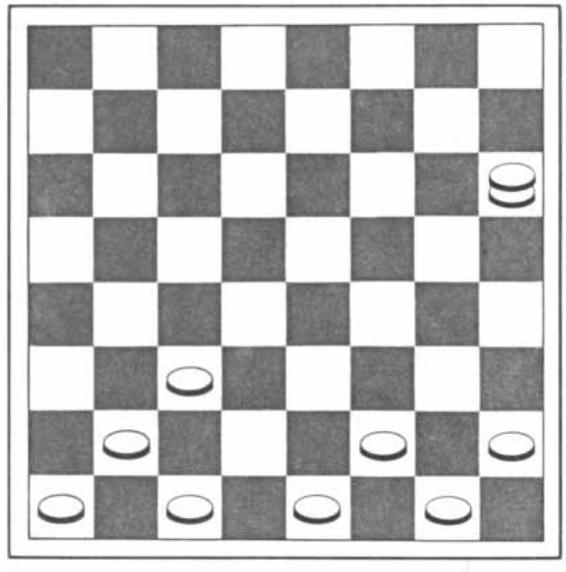


BLACK	WHITE
1. 9-13	24-20
2. 12-16	21-17
3. 10-15	23-18
4. 15-19	18-14
5. 8-12	25-21
6. 4-8	29-25
7. 6-10	27-23
8. 10-15	23-18
9. 2-6	31-27
10. 6-9	27-24
11. 1-6	32-27
12. 6-10	27-23

The shortest no-capture checkers game

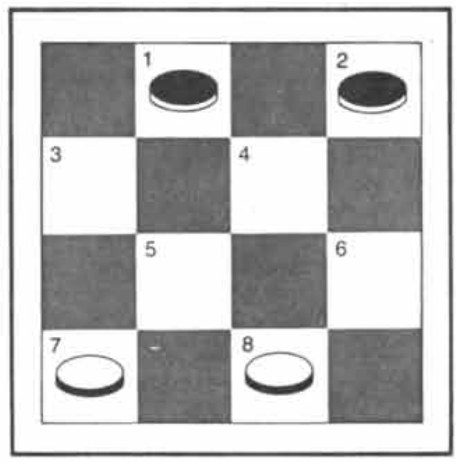
Grand-master checkers players, like their chess counterparts, take a much dimmer view of the quality of computer programs. They all agree with W. Burke Grandjean, secretary of the American Checker Federation, who considers the optimism of the Duke group to be ludicrously naive. Backed by the federation, Tinsley has a standing bet of $5,000 that in a stake match of 20 games he can beat any computer program devised over the next five years. Biermann thinks Tinsley would now defeat the Duke program, but not without losing a few games. Grandjean says he can easily find 20 players who would make side bets of $100 each that Tinsley would not lose a game, and he personally believes Tinsley would win all 20. Perhaps we shall soon find out. (Readers interested in joining the American Checker Federation and receiving its monthly *Bulletin* can write to Grandjean at 3475 Belmont Avenue, Baton Rouge, La. 70808.) Fidelity Electronics now has on the market Checker Challenger 2, an inexpensive solid-state machine that plays on two levels, and also Checker Challenger 4, which plays on five levels, although its top level is considered below the levels of the Samuel and Duke programs.

In chess it is easy to prove that the "fool's mate," in which the second player checkmates on his second move, is the shortest possible chess game. Surprisingly, the shortest checkers game is not yet known. Until two years ago it was thought to be the 24-move blocked game shown in the illustration above, in



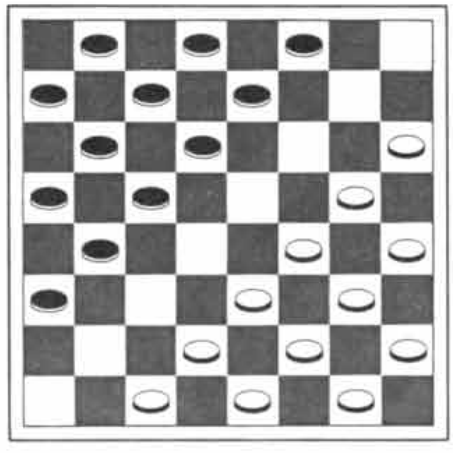
BLACK	WHITE
1. 11-16	21-17
2. 10-14	17x10
3. 6x15	23-18
4. 2-6	18x2 (K)
5. 9-14	2x18
6. 3-7	24-20
7. 1-6	20x2 (K)
8. 12-16	2x9
9. 5x23	26x3 (K)
10. 4-8	3x12

The shortest checkers game known

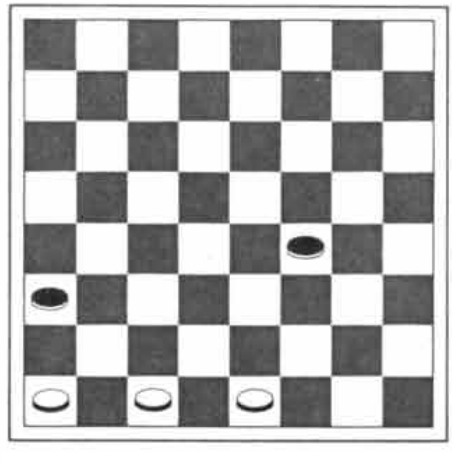


A minicheckers puzzle

which the board displays the final position. There are many sequences of 24 moves that lead to this position, but the position itself is thought to be unique. In the line of play given every White move is symmetrically opposite (with respect to the board's center) to Black's preceding move. I do not know who first put the play in this symmetrical form. The version I give, worked out by Rudolf Ondrejka of Linwood, N.J., begins with the



Diagonal checkers



A hustler's checkers bet

two-move Edinburgh opening. Because 9–13, a favorite first move among tyros, is considered the worst possible start for Black, the symmetrical game is more often started with 10–15, 23–18, an opening known as the Kelso Cross.

Sam Loyd, in his *Cyclopedia of Puzzles* (1914), page 379, using an eccentric notation that incorrectly assumes the board has been rotated 90 degrees, records a nonsymmetrical sequence of moves ending with the same pattern. Loyd states flatly that it is the "shortest possible game." The 24-move blocked game is indeed (as can be proved) the shortest game in which there are no captures. In 1978, however, Alan Malcolm Beckerson, problems editor of *English Draughts Journal,* discovered that White could win on his 10th move (20 moves in all) by capturing all Black's pieces! This is now the shortest checkers game known, although no one has yet proved that no game can be shorter. Beckerson found other 20-movers that capture all the black pieces, as well as some 20-movers that end in blocked games after some captures. The version given in the bottom illustration on page 25, with the board showing the final position, was first published in the British monthly *Games and Puzzles* for March, 1978. Its two-move opening is known as the Newcastle.

Many other minimum-move checkers tasks are far from settled. In how few legal moves can a game produce 24 kings? The best-known solution, in 90 moves, appeared in *Journal of Recreational Mathematics* (Vol. 9, No. 1, page 45; 1976). In how few moves can Black and White reverse their initial positions? It takes at least 60 moves for either side, alone on the board, to occupy the opposite starting cells, and so it follows that 2×60, or 120, is an absolute lower bound. A solution in 172 moves is given in a late-19th-century English book, *The Draughts-Player's Guide and Companion,* by Frank Dunne, pages 94–95. At the finish each side has six kings. It seems likely that 172 moves can be considerably lowered.

It is interesting to try this problem on smaller boards. The 3-by-3 is trivial, but the 4-by-4 presents a pleasant puzzle. Starting as is shown in the top illustration on the opposite page, the task is to interchange the two sides in a minimum number of legal moves. Captures are of course compulsory. At the finish all four pieces will necessarily be kings. I shall give a solution next month. Incidentally, five moves are needed for the shortest game on this miniboard. If both sides play to win and follow their best strategy, the game is a draw.

As with chess, endless ways of playing checkers have been proposed by varying the shape of the board, the starting position, the rules and so on. A privately published French book, *Les jeux de*

dames non orthodoxes et autres jeux à pions, by Joseph Boyer and Vern R. Parton, gives more than 100 such variants. Some are played on triangular or hexagonal tessellations and some on three-dimensional boards; some mix chess pieces with checkers and some allow three or four players to compete at once. As one would imagine, it is hard to draw a line between a game similar enough to checkers to be called a variant and one so different from checkers that it is best regarded as another game altogether. The so-called Turkish checkers, for example, has almost no resemblance to checkers except that it is played on an 8-by-8 board with counters of two colors. One simple way to vary standard checkers is to start with the men positioned as is shown in the middle illustration on page 26. All checkers rules hold. The opening moves quickly lead to patterns not encountered in orthodox games.

The most eccentric variation of checkers, about which one would like to know more, is "supercheckers," invented by Charles Fort, the Bronx collector of science anomalies who had a marked influence on science fiction as well as on the current epidemic of interest in the paranormal. According to Fort's biographer Damon Knight (*Charles Fort,* Doubleday, 1970), supercheckers was played "with armies of men on a board with thousands of squares. Fort used bits of cardboard with carpet-tack handles for the men, and a piece of checkered cloth for the board."

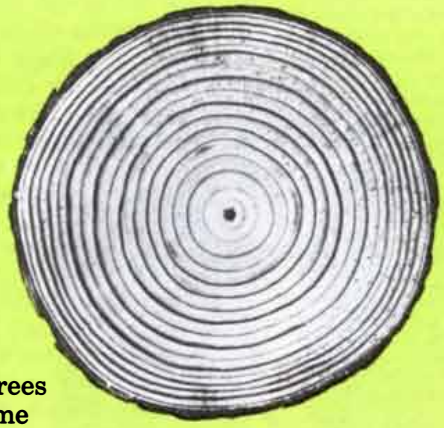
The two players start with their forces in any agreed-on formation that has a space between the two armies. If a player moved only one man at a time a game might last for weeks, and so Fort allowed for movements en masse. Here is how he put it in a letter: "Let *A* start out, moving until *B* tells him to stop—say a hundred moves. Then *B* makes a hundred moves. *A* may want to make another hundred moves, but *B,* sizing up the situation, tells him to stop, say at thirty. Then perhaps occurs 'fighting,' at close quarters, one move at a time, as in ordinary checkers. But, at any time, if either player wants to make a 'mass movement,' that is a matter of obtaining permission from his opponent."

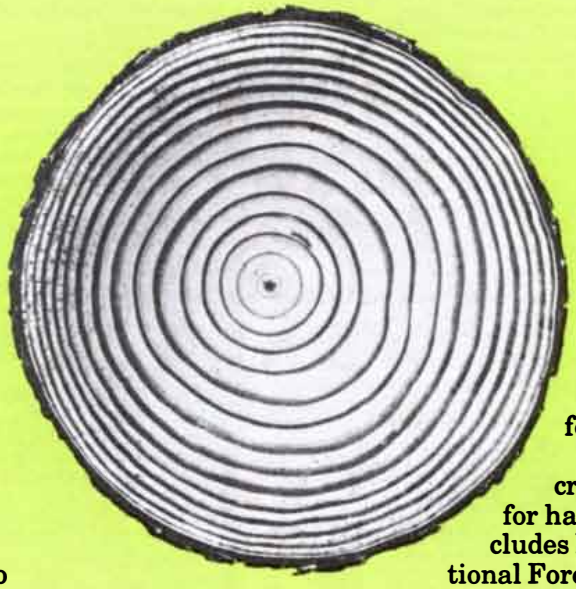
A game usually lasted all night. In 1930 Fort wrote to Tiffany Thayer, who edited the first Fortean magazine, *Doubt:* "Supercheckers is going to be a great success. I have met four more people who consider it preposterous."

In Britain and the U.S. the most popular variant of checkers is "giveaway." It differs from the standard game only in that the object is to be the first to lose all one's men. In Dunne's book cited above, pages 91–92, there is a fantastic giveaway "sucker bet," presumably devised by British checkers hustlers. White begins with his 12 men in the usual starting

A report on the great American forest.

Look what happens when Mother Nature gets a little help.





These log slices dramatically show the value of modern, scientific forest management. Both (shown 36% of actual size) are from trees of about the same age. The smaller one grew in an unmanaged, overcrowded stand where it had to compete for sunlight and moisture. The other is from a grove that was thinned to give the best trees room to thrive.

This is just one of many ways forest management can help meet a predicted doubling of domestic demand for wood and paper products in the next 50 years.

The forest industry has learned how to help Mother Nature grow more trees, faster, by applying scientific forest management techniques: encouraging natural regrowth, planting superior seeds and seedlings, fertilizing, protecting against disease and insects, thinning, watching carefully until time to harvest — then starting the cycle again.

Demand keeps rising.

Forest management like this is vital because every year Americans need more homes, more paper products, more packaging and containers, more fuelwood, more of the thousands of other essential products that only the forest can provide.

So far, the country's commercial forest has been able to keep up with demand. (*Commercial forest,* as defined by the U.S. Forest Service, is *all* forestland—whether owned by individuals, government or the industry—that is capable of, and potentially available for, growing repeated crops of trees for harvest. It includes land in National Forests but not in National Parks or Wilderness areas.)

But trees aren't grown equally fast by all commercial forest owners. Though industry has made striking advances on its lands, productivity is lower on privately owned lands and seriously lagging in National Forests.

Productivity not up to potential.

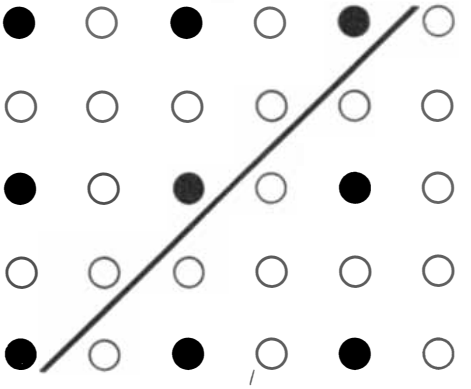
Overall, the U.S. Forest Service estimates average productivity of all commercial forestland is only 61 percent of potential. And at the same time, actual acreage in commercial forests keeps shrinking, as land is withdrawn for homes, highways and other needs of an expanding population.

So we as a nation still have a long way to go if our wood and paper products are not to become scarce and expensive.

If you'd like to be better informed on how important it is to keep America's forests productive, write American Forest Institute, P.O. Box 873, Springfield, VA 22150 for a free booklet, "The Great American Forest."

The great American forest. Trees for tomorrow. And tomorrow. And all the tomorrows after that.

Trees. America's renewable resource.



An Aha! solution

position. Black has only a king on cell 7. Black wins if he loses his king. White wins if he loses all 12 men. Dunne shows how White can always win and gives three similar wagers in which Black begins with a single uncrowned piece on cell 1, cell 4 or cell 5.

Among hundreds of hustler wagers, one of the best begins with the position shown in the bottom illustration on page 26. (I am indebted to Mel Stover for passing it along.) It is Black's turn. White wagers that Black will not be able to crown the piece he moves first. Clearly Black should not move the piece on cell 21 because he would lose it immediately, so that the question is whether Black can move the piece on cell 19 and advance it to his king's row. The more you study the pattern, the more obvious it seems that Black can win the bet easily. Nevertheless, White wins. It is an amusing bet to make with friends. If you cannot fathom White's strategy (and there are no joke catches), you will find the answer here next month.

One final problem. It is widely believed two kings can always win against one king, but that is not invariably true. See if you can place two white kings and one black king on the board in such a way that even though it is White's turn, Black can force a draw.

The first of last month's short problems was answered at the end of that column. The solutions to the remaining problems are as follows:

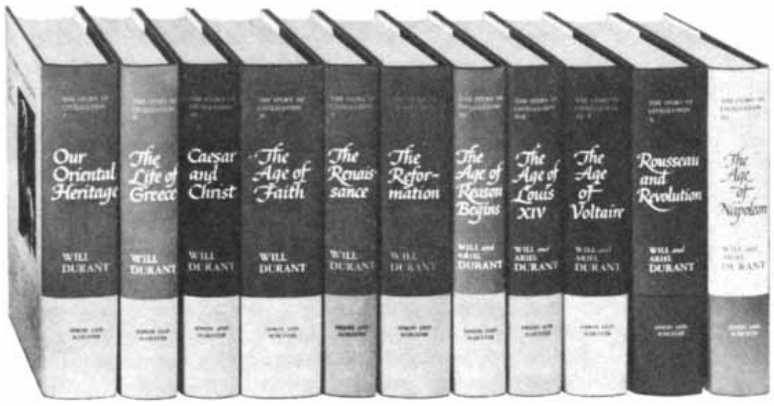
2. The poker puzzle is answered when we consider the fact that a hand with four identical values always has a fifth card. For each four of a kind there are 48 different fifth cards. Consequently there are 48 × 13, or 624, different poker hands containing four of a kind, compared with 40 hands that are straight flushes. It is therefore much less likely that you will be dealt a straight flush, and for this reason a straight flush beats four of a kind. The problem was contributed by M. H. Greenblatt to *Journal of Recreational Mathematics* (Vol. 5, No. 1, page 39; January, 1972).

3. Here is how Raymond Smullyan, in

The best sources.

Choose any one of these invaluable sets as your introduction to membership in the Book-of-the-Month Club. You simply agree to buy 4 books within a year.

The Club is the source for the best of the best sellers. Every Club book is of bookstore quality—or better. You *don't* settle for the altered or inferior editions that some book clubs offer. And when you remain a member after the short trial period, each book you buy earns Book-Dividend® credits, which entitle you to discounts of at least 70% on significant current books. A wealth of sources awaits you in America's Bookstore.

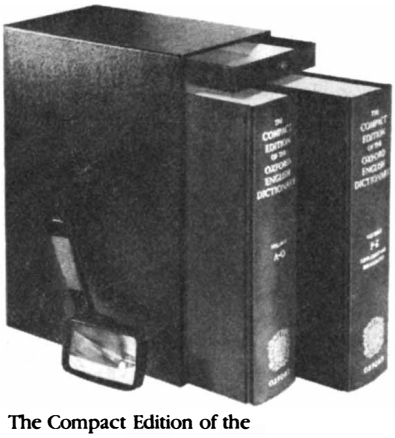


All eleven volumes of

THE STORY OF CIVILIZATION

by Will and Ariel Durant for **$24.95** Publisher's list price: $241.10

This multi-volume history of all mankind—ancient and modern, Occidental and Oriental—traces the development of man's economic and political organizations, science and art. It is one of the great literary achievements of this century and deserves a place in every library.



The Compact Edition of the

OXFORD ENGLISH DICTIONARY

for **$19.95** Publisher's list price: $125

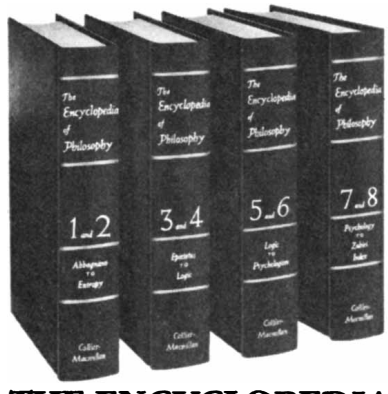
The world's "most complete and most scholarly dictionary of the English language"—*The Christian Science Monitor*. Through a special photo-reduction process, every word of the original 13-volume set has been reproduced in this 2-volume *Compact Edition*. A Bausch & Lomb magnifying glass is included.

The Complete

BEETHOVEN NINE SYMPHONIES

The complete boxed set plus an illustrated sixteen-page booklet for **$10** List price: $62.82

Sir Georg Solti conducts the Chicago Symphony in this nine-record set, hailed by the San Francisco *Sunday Examiner & Chronicle* as "one of the two or three great recording ventures of the century." *Time* called it "a glorious musical combustion" and *Stereo Review* declared, "This Beethoven set ranks first among those of recent date."



THE ENCYCLOPEDIA OF PHILOSOPHY

The complete four-volume set for **$19.95** Publisher's list price: $125

Regarded as the most comprehensive encyclopedia of philosophy ever published, this superb set—compiled in collaboration with the world's foremost philosophers—encompasses all aspects of ancient, medieval, modern, Eastern and Western thought. The four volumes represent an essential aid for students and a rewarding reference source for home libraries.

Prices shown are publishers' U.S. prices. Outside the U.S., prices are generally somewhat higher.

Facts About Membership.

You receive the *Book-of-the-Month Club News*,® 15 times a year (about every 3½ weeks). Each issue reviews a *Main Selection* plus scores of Alternates.

If you want the *Main Selection* do nothing. It will be shipped to you automatically. If you want one or more Alternate books—or no book at all—indicate your decision on the reply form always enclosed and return it by the date specified.

Return Privilege. If the *News* is delayed and you receive the *Main Selection* without having had 10 days to notify us, you may return it for credit at our expense.

Cancellations. Membership may be discontinued, by either you or the Club, at any time after you purchased 4 additional books.

Book-of-the-Month Club, Inc., Camp Hill, Pennsylvania 17012 A170-1

Please enroll me as a member of Book-of-the-Month Club and send me the work I have checked below, billing me for the appropriate amount, plus shipping and handling charges. I agree to buy 4 books during the coming year. A shipping and handling charge is added to each shipment.

CHECK ONE BOX ONLY

- ☐ 904. Beethoven Symphonies $10
- ☐ 912 Compact OED $19.95
- ☐ 913 The Story of Civ. $24.95
- ☐ 917. Ency. of Philos. $19.95

Mr.
Mrs. 0-71
Miss (Please print plainly)

Address ..... Apt. .....

City .....

State ..... Zip .....

BOOK-OF-THE-MONTH CLUB®
AMERICA'S BOOKSTORE®

his new book *The Chess Mysteries of Sherlock Holmes* (Knopf, 1979), proves that, given the chess position shown last month, Green made the first move.

"Red is now in check, hence Green moved last. It remains to determine who moved first, which can be done by figuring whether an odd or an even number of moves have been made.

"The rook on *b*1 has made an odd number of moves; the other three rooks have each made an even number of moves (possibly zero). The Red knights have collectively made an odd number of moves, since they are on squares of the same color, and the Green knights have collectively made an even number of moves. [A knight changes square color on each move.] One king has made an even number of moves (possibly zero), and the other king an odd number. The bishops and pawns have never moved, and both queens were captured before they ever moved. So the grand totality is odd. Thus Green moved first. Hence Green is White and Red is Black."

4. Surprising as it may first seem, both the richest and the poorest classes in Oilaria would prefer pair averaging from the top down. Those in the richest class would prefer to be averaged with the next-richest class before the latter is reduced in wealth by averaging. Those in the poorest class would prefer being averaged with the next-poorest class after the latter has been increased in wealth by averaging.

An example will make this clear. Assume that the wealth of the five classes is in the proportions 1 : 3 : 4 : 7 : 13. Averaging from the bottom up changes the proportions to 2 : 3 : 5 : 9 : 9. Averaging from the top down changes the proportions to 3 : 3 : 5 : 7 : 10.

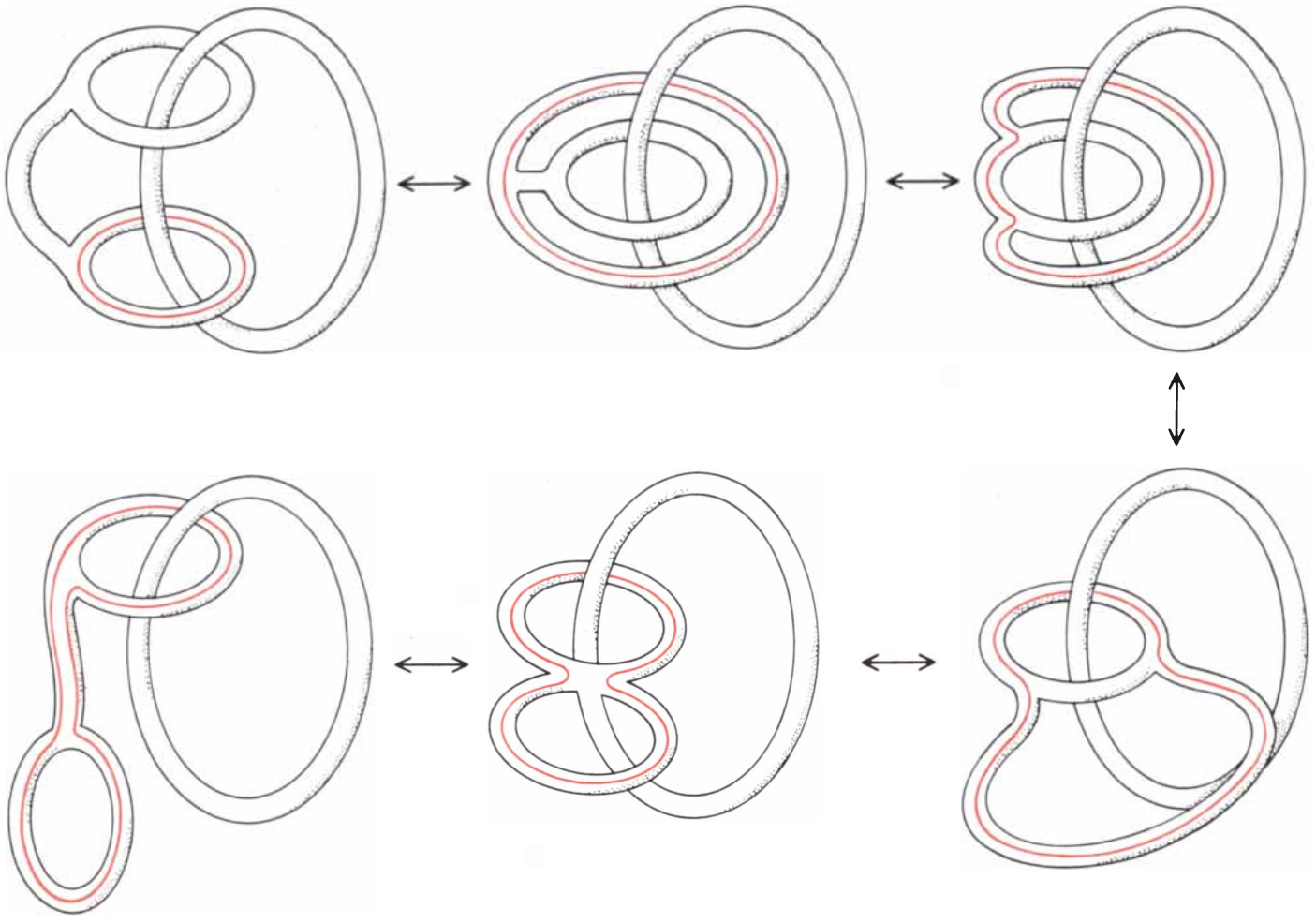
5. Divide the 500-mile track into 10 segments of 50 miles each. If any segment is traversed in one hour, the problem is solved, and so it must be assumed that traversing each segment takes either less than an hour or more than an hour. It then follows that somewhere along the track there will be at least one pair of adjacent segments, one (call it *A*) traversed in less than an hour and the other (call it *B*) traversed in more than an hour.

Imagine an enormous measuring rod 50 miles long that is placed over segment *A*. In your mind slide the rod slowly in the direction of segment *B* until it coincides with *B*. As you slide the rod the average time taken by the train to go the 50 miles covered by the rod varies continuously from less than an hour (for *A*) to more than an hour (for *B*). Therefore there must be at least one position where the rod covers a 50-mile length of track that was traversed by the train in exactly one hour.

6. The Aha! insight that solves the counter-jumping puzzle is to color nine spots as is shown in the illustration on page 30. It is obvious that, no matter how jumps are made, a penny on any colored spot can go only to another colored spot.

There are six colored spots above the line and only three below it. Therefore, by the pigeonhole principle, there must be three pennies above the line that have nowhere to go below the line. The task of moving all the pennies to spots below the line cannot be accomplished unless at least three pennies, on three colored spots, are removed from the top triangular array. Remove any three such pennies and the transfer of the remaining 12 is a simple task.

7. The illustration below shows how a continuous deformation of the two-hole torus will unlink one of its holes from the single-hole torus. The argument given last month for the impossibility of this task fails because if a ring is painted around one hole (as is shown by the colored line), the ring becomes distorted in such a way that after the hole is unlinked the painted ring remains linked through the one-hole torus.



Solution to the linked-toruses problem

24 COLOR SLIDES of the Great Impressionist Masterpiece for only $1

(comparable value $18 to $24)

Includes this Deluxe Album and 48-page Gallery Lecture Guide—at no extra cost

GAZE in astonishment as Renoir's lovely "Seated Bather" (shown above) appears on your wall aglow with all the vibrant reds, blues and greens of the Impressionist palette. Or, thrill to Rembrandt's famed "Aristotle Contemplating the Bust of Homer"—one of the world's costliest works of art. Or, marvel as you view Titian's allegorical masterpiece "Sacred and Profane Love" just as it appears in Rome's Borghese Gallery.

Until today, unless you had the means to travel among the world's great museums and private collections, you simply could not know these paintings so *intimately*. Now, as a member of the *McGraw-Hill Color Slide Program of Art Enjoyment,* they are but a few of the magnificent art works you will enjoy and more richly understand as you view them in dramatic *full-size* and glowing *original colors!*

Accept this $1 demonstration offer

If you mail the attached card, you will receive a unique Album containing 24 color slides of priceless works by Renoir, Monet, Degas and other important Impressionist painters. These exquisite slides are produced to exacting specifications in Florence on 35mm. Kodak film. They come in individual 2″ x 2″ mounts and can be used in any home slide projector. (If you do not own a projector, see special offer below.)

The color slides, however, are only one part of this elaborate Album. Bound into the center is an illustrated 48-page book by Albert Chatelet, Director of France's renowned Lille Museum—with a guide to each of the slide pictures and the complete story of the Impressionist era.

If this demonstration pleases you, you will be offered a new slide-and-commentary Album approximately every six weeks. Among the Albums in this series are *Italian Renaissance Painting*, the *Origins of Modern Art—Romanticism* and *Realism*—and other Albums covering every major art period.

The cost is surprisingly low

The least you might expect to pay for 24 color slides of comparable quality—if the subjects were available from museums and private collections—is $18-$24. Yet, as a subscriber to this new Program, you pay only $12.95 for each complete Album—including slides and illustrated guide book. You may take as few or as many as you like, and may stop whenever you wish.

Your Introductory Album only $1
No obligation to continue

Send no money now. We will ship your introductory Color Slide Album of Impressionist Paintings and enroll you as a trial subscriber. If you are thrilled with the demonstration, send only $1, plus a few cents for tax and shipping. Thereafter, you will be offered a new Album approximately every six weeks at the subscribers' low price of only $12.95. If not delighted, however, you may return the Introductory Album within 15 days, and owe nothing. Mail the attached card today.

Here's what you get for only $1

Luncheon on the Grass—Edouard Manet
Women in the Garden—Claude Monet
The Green Parasol—Berthe Morisot
Nude in the Sun—Auguste Renoir
The Seated Bather—Auguste Renoir
The Pedicure—Edgar Degas
The Saint-Lazare Railroad Station—Claude Monet
The House of Pere Lacroix—Paul Cezanne
The Saint-Martin Canal—Alfred Sisley
At the Races—Edgar Degas
The Red Roofs—Camille Pissarro
The Loge—Mary Cassatt
. . . plus 12 other Impressionist classics!

The Color Slide Program of Art Enjoyment

If you do not have a projector: here's an unusual opportunity. We can offer you the sensational Kodak Carousel BC 604 Projector supplied with a 500-watt lamp, and 4-inch f/2.8 Ektanar lens. This top-rated projector shows all regular 2″ x 2″ 35mm. slides up to 1/16-inch thick. It has dependable gravity feed and you have instant access to any one of its 140-slide capacity at any time. List price $129.50—subscribers' price only $94.50, plus tax. Remittance must accompany trial subscription form. If you are not completely delighted with the projector, you may return it in 15 days for a full refund.

If Reply Card has been removed, write McGraw-Hill Color Slide Program of Art Enjoyment, Dept. BC-110, 1221 Ave. of the Americas, N.Y., N.Y. 10020.

The Sun shines bright

A solar energy progress report from the designers of America's first solar thermal electric generating plant.

Sometime in the early 1980s, somewhere in or near Barstow, California, someone will flip on a light switch, a hair dryer, or a TV set—and make history—by drawing on commercially-available solar-generated electricity, produced by Solar One.

Solar One: What it is, and what it's not.

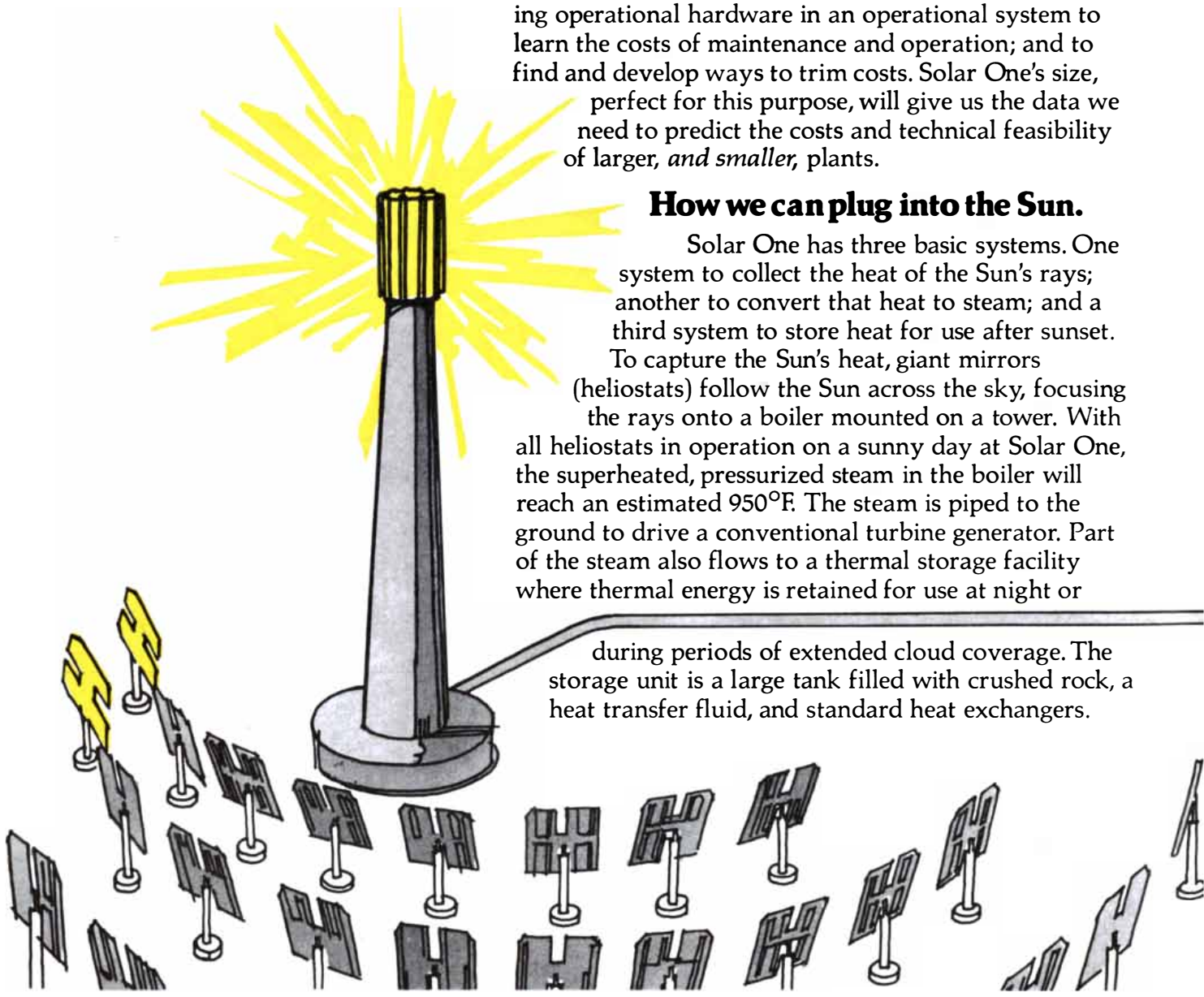
Designed by McDonnell Douglas and being built with our help, Solar One will provide 10 megawatts of power to the Barstow area electric grid. But that's not by any means the purpose of this plant.

Solar One is a pilot plant. A means for demonstrating operational hardware in an operational system to learn the costs of maintenance and operation; and to find and develop ways to trim costs. Solar One's size, perfect for this purpose, will give us the data we need to predict the costs and technical feasibility of larger, *and smaller,* plants.

How we can plug into the Sun.

Solar One has three basic systems. One system to collect the heat of the Sun's rays; another to convert that heat to steam; and a third system to store heat for use after sunset.

To capture the Sun's heat, giant mirrors (heliostats) follow the Sun across the sky, focusing the rays onto a boiler mounted on a tower. With all heliostats in operation on a sunny day at Solar One, the superheated, pressurized steam in the boiler will reach an estimated 950°F. The steam is piped to the ground to drive a conventional turbine generator. Part of the steam also flows to a thermal storage facility where thermal energy is retained for use at night or during periods of extended cloud coverage. The storage unit is a large tank filled with crushed rock, a heat transfer fluid, and standard heat exchangers.

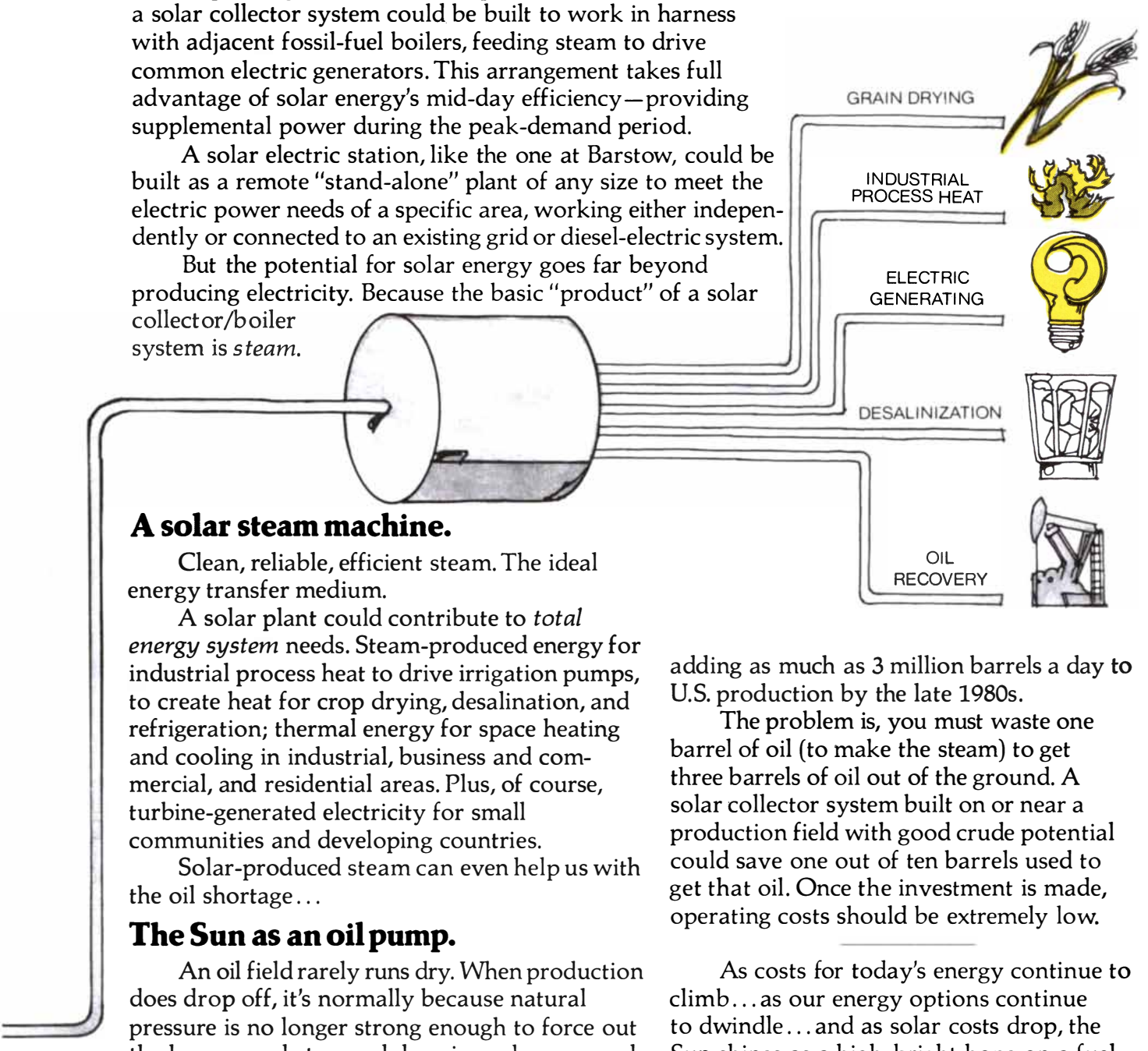


on a fuel-hungry world

Depending on the size of the plant and the needs to be met, a solar collector system could be built to work in harness with adjacent fossil-fuel boilers, feeding steam to drive common electric generators. This arrangement takes full advantage of solar energy's mid-day efficiency—providing supplemental power during the peak-demand period.

A solar electric station, like the one at Barstow, could be built as a remote "stand-alone" plant of any size to meet the electric power needs of a specific area, working either independently or connected to an existing grid or diesel-electric system.

But the potential for solar energy goes far beyond producing electricity. Because the basic "product" of a solar collector/boiler system is *steam*.



A solar steam machine.

Clean, reliable, efficient steam. The ideal energy transfer medium.

A solar plant could contribute to *total energy system* needs. Steam-produced energy for industrial process heat to drive irrigation pumps, to create heat for crop drying, desalination, and refrigeration; thermal energy for space heating and cooling in industrial, business and commercial, and residential areas. Plus, of course, turbine-generated electricity for small communities and developing countries.

Solar-produced steam can even help us with the oil shortage...

The Sun as an oil pump.

An oil field rarely runs dry. When production does drop off, it's normally because natural pressure is no longer strong enough to force out the heavy crude trapped deep in rock pores and fissures. When that happens, external force, usually steam, is forced into the well. The heat makes the oil less viscous, and allows the oil to flow toward the production pumps.

U.S. petroleum engineers estimate that up to 76 billion barrels of crude could be recovered, adding as much as 3 million barrels a day to U.S. production by the late 1980s.

The problem is, you must waste one barrel of oil (to make the steam) to get three barrels of oil out of the ground. A solar collector system built on or near a production field with good crude potential could save one out of ten barrels used to get that oil. Once the investment is made, operating costs should be extremely low.

As costs for today's energy continue to climb...as our energy options continue to dwindle...and as solar costs drop, the Sun shines as a high, bright hope on a fuel-hungry world.

We want you to know we're working on it. And that we would welcome your interest. For more information, write: McDonnell Douglas, Box 14526, St. Louis, Missouri 63178.

Solar Power

MCDONNELL DOUGLAS

BOOKS

Bumblebee economics, new maps and mapping, the night sky in 1980 and Nigerian archaeology

by Philip Morrison

BUMBLEBEE ECONOMICS, by Bernd Heinrich. Harvard University Press ($17.50). It might be anywhere in the temperate U.S. where morning sunlight falls full on some moist meadow. A shadowed woodland is not far off, and nearby too may be a quiet pond, where beavers silently glide. In the spring a flycatcher calls; later on stands of high-bush blueberry and jewelweed come to decorate the spot. Insistent little furry aerial forms in orange and black dart among the blossoms. The simple sense of leisure is our own sunny reverie; here we are watching the work of a hard-pressed economy. Bumblebees (there are 50 species over the U.S. and Canada, all neatly painted in color in an identification key useful in the field) have been inlaid within the complex of tundra life for tens of millions of years by the deftest of invisible hands. Found now in all kinds of open areas—field, roadside and mountaintop—bumblebees are best adapted to the flora of the slowly changing bogs that mark the cool woodlands, relicts of the ancient tundra.

The free energy that drives this urgent economy is held in tiny droplets of a few milligrams of nectar within a myriad of blossoms dotted over the landscape. The bumblebee must visit these, flying in cool air effectively enough for the foraging investment to return an energy profit, with allowance for the indispensable moiety of protein, fats and the like that the flower pollen offers for forming capital: new bees. To this end bees of the genus *Bombus* have evolved a powerful thermoregulatory system; their furry muscle engine remains above 30 degrees Celsius, running even in a snowstorm, whereas almost all other bees are not able to forage at air temperatures below 16 degrees. Engine losses, kept under elaborate circulatory and nervous control, warm the engine itself and the nest as well.

The bumblebee is a flying machine with a power of about a fiftieth of a watt. On a cold morning it develops as much power again in heat; in midsummer the heat is only the minimal loss. Foraging is usually done at full speed, blossom to blossom, engines always revved up for ready flight, for time is—one almost says—money. Time is heat loss in a short season, and profit is in fast foraging. Hovering costs too much to do often, and walking about on big flowering panicles of goldenrod is a strategy most commonly followed by the drones, which have no obligations to the babes at home. Large loads and long flight paths cost little compared with the cost of lost foraging time.

Bumblebees dwell in colonies of a few hundred, often with many queens as well as many workers, the village demographics changing in response to food supply, established by feedback loops as yet unknown. One queen early in the spring starts it all, typically in some abandoned field-mouse nest, warm and snug in the ground, lined by its original builder with delicate fibrous insulation. In contrast to this nectar pioneer the honeybee swarm, a creature originally of the Old World Tropics, is a "big corporation that goes after the big markets." Honeybees recruit entire flights of foragers to a bonanza, using a dance language; bumblebees are individualist foragers, without communications. The honeybee cities have a huge communal storage for a population of tens of thousands. Honeybee life is plainly adapted to tropical windfalls, some large banana inflorescence loaded with pounds of nectar over a few square feet. The bumblebee is matched to a Temperate Zone scatter of open flowers, species following species during an upland summer. ("A bumblebee colony has a more individualistic cottage-industry approach. It thrives by living hand to mouth, exploiting small, scattered energy sources that in most cases can be taken up by single workers operating individually.") Human beings have domesticated the honeybee to orchard and cropland; the bumblebee remains wild because it thrives on fields where no plow has upset the best-laid plans of the field mouse. The fields of red clover in crop rotation are nonetheless fertilized by bumblebees, a princely gift to the farmer.

The data in this book sample the state of the science, not merely the work of Professor Heinrich. Two of his own recent results do, however, go far to strengthen his thesis. They were secured by novel techniques. An electronic sys-

tem developed in his Berkeley group was able to serve as a time clock for bumblebee workers, identifying single bees and recording their times in and out, with their overall weight. Tiny tags weighing one milligram held microwave resonant circuits, each of which rang to a particular frequency from a wide-band transmitter. A scanning receiver recorded the pip and thus the worker's card number, a form of micro-Taylorism. In another facet of this diverse work Heinrich found that the free-flying bumblebee (six minutes aloft in a temperature-controlled room before capture and thermal probe) showed a temperature change of only some five degrees C. when the air temperature changed 25 degrees. This demonstrated thermostability had been missed in earlier work relying on tethered bees as test animals; the tethered bee does not need to work all out to stay at speed; it has no direct perception of its thoracic temperature, and it is cooled by the ambient air.

What about the bumblebee's coevolutionary partner, the flower, for which the economics of the bee is a system of mating? The book devotes a couple of chapters to such questions, as various as the species are numerous, for example the mutuality of design of the corolla of the desert willow and the daily behavior of the bee. The entire study is not only a model of the adaptive systems within the ecological world, a model as rich in questions as it is in answers, but also an example of good popular exposition. Jargon is at a minimum, sharp metaphor and simple estimates form integral parts of the argument; not even the clear quantitative graphs are necessary for understanding, although they and the rest of the illustrations do enhance it. There is a Maine ambience in this displaced New Englander's writing; every chapter bears an epigraph, often one from a New England poet. The book is itself a concentrate of goodness. Not least is the indirect light it casts on human beings, that single species so swift to change and so varied in social structure that its diverse patterns of behavior worldwide can hardly be the work of the same invisible slow forces that have exquisitely sorted the genes of the bumblebees throughout the Cenozoic.

MAPS AND AIR PHOTOGRAPHS: IMAGES OF THE EARTH, by G. C. Dickinson. Second edition. John Wiley & Sons, Inc. ($39.95). MAPS FOR AMERICA: CARTOGRAPHIC PRODUCTS OF THE U.S. GEOLOGICAL SURVEY AND OTHERS, by Morris M. Thompson. U.S. Geological Survey, U.S. Government Printing Office ($11). The first U.S. mapping agency, the Survey of the Coast, was established by the signature of President Jefferson, and today it remains authoritative and vigorous offshore as the National Ocean Survey. The famous quadrangles that represent the land surface

of the country, however, were not drawn until 1882. Four ambitious Federal surveying expeditions followed the westering after the Civil War, laid out the rail routes, found canyon and ruin, pressed to the shores of California. After years of politicking a single consolidated and permanent organization, the U.S. Geological Survey, was formed by law in the spring of 1879. Its first director, Clarence King, and his dynamic successor John Wesley Powell seized for that agency not only extension to the older lands east of the 100th meridian but also full responsibility for the topographic mapping of the national surface, a task not likely to end while the nation lives.

The uniform mosaic of the most modern quadrangles, at a scale of about two inches to the mile (1 : 24,000), "probably represents the world's largest area consistently and systematically mapped" at such a scale. There are blanks still, and the chief present campaign of the U.S.G.S. is to bring the entire country (save great empty Alaska) onto the map, in the format familiar and delectable to every wanderer since the 1950's. The work goes on apace, about 100,000 new square miles each year, and the National Mapping Program inaugurated in 1975 promises to be completed by the mid-1980's. The Sierras and the Cascades, the forests of Maine and Michigan, and a good deal of flatland east and west remain to be done at the large scale.

These two books provide a wide look at maps today. Dr. Dickinson is an English geographer. His book is an introductory text, with international interests. Mr. Thompson, a veteran U.S.G.S. cartographer, has prepared an inexpensive, useful and colorful volume celebrating the centennial of the survey, a book aimed at users and buyers of maps in the U.S. Centered on the work of the survey, as its anniversary warrants, the book is by no means narrow in coverage. On the contrary, it accepts with intelligence and concern the function of helping users whose interest is not bureaucratic boundaries but data on the land. One chapter lists all the Federal agencies that make maps, with the nature of their diverse products, and it effectively reminds the reader of state, local and commercial sources of maps as well. Included is a tabulation of sources by the product sought and a careful list of agency addresses, bound to be a valuable reference.

Maps are old, of course; the first atlas was perhaps that of Ptolemy. It was the Renaissance, however, that nourished their flowering. Seamen had coastal maps early on, and landholders had long used surveyors well. The overall representation of their lands fell to the forming national states. The military need for maps is clear, and it still plays a major role. Nowadays even absolute geodetic accuracy is a military goal, in a time of far-flying strategic weapons. The careful discrimination of types of bridges in some European national mapping systems is logistic information; the woodland markings on the maps of the U.S.G.S. staff take into account whether or not the vegetation provides cover for troops. Perhaps the fullest coverage of all national maps is that of the British Ordnance Survey, which has been responsible for topographic maps of the realm since the days of that revolution over in France, when the Master General of His Majesty's Ordnance undertook to image the kingdom.

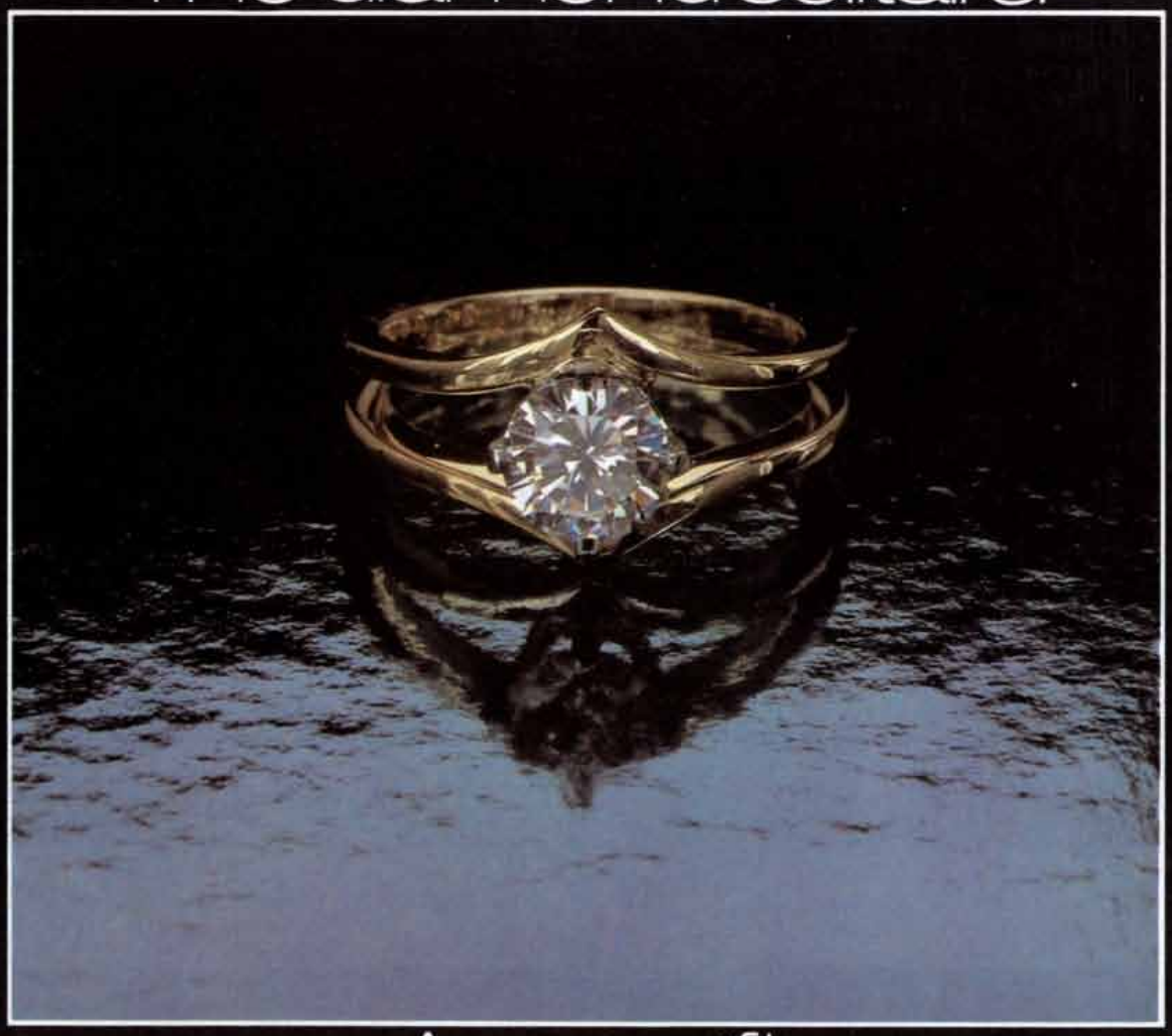
A fine period photograph in the U.S.G.S. book provides the canonical image of the mapmaker. One sees a man jaunty in slouch hat and high boots bending over his tripod, set carefully on shingly rock above the timberline, with the snowy peaks of the Rockies behind him. The apposite epoch is Theodore Roosevelt's time. Nowadays, of course, it is the camera flying high in the air, or even in orbit, that supplies most of the mapmaker's data. Almost every quadrangle bears a date of the primary photography. Indeed, the newest cartographic product is the orthophotomap, a direct use of the aerial photograph as a published map, with only its distortion errors removed. Photographs in general, even neglecting the small errors of optical systems, are marked by two types of intrinsic distortion, one type due to the perspective effects of a camera angle that is not perfectly vertical and the other due to changes in the distance between the ground and the lens (the latter being the consequence less of changes in the altitude of flight than of the changing elevation of the ground itself).

A revealing pair of figures in *Maps and Air Photographs* shows on one page a normal vertical aerial photograph inscribed with a rectangular grid and on the opposite page the same photograph and the same grid corrected for the height and tilt displacements of careful flight. The grid lines wander and bend like a crumpled piece of fishnet. It was the task of photogrammetry to apply those corrections to the air coverage. A map represents the true plan of the land, projected onto a single horizontal plane within some stipulated limits of tolerance. Between a series of photographs and a map a long process of measurement, reprojection, computation and careful fitting has always intervened. The orthophotomap is the rectified product of the faulty air coverage. What has been done is in effect to scan the film systematically, one tiny rectangular area at a time. Each element is adjusted for the local height of the terrain and for any tilt, by means of the information available in the stereoscopic pair. At first this differential process was carried out visually and manually with ingenious optical and mechanical analogue instruments; these days the entire process is digital and computerized. The only remaining human skills are surveying on the ground to fix the all-important control points for accuracy and a similar process of setting things up on the consoles; the rest is done by machine, from survey to map.

The most ambitious program of this kind now under way is a map of Australia in 3,000 sheets. A number of U.S.G.S. orthophotomaps, at various levels of detail, are shown in the Thompson book. They seem excellent for such flat and yet intricate places as the marshy shores of Prudhoe Bay in Alaska, but the contour lines and the photographic image interfere severely in places of strong relief. The symbolic world of the drafted map still has a great deal to offer. There are 100 million bits in a typical quadrangle; digital replacement, even simple storage of the contours on magnetic tape, is immature. The terrain contours of the U.S. have been digitized rather roughly, with elevations compiled from contours at 100- or 200-foot intervals, on maps at a scale of 1 : 250,000. Those digitized terrain tapes are available to the public; airplane and missile guidance require more data. The future will see more detail: automatic line following and interactive editing of the data are here now, on an experimental basis, and many specialized digital-data map systems have been set up, for example the network of public-land boundaries over 400 quadrangles in the California desert. In the 1980's elevations and boundaries will be digitized by automatic systems over wide areas.

Cartography involves a kind of social self-reference. More than in most technological fields, progress is governed by what has actually been done, by what conventions have been adopted and by what needs have been realized. Both books reflect the point clearly, the British introduction to techniques, utilization and history perhaps less so than the American volume, which is really a user's guide to available information on maps rather than an explanation in detail of how they are made. Remote sensing from airplane and satellite presages a new cartography, richer in every way. The evidence is nonetheless clear that the visual presentation of data will continue to require the subtle intervention of mind and hand in the long chain of mapmaking. Aerial photograph and computer program are not enough. Meanwhile it is a pleasure to note how accurately the form of the world we now see swiftly from far above was caught in the net cast by the slow, surface-bound mappers of the past. The Bureau of the Census published a wonderful map of the distribution of the population of the U.S. presenting the head count of 1970 in white dots on a black ground, simulating "a night view of the United States from a satellite." The actual satellite photograph is now familiar to many; the

The diamond solitaire.



A rare gift.

A one carat diamond. Set simply and elegantly.
To sparkle on its own. Of lasting value,
because diamonds of about one carat and up are rare.
(The one shown here is worth about $4,950.)
The diamond solitaire.
It is the gift that makes a rare and beautiful
moment last a lifetime.

A diamond is forever.

The one carat diamond shown is enlarged for detail.
Prices may change substantially due to differences in diamond quality and market conditions. De Beers.

distribution of the population is close to the brightness of the lights at night but is not identical with it. The U.S.G.S. book reproduces the eastern part of that census map. It is a paradigm of inference.

THE VIEW FROM THE EARTH: A SIMPLE GUIDE TO THE NIGHT SKY AND THE EVENTS THAT WILL HAPPEN THERE, by Guy Ottewell. Department of Physics, Furman University, Greenville, S.C. 29613 ($4). THE AMATEUR ASTRONOMER, by Antonín Rükl. Consultant editor, John Gribbin. Octopus Books Limited, available from Mayflower Books, Inc., 575 Lexington Avenue, New York, N.Y. 10022 ($9.95). THE SOLAR SYSTEM, by John A. Wood. Prentice-Hall, Inc. ($10.95). PROTOSTARS & PLANETS: STUDIES OF STAR FORMATION AND OF THE ORIGIN OF THE SOLAR SYSTEM, edited by Tom Gehrels, with the assistance of Mildred Shapley Matthews. With 51 collaborating authors. The University of Arizona Press ($17.50). In these days when our earth starts uphill in its annual orbital round (its closest approach to the sun is near January 2) it is a custom here to comment on books that look to the skies within the sun's system. This year four books will be assessed, a progression in depth from the visible and introductory to a systematic review at the frontiers of knowledge.

Guy Ottewell, an artist-astronomer, has made a book of sky maps, an evening sky map for each month, done by his own hand, white on black, with a level of constellation detail and textual weight about right for serious-minded novices and a "lower limit" for children in the upper elementary grades. Each map bears an extended legend that comments on the month's stars (January, say, brings Orion, the easiest constellation), and on the opposite page there is an account of the month's transient events: planetary motions, meteors, eclipses and even the estimated arrival times for a few planetary probes. A calendar of the month with the moon phases is also presented, with candid remarks on stargazing opportunities. A few pages at the end explain the views in their true spatial context, in the author's imaginative pedagogical style.

The sky is drawn as from latitude 40 degrees north, good enough anywhere in the "thickly inhabited north-temperate zone of the earth." If you call it home there, the book is a useful guide, neither the simplest nor the cheapest to be had but handsome, easy enough and genuinely helpful. Look up at the eastern predawn sky in January and February or watch the sunset fade from March through June and you will see an unusual stately dance of the three superior planets (Mars, Jupiter and Saturn), to whose to-and-fro close passages the bright star Regulus adds a position of rest.

A notable bargain, *The Amateur Astronomer* is a product of the expert educational publishers of Czechoslovakia. A well-bound book of nearly 200 large pages with 75 pages of colorful paintings and plenty of useful data in color graphics, it comes to us after translation (by way of London) through a new U.S. publisher. Uncoated paper and an appropriate graphic design present the pictorial material economically and well. There are no photographs, but their place is sturdily held by the watercolors and airbrush work of the artist, who manages to represent convincingly the cratered surface of the moon, the look of the Milky Way at night and the sweeping tail of Comet Ikeya-Seki.

In addition to such paintings-after-photographs the book offers many useful astronomical diagrams in color and a suite of carefully labeled up-to-date topographic maps of the inner planets. The work is an atlas and a handbook, from the compelling page of colored graphs that display the hours of day, night and twilight at four latitudes, an artist's view of our galaxy edge on and face on, a page of spectral distributions to show how temperature determines the colors of stars, a few nebulas seen as in a telescope, to a general biography of a comet: the true orbit of Halley's comet shown to scale with its speeds and latest schedule. Pavel Büchler has carried out a useful graphic design, although his typographic symbols are rather oversize for what is needed. The book is not difficult to follow visually, even though it is rich in detail; the text is well translated by Olga Kuthanová.

The last 50 pages are a valuable bonus. There is a simplified map of the night sky, stars white on blue, which is then elaborated over seven more full-page maps, with positions entered for the usual sky sights of the amateur at his telescope. Their numerical coordinates, however, are not given. The opposite pages list the sights with some physical details. Then there is an unusual graphical display, easily usable once the reader has a little understanding of the geometry of the skies, that offers eight pages of curves marking the positions of the five bright planets over the next 20 years. The accuracy is good enough for locating the object among the stars and in relation to the sun. The sinuous path of Mercury around the position of the sun (with Mercury coded as *H,* for Hermes) suggests the serpent-entwined staff, the caduceus, certainly by no accident. Tables announce the phases of the moon, the eclipses of the sun and the moon and the close approaches of the planets, all through A.D. 2000. Although the book is not a primer, it is an outstanding buy as a reference work and an enjoyable visual volume for anyone beyond the tyro stage, and of course for any library.

In *The Solar System* John Wood has written a little text for college students, but the fare is rich enough for any general scientific reader. In it he acts as intrepid guide to a zoo of sciences, those whose quantitative findings shape and constrain our understanding of the solar system, going much beyond our rough perceptions of the sky. The objects in the solar system fly about by no means haphazardly; they have solid surfaces, and some have a gaseous atmosphere and a high-pressure interior; samples arrive here, by Newtonian chance or Apollonian virtuosity, to be scrutinized by analysts who have thrown the book of instrumentation at their precious bits of sun stuff or moon stuff; the sun and its kindred stars demand optical attention. From all these facts a theory of the solar system will mature; the fruit is still small, green and a bit acid.

A few of the fascinating graphs and charts in the book will suggest its level and value: the Titius-Bode law is expressed as a pretty good straight-line fit between the outward order of the planets and the logarithm of the mean solar distance, with each orbit about 75 percent larger than the preceding one. Here is seen less of origin than of persistence; closer spacings simply do not endure against mutual gravitational perturbation. A wonderful graph plots the orbits of many objects according to inclination to the mean plane and to shape, measured by degree of circularity. There are three clear groupings: most planets hug the origin and orbit circles close to the plane, whereas the many near-parabolic comets are spread out along a well-defined line. The comets fall in and fly out again at all directions to the plane, but their orbits are the opposite of circular. The asteroids of the asteroid belt cluster tightly around an implicit orbit, and a scatter of mixed near-earth asteroids and periodic comets reveals a system in the process of slow change.

Geologic maps of Mercury, the earth, the moon and Mars are compared, and a range of models for the internal structure of Mars, six on a page, are displayed. Another cleverly managed page exhibits to scale the Jovian planets and their satellites. Modern chemical and mineralogical work on the samples from space is well described by an author whose special expertise lies here, and the reader is shown a good deal of order in a confusing landscape. The noble antiquity of the volatile-rich carbonaceous chondrites, sun stuff as soft and organic as a sooty cheese, is made plain, and so is the reworked nature of the lunar samples, pieces not of astronomy but of planetary geology. A too brief elementary review of star physics is followed by a modest and helpful summary of the features of origin theory, all variants of the evolution of a dusty nebula, co-evolving with the ancient sun. Dr. Wood does not oversell the arguments: there is a lot to learn.

Protostars & Planets makes pretty clear how far there is to go. Its three

SETTLE FOR MORE

The Volkswagen Dasher isn't a down-sized something else. It wasn't wrenched into shape overnight.

It's exactly the size we had in mind, and so we've had plenty of time to make it more of a car instead of less.

For example, you have more than one choice: the 2-door hatchback, the 4-door hatchback or what we feel is the world's most fascinating station wagon.

Whichever you choose, you also get to choose either the fuel-injected gasoline engine or the optional diesel engine.

The Dasher Diesels deliver an EPA estimated 36 mpg and an extraordinary 49 mpg highway estimate. And the gasoline engine is no slouch, either, with 23 estimated mpg, 35 mpg highway estimate. Use "estimated mpg" for comparisons. Mpg varies with speed, trip length, and weather. Actual highway mpg will probably be less.

Don't go away. There's more performance than you thought: a gasoline Dasher Sedan will propel you from 0 to 50 mph in only 8.1 seconds*, quicker than a few "sportscars."

More room than you thought: the Dasher wagon holds 31 cubic feet of anything. Need still more? In 30 seconds, the rear seats fold down and you have 51.6 cubic feet.

More luxury than you thought: the carpeting in every Dasher is really plush and really everywhere. The AM/FM stereo radio is standard. So is the electric rear window defogger. The steel-belted radials. The quartz clock.

There's nothing cheap about the Dasher, including the price. But the fit and finish are equal to just about any car made.

The only thing you give up are some old-fashioned ideas about what a car ought to be.

*Dasher Diesel Sedan 0-50 mph in 13.0 sec.

VOLKSWAGEN DOES IT AGAIN

dozen papers arise from a determination to initiate a new discipline: in the absence of any other solar system—yet—there is a plain need to bring together the deep-sky astronomers, who observe and calculate on how stars form, and the solar-system astronomers, who have enormous power over detail but who must study a big show that closed four billion years ago, with only minor scenes playing here and there ever since. The dusty clouds in space have a complex chemistry and a poorly understood self-gravitating dynamics. Exposure to a number of outside inducements toward collapse seems likely; observations show both new stars that have flared up in a cloud of dust and supernovas that seem to have triggered the formation of a ring of stars around them.

All of this is examined and modeled in contributions by an international coterie of diverse experts. The degree and consequence of uniformity, subdivision and mixing in the ancient solar nebula are the chief local themes, as the role of outside influence is the main remote theme, in this double look at two joint mysteries: how distant stars form now and how nearby planets grew then. At the level of the research student this big volume will be a starting point for many. Perhaps the most important novelty since the smooth and thoughtful compilation of this tome in 1978 is the growing evidence in our galaxy for a decisive role of external influences in forming new stars, as the isotopic anomalies in the moon, the earth and meteorites have offered local evidence for the fathering of our sun on some ancient supernova. This is no book for beginners, but it is a fine one for those who would embark on stormy seas of inference.

In all these books only one drawing in Wood offers information on such a homely, if immaterial, point as the date of the earth's annual perihelion passage.

NIGERIA: ITS ARCHAEOLOGY AND EARLY HISTORY, by Thurstan Shaw. With 147 illustrations. Thames and Hudson Ltd. Distributed by W. W. Norton & Co., Inc. ($16.95). The scientific spade has borne eloquent witness to the human past for a century and a half or more in some lands. In Nigeria, the most populous of all the nations of Black Africa, with a territory and a population as large as those of France and Spain together, archaeology hardly began before World War II. Still young, it has the buoyancy of youth; recent finds are abundant (even in this brief introductory volume some objects are shown that have not yet been formally published) and new excavators are enthusiastically at work, Nigerians trained at and responsible to Nigerian universities and museums. In a new and diverse nation one impetus to the study of a buried past is clear. It is expressed neatly in the epigraph of this exciting book, ascribed to a well-known pop singer and social critic of Lagos, Fela Anikulapo-Kuti: "You will never know where you are going unless you know where you are coming from."

Our guide here to where Nigerians in all their variety come from is no newcomer; he is the dean of Nigerian archaeologists, who has to his credit digs and publications in West Africa for nearly 40 years. The reflective survey he offers not only brings material fresh to most readers but also is marked by a catholic breadth of interest, from climatology to ethnobotany and linguistics, which give his book a value beyond the topic directly at hand.

The spade can find order, not make it; the book of sediments is hard to read where it is jumbled and awry. Across Africa in the Rift Valley there are famous sites, such as Olduvai Gorge, where the pages of the past lie one on another in neat order, cut into and exposed to our benefit. There is no Olduvai in Nigeria; the closest to that treasury is the cool grassy Jos Plateau in central Nigeria. There the granite outcrops have weathered and broken down to fill the complexly cut streams with alluvial gravels, whose content of heavy lumps and grains of tin oxide have drawn prospectors and miners on a scale large and small for 50 years. Their burrowings have found many works of human hands in the gravels, but they are rarely ordered and datable. More often they have been washed from the sides of the valley into the loosened gravels of the fill below. Nowhere else in Nigeria has the subsurface been so widely sampled.

In the Jos region and downstream from it hand and mind can be traced for a very long time. Hundreds of scattered flint hand axes lie there, hard to date or even to sort but plainly showing occupancy by our ancestor *Homo erectus* perhaps as much as a couple of hundred thousand years ago. It can only be hoped that an assemblage of tools will be found in place, marking some camp of the elephant hunters, like the tools encountered on the quieter shores of the great lakes of East Africa. In the fans of loose gravel at the base of flowing waterfalls the tin miners have found smaller flakes made as spearpoints by people of the Middle Stone Age, our own species *Homo sapiens;* at a guess the flakes are some 30,000 or 40,000 years old. In the same area a rock-shelter has yielded ground stone axes and pottery shards, 6,000 or 8,000 years old, datable by accompanying samples of carbon. Still more is owed to the prospectors eager for tin and tantalum: iron objects, slag and finally a dozen old iron furnaces, located by the magnetometer. There, at Taruga in about 300 B.C., there was a "fully evolved iron technology." In a nonalluvial site, Taruga begins to tell a coherent cultural story.

These later people are known to us by more than their technology. Their portrait heads and now figurines of all kinds in terra-cotta have also been by-products of the search for tin. The Nok culture, as it is called, still awaits full publication, although the first few extraordinary heads, found a generation ago, are known to everyone who has seen a text on African art or archaeology. The look of these life-size heads can be matched unmistakably to many a Nigerian head today. Agriculturists with their own novel domesticates (such as sorghum), artists, potters, smelters of iron, these people are not so unlike their contemporaries of western Europe.

What is striking are their physical characteristics, made plain by their own hands in portraiture. They were negroid, as are almost all Nigerians today; such a gene pool, of course with a statistical spread of traits, can be documented in the Nigerian subsoil. In one rock-shelter well south of the Jos Plateau a Middle Stone Age skeleton has been dated to about 9000 B.C., "the earliest skeleton so far discovered" that shows physical characteristics prototypical of present African populations. The modest present degree of human polytypy seems to have developed only rather late in the history of our species. Ethiopian skulls of the *sapiens* type indicate that such "differentiation had not taken place before 40,000 B.C." We are a single, complex, cunning species, north and south, our common ancestry strongly marked on us all, probably emigrants out of East Africa some 50 millenniums back.

Professor Shaw of course includes the protohistoric archaeology of Nigeria: the art, both in terra-cotta and in copper and its alloys, that has astonished viewers for a long time. Three remarkable styles are detailed: the dig Professor Shaw himself conducted in Igbo-Ukwo, whose cast objects of "strange rococo, almost Fabergé-like virtuosity" in a local burial seem Roman in their lightness and mimicry of ropework; the copper heads of Athenian naturalism found and still being found in Ile-Ife, the holy city of the Yorubas, and the famous brasses of Benin, brought back by Portuguese travelers as impressed by that great city as learned Arab visitors had been before them. These works, all manifestly indigenous although made from traded copper, probably under the influence of metallurgical techniques from far away, span the years from the fifth to the 17th centuries A.D. The city walls, the shard pavements, the chief's insignia, the clay leopard bearing a human leg in its mouth—all abundantly reward the reader in interest and beauty. Here one can turn the first illuminated leaves of a work in progress: the unfolding story of where the Nigerians come from. In this logical land, where the age of costly bronze appears much later than the age of common iron, the reader finds much to admire and to learn.

Welcome to the expanding universe . . .

The known universe is expanding at the rate of 40 trillion cubic light-years per second. And, not only is the universe *physically* expanding, but our understanding of it is undergoing an unprecedented explosion as well: we have gained more *new* knowledge of the universe in the last two decades than in *all prior human history!*

The pace of discovery is now so rapid that most – if not all – of the astronomy books in your local library are badly out-of-date – *even if they are only a few years old!*

STAR & SKY is a magazine that pays particular attention to the fast-breaking news and keeps you on top of it like an astronomy newspaper . . . in it you can find consistently accurate, detailed reports of what's going on today (and *tomorrow*) in space exploration, astronomical research and discovery.

STAR & SKY has reporters at mission control for the latest planetary encounters so we can bring you the detail and full color illustrations that the general science magazines don't seem to have room for. We reported on what the Voyager and Pioneer spacecraft actually found at Jupiter and Saturn and in later issues augmented it with late-breaking data emerging from further analyses.

We cover the scientific meetings at the world's great observatories and universities to bring you details of the most recent discoveries about quasars, galaxies, black holes and such mind-stretching concepts as the origin and destiny of the universe. In addition, our monthly report on manned space activities by James Oberg at NASA's Johnson Space Center ensures you that you won't miss anything.

UNLIKE ANY OTHER MAGAZINE IN THE FIELD we often publish *Special Issues* – entire issues devoted to in-depth treatment of *single topics*. Our December 1979 *Special Issue* is an all-inclusive examination of those cosmic enigmas, BLACK HOLES.

Six feature articles survey the most up-to-the-minute black hole facts and notions, written by authors who are among the best-known names in the field:

- **SPACE, TIME AND BLACK HOLES *by William J. Kaufmann:*** *By severely warping space and time, black holes can create a totally alien realm that may lead beyond the universe.*
- **THE ULTIMATE BLACK HOLE *by Isaac Asimov:*** *What would happen if all the matter in the universe were compressed into one enormous black hole?*
- **THE MAW OF THE ABYSS *by Wallace Tucker:*** *Colossal black holes with masses greater than a billion suns may lurk in the hearts of galaxies and quasars.*
- **BLACK AND WHITE HOLES: MINIS, MAXIS AND . . . WORMS! *by Barry Parker:*** *Are black holes gateways to other universes?*
- **BLACK HOLES OF THE MIND *by Philip Morrison:*** *A leading physicist cautions against confident acceptance of black hole reality.*
- **STEPHEN HAWKING, *BLACK HOLE REVOLUTIONARY by Ian Ridpath:*** *Wherever black holes are discussed, the name of Stephen Hawking is sure to emerge. This physicist has created a revolution in our understanding of black holes.*

These ***Special Editions*** give you the most thorough, in-depth and up-to-date accounts available anywhere: the facts, the findings, old and new theories, photos, computer-generated art, schematics and diagrams, upcoming projects and the riddles they hope to solve – all presented to you by the greatest astronomy minds on Earth!

A BALANCE OF VIEWPOINTS AND AN ARRAY OF FACTS *ALL IN ONE VOLUME AVAILABLE NOWHERE ELSE – BOOKS OR MAGAZINES!*

STAR & SKY is a blend of visuals and precise but easy-to-digest text. Jargon is held to a minimum yet we don't skimp on comprehensive coverage.

And art: What does a black hole look like? How does it work? How do the rings of Jupiter, Saturn and Uranus differ? Where photography is impossible, we turn to art. Intelligent renderings–made by astronomer-artists who keep abreast of the newest theories and discoveries – portray visually the probabilities and possibilities we can't see.

STAR & SKY brings you top authors in both professional and amateur astronomy: Isaac Asimov, James Oberg, Richard Hoagland, William J. Kaufmann, Barry Parker, Wallace Tucker, John Gribbin, Patrick Moore, John Sanford, Leo Henzl, Terence Dickinson, Ben Mayer, Sam Bleecker, Trudy E. Bell, George T. Keene, James Mullaney, and many others.

STAR & SKY **covers the entire field of astronomy: IT IS THE COMPLETE REFERENCE FOR EVERYONE INTERESTED IN THE UNIVERSE.**

FOR THE PRICE OF A TYPICAL HARDCOVER BOOK YOU CAN ENJOY 12 DETAILED AND UP-TO-DATE ISSUES OF STAR & SKY . . .

AND as a Special Offer available only to SCIENTIFIC AMERICAN readers, the Special BLACK HOLE Issue (a $2.25 value) is yours – FREE – by subscribing to *STAR & SKY*. If you enclose payment with your order, we will send your BLACK HOLE Issue out immediately. If you prefer to be billed, we will send your BLACK HOLE Issue out when your payment is received.

Act today to reserve your Special BLACK HOLE Issue – a definite collector's item – as supplies are limited.

STAR & SKY Magazine, P.O. Box 5172, Westport, CT 06880

I don't want to miss any reports of Voyager's encounter with Saturn; the first launch of the Space Shuttle; preparations for the return of Halley's Comet; plans for the launch of a 94-inch telescope into Earth orbit; etc., etc. . . . but I especially don't want to miss any future ***Special Issues*** *on important and timely topics in astronomy!*

☐ Yes, send me my ***Special*** **BLACK HOLE *Issue*** right now. I am enclosing my check or money order.

☐ Please bill me. I understand that when my payment is received, ***STAR & SKY*** will send me my ***Special*** **BLACK HOLE *Issue.***

- ☐ 3 years for $33 (36 issues) – a 54% saving over the newsstand price.
- ☐ 2 years for $27 (24 issues) – a 44% savings over the newsstand price.
- ☐ 1 year for $15 (12 issues) – a 38% saving over the newsstand price.

Canada and Mexico add $3 – all other foreign add $6 PER YEAR OF SUBSCRIPTION.

Name_____

Address_____ Apt. #_____

City_____ State_____ Zip_____

ISAI

Established 1845 SCIENTIFIC AMERICAN January 1980 Volume 242 Number 1

The Next Generation of Particle Accelerators

The smallest of objects can be perceived only with the largest of instruments. Penetrating deeper into the structure of matter may require accelerators built with multinational sponsorship

by Robert R. Wilson

For some 50 years the effort to understand the ultimate structure of matter has proceeded almost entirely through a single experimental technique. A particle of matter is brought to high speed and made to strike another particle. From an examination of the debris released in the aftermath of the collision, information is gained about the nature of the particles and about the forces that act between them. To carry out a program of such experiments it is necessary to have a source of energetic particles. Cosmic rays provide a natural source, but the flux of particles is diffuse and is beyond the control of the experimenter. A more practical source is a particle accelerator, the device for increasing the speed of a particle and hence also its energy.

One of the first particle accelerators, built by Ernest O. Lawrence in 1928, was made of laboratory glassware a few inches in diameter. Most of the accelerators in service today are lineal descendants of Lawrence's device, but they have grown enormously in size and complexity. The largest extend over many square kilometers and indeed are so large that the availability of real estate has become a significant consideration in their design. The particle accelerator is no longer an instrument installed in a laboratory; instead the laboratory is assembled around the accelerator. Building such a machine costs hundreds of millions of dollars; operating it requires a staff of about 1,000 people and dozens of digital computers.

A new generation of particle accelerators is now in prospect. The first few are just coming into operation; several more are under construction; others are still being planned, and their characteristics are not yet fixed. For both the physicist and the layman the principal interest inspired by these new machines is in the results of the experiments they will make possible, but the accelerators themselves also merit notice. In the physics of elementary particles the highest available energy represents a frontier marking one of the boundaries of experimentally verifiable knowledge. Several of the new accelerators will be capable of attaining higher energies than any existing machine, and so they will push the frontier into unexplored territory. In order to reach those energies the accelerators will of necessity be larger, more complicated and more expensive than their predecessors.

Largely because of the cost, the construction of an accelerator today requires the resolution not only of technical problems but also of political, economic and managerial ones. Money for scientific research is a scarce resource, and it is imperative that it be used as efficiently as possible. Technical innovations have brought a substantial reduction in the cost per unit energy of accelerating a particle. It is encouraging to note that another means for minimizing the total world expenditure is now emerging: through international cooperation the unnecessary duplication of facilities can be avoided, and projects too large for any one nation can be undertaken by regional groups of nations and perhaps eventually through a worldwide collaboration.

The large investment now being made in instruments for high-energy-physics research can be justified only because the preceding generations of accelerators have already proved their worth. Fifty years ago only two kinds of apparently indivisible particle were recognized: the electron and the proton. The remaining constituent of the atom, the neutron, was discovered in 1932. In subsequent years, through experiments with cosmic rays and with early accelerators, several additional particles were identified. One of the first was the positron, the antiparticle of the electron. Others were the neutrino, a particle without mass or electric charge, and the muon and the pion, which have masses intermediate between those of the electron and the proton.

In the 1950's, when more powerful accelerators began operating, there was an unexpected and in some respects alarming proliferation in the number of known particles. Within a few years the list extended to more than 100, most of them classified as hadrons, or nuclear particles. Among the hadrons were several with the new property of matter called strangeness. Five years ago it was necessary to add another class of hadrons, bearing another whimsically named property, charm. The pace of discovery continues to increase. Particles that apparently signal the existence of two more classes have been observed. These newest classes, which have only begun to be catalogued, are distinguished by properties called truth and beauty or top and bottom.

For a time it seemed that all of these particles might have to be accorded equal status as elementary objects. That possibility was deeply troubling, as it was difficult to reconcile with the conviction that the laws of nature should be

reasonably simple. It was subsequently discovered, however, that all the hadrons could be arranged in logical patterns, some of which have a lovely snowflake form. Moreover, the existence of such patterns could be understood if it was assumed that the hadrons are not elementary but are made up of the more fundamental entities that have been given the name quarks.

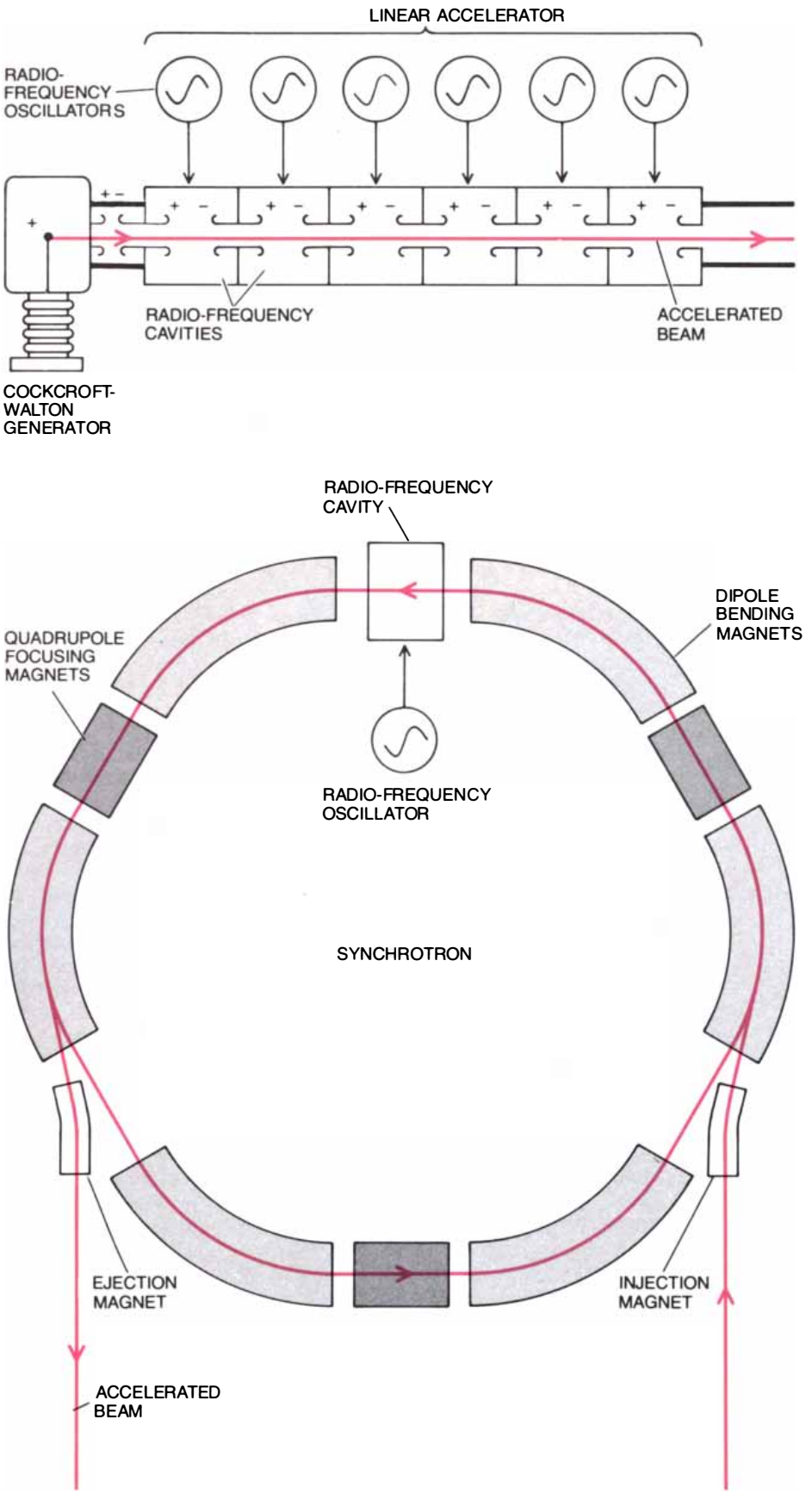
In the view that now prevails among physicists there are just two kinds of elementary particles: leptons and quarks. Among the leptons the most familiar particle is the electron. Also included in that class are the muon and two kinds of neutrino, one associated with the electron and one with the muon. A few years ago a new lepton was discovered and given the designation tau. Presumably the tau also has an associated neutrino, so that there should be six leptons altogether.

There also appear to be six kinds of quarks, labeled up, down, strange, charmed, top and bottom. (As yet there is no experimental evidence for the top quark, but because all the other quarks and leptons come in pairs it is assumed that the bottom quark also has a partner.) No one has observed a quark in isolation, but there are substantial reasons for believing in their existence. Every known hadron (and there are now a few hundred) can be explained as a combination of quarks or of quarks and antiquarks, formed by explicit rules.

Cataloguing the elementary particles constitutes only half the problem of understanding the structure of matter. It is also necessary to understand the forces that hold the particles together and give rise to their motions. Four basic forces are generally recognized. In order of increasing strength they are gravitation, the weak force (which is responsible for radioactive beta decay), electromagnetism and the strong force. Gravitation and the weak force are universal: they act between all kinds of particles. Electromagnetism has a direct influence only on particles that carry an electric charge. The strong force acts only on hadrons or on their constituents, the quarks.

Each force is thought to be transmitted from point to point by the exchange of an intermediary particle. For electromagnetism the intermediary particle is the photon: the quantum of light and of other forms of electromagnetic radiation. Thus the repulsion between two electrons is described as coming about from the exchange of photons, which are emitted by one electron and absorbed by the other. (A similar repulsion would be observed between two children on ice skates throwing a ball back and forth.) The mechanism of interaction is similar for the other forces, except for the identity of the exchanged

TANDEM ACCELERATORS occupy a single tunnel at the Fermi National Accelerator Laboratory (Fermilab) near Chicago. The blue objects in the upper tier are the beam-bending magnets of the original "main ring" proton synchrotron, which began operating in 1972 and has reached a peak energy of 500 billion electron volts (GeV). Lower tier of red magnets is the first segment of a new proton synchrotron, called the Tevatron, that is designed to reach one trillion electron volts (TeV). The Tevatron magnets have superconducting windings. The scale of the accelerators can be judged from the slight curvature of the tunnel, which has a circumference of 6.3 kilometers.



ACCELERATION OF A PARTICLE results from the force applied by an electric field. A static high voltage supplies the field in a Cockcroft-Walton generator, but the maximum energy that can be reached by this method is about a million electron volts (MeV). Higher energies call for many stages of acceleration, in which the electric field is generated by radio-frequency oscillators. In a linear accelerator, or linac, many radio-frequency cavities are arranged in sequence and synchronized so that a particle receives a small push at each stage. In a synchrotron the many radio-frequency stages can be replaced by one cavity, through which the beam passes many times. Bending and focusing magnets confine the particles to a circular path; the magnetic fields must be adjusted continuously, however, as energy of the particles increases.

particle. The quantum of gravitation is the graviton, and the strong force between quarks is transmitted by the particles called gluons. For the weak force the mediating particle is the intermediate vector boson, now also known as the weakon; it comes in three charge states, designated W^+, W^- and Z^0. Whereas the photon, the graviton and presumably the gluons are all massless, the weakon is expected to be very heavy.

One of the most far-reaching theoretical developments of the past several years was a demonstration that the weak force and the electromagnetic force can be unified. Although the two forces are quite dissimilar in their observable characteristics, they can now be understood as manifestations of a single underlying phenomenon. The unification is precisely analogous to the 19th-century realization that electric and magnetic forces are merely different manifestations of charge. Work is now under way toward the construction of a "grand unification" that would include the strong force with the unified weak and electromagnetic forces. The ultimate source of force, however, remains as mysterious today as the source of the Nile was in the 19th century, and the search for that source is just as romantic.

However much has been learned in the past 50 years, it would be misleading to suggest that the present understanding of elementary particles is even approaching finality or completion. The status of the field is tantalizing rather than satisfying; there is no shortage of questions to be answered. A first order of business for the new accelerators will be filling in the blanks in the catalogue of hadrons, particularly those that incorporate top and bottom quarks in their structure. It is also important to find out whether the list of quarks and leptons ends with the six of each that are now known or whether more will be found at higher energies. In a sense six quarks and six leptons are already too many; all of the ordinary matter in the universe could be constructed out of just four elementary particles: the electron, the electron neutrino and the up and down quarks. The existence of the other leptons and quarks, which appear only in high-energy-physics experiments, is a puzzle.

Another puzzle is the failure of all attempts so far to detect a free quark. Various theoretical constructs have been offered, after the fact, to explain why quarks should be permanently confined to hadrons. The possibility remains, however, that a quark can be knocked loose from a hadron if enough energy is supplied. Future experimental programs are therefore certain to include quark searches.

One of the trophies that will be hunted with the new accelerators is the weakon, the transmitter of the weak force. The

three kinds of weakon are estimated to have masses roughly 100 times the mass of the proton; creating particles that heavy is beyond the capabilities of any existing accelerator, and it may not be possible for several years.

The historical development of elementary-particle physics might well be read as an extended lesson in skepticism. Over the course of the past century the realm of inquiry has progressed from the atom to the atomic nucleus to the hadrons that make up the nucleus to the quarks that make up the hadrons. Each of these objects has been considered for at least a time to be an elementary particle, one with no internal structure. It may be too soon to declare an end to the progression by assuming that the quarks (and the leptons) are truly elementary. They too could be composite objects, made up of simpler components. Indeed, one can imagine the endless regression that was visualized by Jonathan Swift:

> So, naturalists, observe, a flea
> Hath smaller fleas that on him prey;
> And these have smaller still to
> bite 'em;
> And so proceed *ad infinitum.*

Theorists are already contemplating the next step in the regression. For example, Haim Harari of the Weizmann Institute of Science in Israel has suggested that both the quarks and the leptons could be constructed of just two kinds of "ultimate" particles. He calls them rishons, from the Hebrew word for elementary.

The rationale for employing particle accelerators to explore the structure of matter is straightforward. Smashing two objects together is merely a means for breaking them down into their component parts. If the collision could be made sufficiently violent, the particles might be reduced to their ultimate and unbreakable constituents. An interaction between high-energy particles is not entirely like an automobile accident, however. Particles not only are knocked out of the target and the projectile but also can be created anew out of the energy brought to the collision by the accelerated particle. For example, the upsilon particle, which is thought to include a bottom quark, can be made in collisions between energetic protons even though the mass of the upsilon is 10 times that of the proton.

Another way of considering the functioning of an accelerator is by analogy with a microscope. The ultimate limit to the resolution of a microscope is the wavelength of the radiation with which the specimen is illuminated. Patterns much smaller than the wavelength cannot be resolved; as a result the light microscope cannot distinguish objects smaller than about 10^{-5} centimeter. In quantum mechanics a particle of matter can be described as a wave, which has a wavelength inversely proportional to the momentum of the particle. If an accelerator is conceived as a large microscope, the motive for increasing the energy is to reduce the particle wavelength and thereby improve the resolution. The largest accelerators now operating have an effective resolution of about 10^{-16} centimeter, which is a thousandth of the diameter of the proton.

The newest accelerators exploit the same fundamental principles as the first ones. The force employed to accelerate the particles is electromagnetism, and so only particles that have an electric charge can be accelerated; they are usually either protons (with a charge of $+1$) or electrons (with a charge of -1). The particles are injected into a vacuum chamber, the vacuum being necessary to prevent the moving particles from colliding with air molecules. An electric field then sets the particles in motion. In the simplest case a high voltage is applied across a pair of electrodes. Electrons are propelled toward the positive electrode; protons move in the opposite direction, toward the negative electrode. A simple accelerator of this kind is found in the picture tube of a television receiver.

The standard unit of measurement for particle energy is the electron volt, abbreviated eV. One electron volt is the energy acquired by an electron when it is accelerated through a potential difference of one volt. The same unit serves to measure the energy of moving protons or of any other particles. For convenience various multiples of the electron volt are employed in specifying the energy of accelerators. The kiloelectron volt (keV) is 1,000 electron volts, the megaelectron volt (MeV) is a million, the gigaelectron volt (GeV) is a billion and the teraelectron volt (TeV) is a trillion, or 10^{12} electron volts. Because mass and energy can be freely interconverted it is customary to give the mass of a particle in terms of its energy equivalent, measured in electron volts. Thus the "mass" of the proton is 938 MeV.

In principle any energy could be achieved with a simple accelerator made up of two electrodes, merely by raising the voltage to the appropriate level. In practice the maximum potential that can be sustained across a pair of electrodes is several million volts, so that such accelerators are confined to energies of no more than several MeV. The limit is determined by the onset of arcing between the electrodes or by the breakdown of insulators.

In order to reach higher energies it is necessary to accelerate a particle in stages, giving it a sequence of small pushes instead of one big push. The most obvious way of arranging for such a gradual acceleration is by stringing together many small accelerator stages one after the next. That is the principle of the linear accelerator, or linac. Instead of applying a continuous high voltage to the electrodes in each stage, an alternating electric field is set up by an oscillator connected to each set of electrodes, forming the structure called a radio-frequency cavity. The oscillators for successive cavities are synchronized in such a way that the electric field always has the correct polarity to accelerate, rather than retard, the moving particle. In effect an electromagnetic wave travels continuously through the vacuum tube and the particle rides the electrical wave as a surfer rides a water wave.

There is only one large linac now operating. It is at the Stanford Linear Accelerator Center (SLAC) near Stanford University. Completed in 1961 at a cost of some $115 million, it is two miles long and made up of 82,560 radio-frequency cavities. The SLAC device was designed to accelerate electrons to an energy of 22 GeV; a program for replacing the radio-frequency oscillators with more powerful units is raising that energy to 30 GeV. Even without the upgrading of the machine SLAC would have remained for some time to come the most powerful electron accelerator in the world.

The limit to the maximum practical energy of a linear accelerator is the cost of the thousands of accelerator cavities and their associated radio-frequency power supplies. The way to avoid that cost is to employ only a few cavities but to make each particle pass through them many times. Under the influence of a magnetic field an electrically charged particle follows a curved trajectory. By arranging many magnets in a ring the particle can be made to follow a circular orbit, or any other closed curve. A bunch or cluster of particles can circle the ring several million times, passing through the radio-frequency cavities and gaining energy on each revolution. An accelerator built in this way is called a synchrotron.

All the large new accelerators that are now planned or under construction are synchrotrons. It is therefore worthwhile to consider their operation in somewhat greater detail. The magnets that make up the ring are of two kinds. The dipole magnets, which have two poles (one north and one south), generate a uniform magnetic field; they accomplish the bending of the particle trajectories. Quadrupole magnets, which give rise to a field with two north poles and two south poles, do not deflect the particles but focus them into a narrower beam, acting much like lenses. Interspersed among the magnets are the radio-fre-

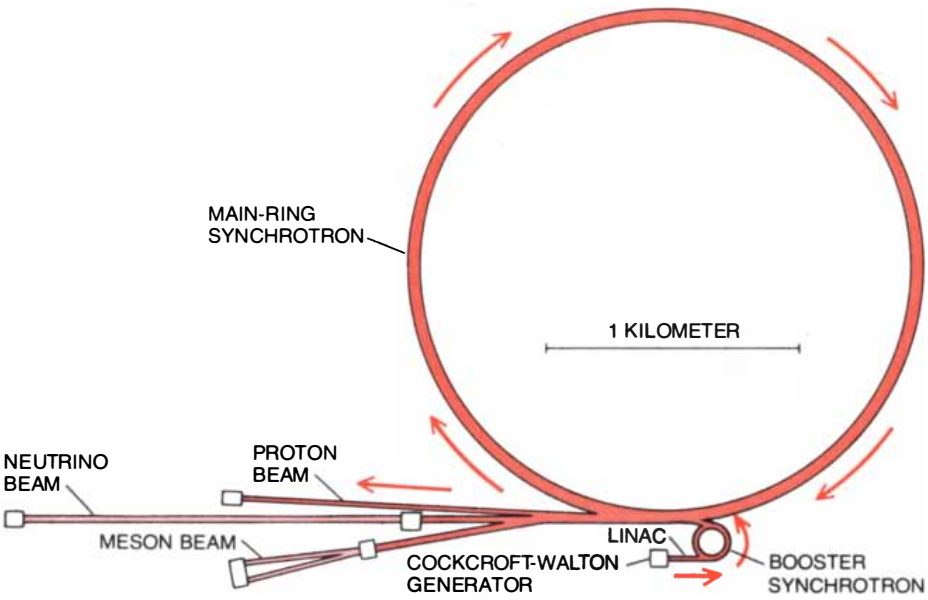
quency cavities, where the actual acceleration takes place. Specialized magnets and electrodes must also be provided for injecting the particles into the ring and for extracting them from it.

The synchrotron operates in cycles. When a bunch of particles is first injected, the fields of the bending magnets are adjusted so that the particles precisely follow the curvature of the vacuum tube. As the energy of the particles increases on each revolution the field strength in the bending magnets must also be smoothly increased. When the maximum energy is reached, the beam is extracted; then the magnetic field is allowed to fall to its original level in preparation for the next bunch of particles. The accelerator is called a synchrotron because the particles automatically synchronize their motion with the rising magnetic field and the rising frequency of the accelerating voltage.

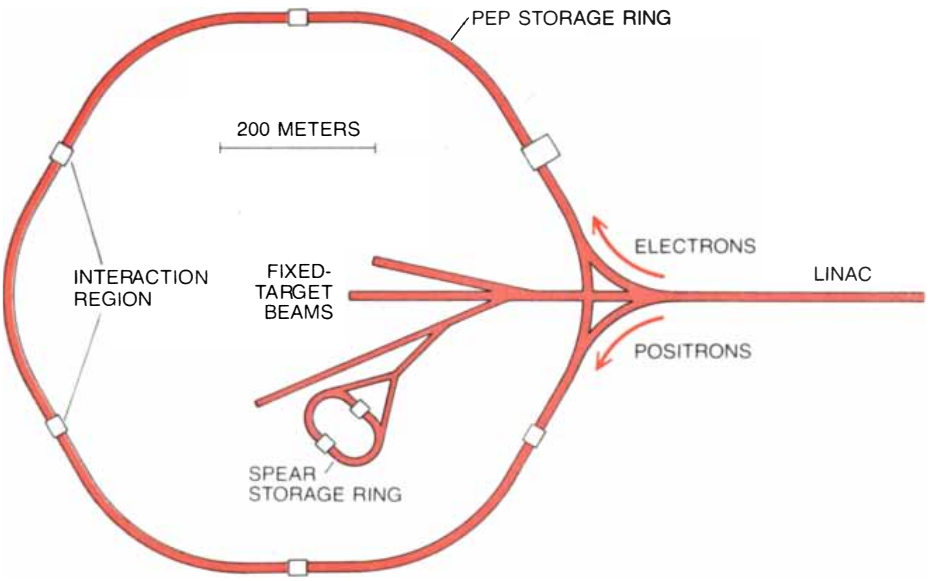
The highest energies are not attempted with a single machine; instead several machines are lined up in series. Each one augments the particle energy by a factor of 10 or even 100, then passes the beam on to the next accelerator in the sequence. In several instances older accelerators serve as injectors or preliminary stages for newer and for more powerful machines.

In a proton accelerator the first stage is most often a device of the kind built in 1928 by John D. Cockcroft and Ernest T. S. Walton at the Cavendish Laboratory of the University of Cambridge. It is a large transformer and rectifier that generates a potential of about a million volts between an inner electrode and an outer shell. Protons, obtained by ionizing hydrogen atoms, are released at the inner electrode; when they emerge (through a hole in the shell), they have an energy of about 1 MeV.

The next stage is often a linac, which typically raises the energy per proton to 50 or even 200 MeV. From the linac the protons can be injected into a synchrotron, which may be the final link in the chain or may serve merely to boost the energy of the protons for injection into a larger synchrotron.



FIXED-TARGET ACCELERATOR delivers particles to an external target, where detectors and other experimental apparatus can be set up. The accelerator shown is the one at Fermilab. Protons are accelerated in four stages: a Cockcroft-Walton generator (.75 MeV), a linac (200 MeV), a booster synchrotron (8 GeV) and the main-ring synchrotron (400 to 500 GeV). Experiments can be carried out not only with the accelerated protons but also with beams of secondary particles, such as mesons and neutrinos, knocked out of the target by the impact of protons.



COLLIDING-BEAM DEVICES store particles at high energy and bring them into collision head on. The storage ring shown is called PEP; it is now under construction at the Stanford Linear Accelerator Center (SLAC). Electrons and positrons from the two-mile linear accelerator at SLAC will be injected into the PEP ring in opposite directions and maintained at an energy of up to 18 GeV. Collisions between the counter-rotating beams will take place in six interaction zones around the circumference, where detectors can be built around the beam pipe. Smaller SPEAR ring has been operating since 1972 at center-of-mass energies of up to 8 GeV.

The first large synchrotrons were the Cosmotron, built in 1952 at the Brookhaven National Laboratory, and the Bevatron, completed in 1954 at the University of California at Berkeley. They reached energies of 3 GeV and 6.2 GeV respectively. In design they differed from more recent practice chiefly in the disposition of the magnetic field, which provided only weak focusing of the beam. In building the next generation of larger and more powerful synchrotrons, a new system of strong focusing was introduced. The shape of the magnetic field can be described mathematically as being partly uniform (the dipole component) and partly a gradient in a direction transverse to the orbit of the beam (the quadrupole component). The quadrupole component was made stronger and was alternated in sign, so that oscillations of the particles around the desired orbit were more frequent but of smaller amplitude. As a result of this alternating gradient the aperture of the magnets and the bore of the vacuum chamber could be made smaller. It is the invention of the synchrotron and of strong focusing that has made the very large accelerators of today economically feasible.

The strong-focusing principle was first applied to synchrotrons built at Brookhaven and in Europe. The Brookhaven machine, which came to be known as the Alternating Gradient Synchrotron, or AGS, was completed in 1961 and eventually reached an energy of 33 GeV. The AGS has had a distinguished career. It was an experiment with the AGS that first revealed the exis-

tence of two kinds of neutrino, one for the electron and one for the muon. In 1974 an AGS experiment was one of two that discovered the particle labeled *J* or psi, which provided the first evidence of charm. The AGS is still operating and, as I shall explain below, there are plans for its further utilization.

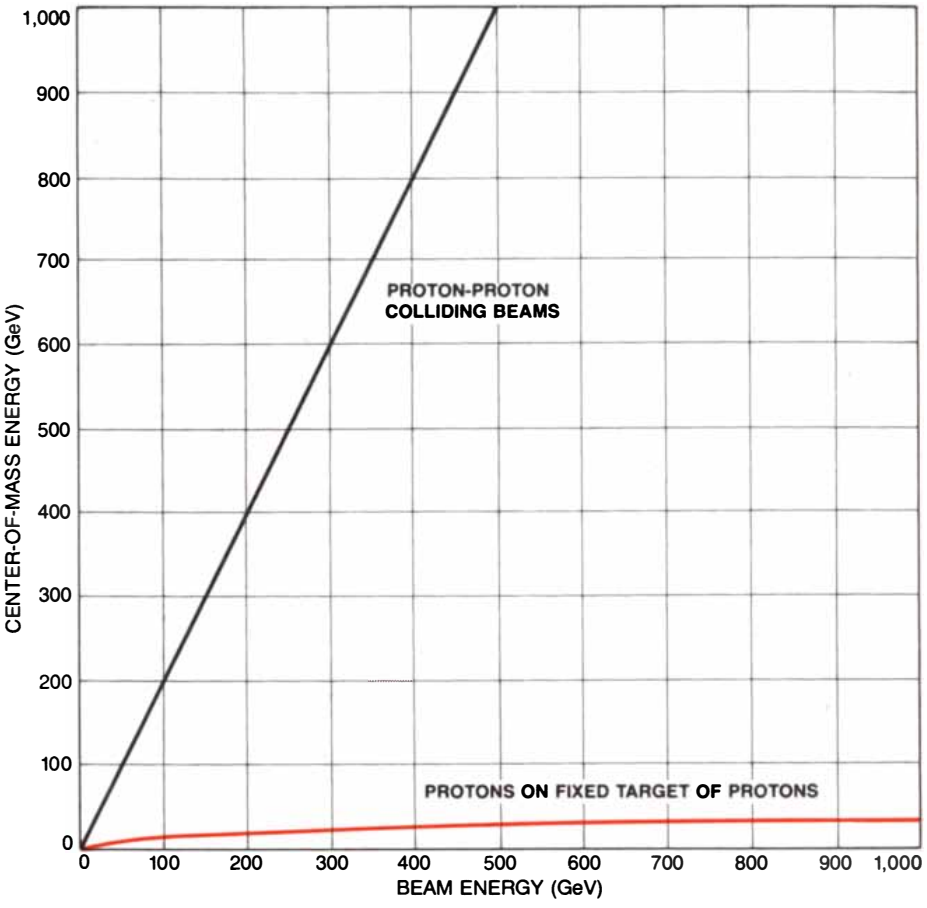
The European strong-focusing synchrotron has turned out to be part of a much larger endeavor, what is indeed now the largest of the world's high-energy-physics laboratories. The project was undertaken by a consortium of European nations, which formed the Conseil Européen pour la Recherche Nucléaire, or CERN. (The name has since been changed to European Organization for Nuclear Research, but the abbreviation remains.) It is a model of international cooperation, which has broken new ground in solving the problems of language, money, custom and national self-interest. Today there are 12 member nations. The laboratory itself now straddles the French-Swiss border a few miles north of Geneva.

The Proton Synchrotron, or PS, at CERN was completed about a year ahead of the Brookhaven AGS. The PS too has made distinguished contributions. Most notably it was the instrument with which CERN physicists discovered a new aspect of the weak force, called the neutral weak current, which provided the first evidence supporting the unification of the weak and electromagnetic forces.

Research is still carried out with the proton beams of the PS, but what is now probably the most important function of that machine is to serve as the injector for a still larger accelerator, the Super Proton Synchrotron, or SPS. The initial plans for the SPS called for a 300-GeV accelerator to be built somewhere in Europe other than the CERN site. In 1965, however, France made available a tract of land adjoining the original laboratory in Switzerland, so that the SPS could be built next to the PS. As it happened, there was no need to disturb the surface over most of the area available. The tunnel for the SPS, which is almost seven kilometers in circumference, was bored underground with a mining machine, at an average depth of 40 meters. The beam comes to the surface only in the experimental halls.

When the SPS began operation in 1976, its energy was not 300 GeV but 400, and it has the capability of reaching 500 GeV. There is no question that the availability of the PS as an injector speeded construction and made the higher energy economically feasible. The highest-energy protons at CERN pass through five accelerators: a Cockcroft-Walton generator (550 keV), a linac (50 MeV), a booster synchrotron (800 MeV), the PS and the SPS.

At about the same time planning began for the SPS, an American project of comparable scale was undertaken. After considerable bickering among states competing for the new laboratory, a site was chosen in 1967, on farmland 30 miles west of Chicago. The facilities built there are now called the Fermi National Accelerator Laboratory, or Fermilab. It does not have the cachet of international sponsorship, but at least it is interscholastic: the operating authority is an association of 53 universities. Of course, physicists of many nations do experiments at the laboratory.

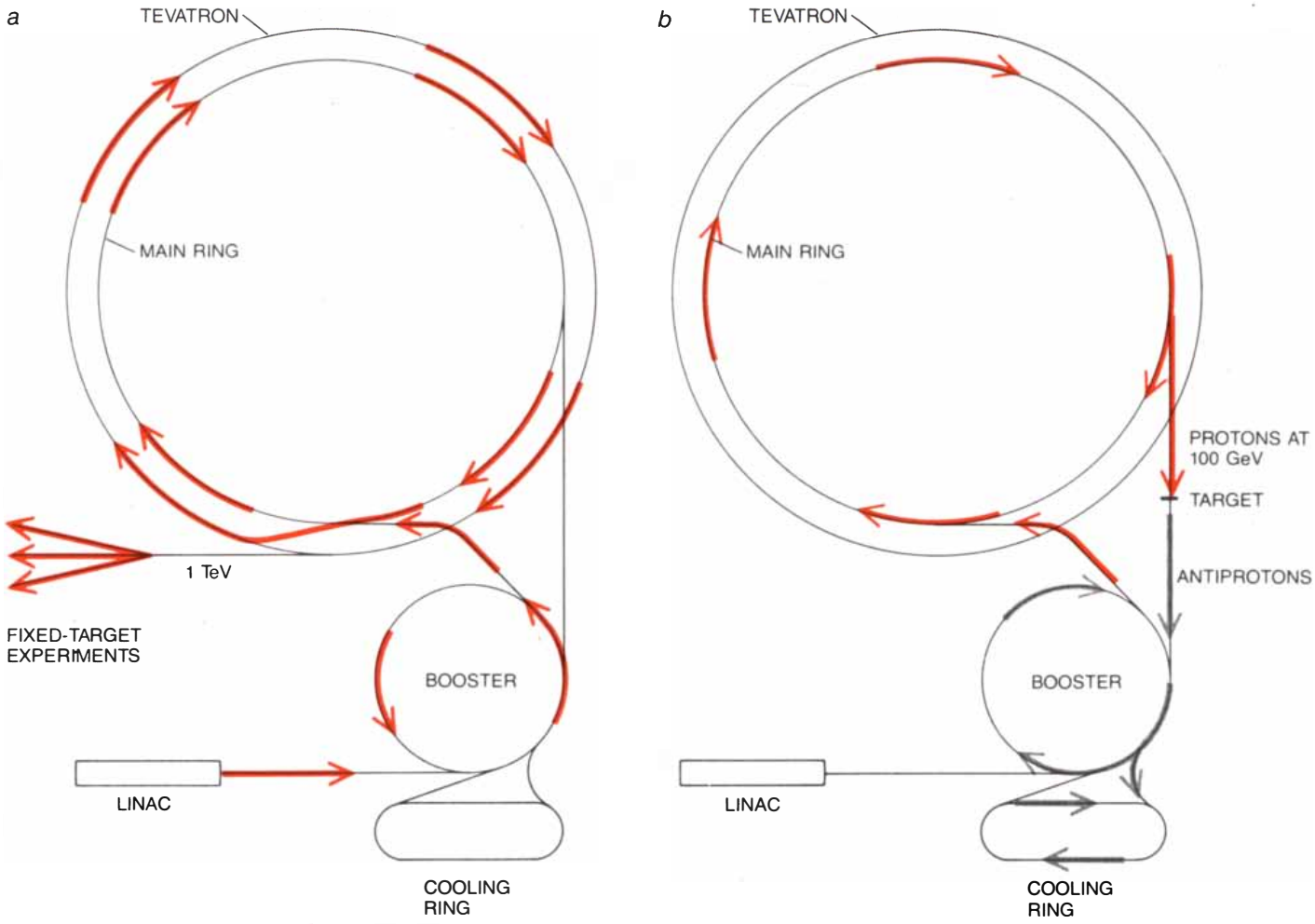


EFFECTIVE COLLISION ENERGY is much greater in colliding-beam devices than it is in fixed-target accelerators. The energy available for creating new particles is the energy of the collision when the collision is viewed in the frame of reference of the center of mass of the colliding particles. In a storage ring the center-of-mass energy is simply twice the beam energy. (It is assumed that both particles have the same rest mass and are accelerated to the same energy.) In a fixed-target accelerator at low energy the center-of-mass energy is proportional to the square root of the beam energy, and at high energy it rises even more slowly because of relativistic effects. Fixed-target accelerators have the redeeming property of higher luminosity, which is a measure of the rate at which particle interactions can be observed. Fixed-target machines also yield a variety of secondary beams, which cannot be generated by a colliding-beam device.

Work at Fermilab proceeded faster than at CERN, and the accelerator began operating in 1972. The original plan had called for a 200-GeV proton synchrotron, but it was possible to build a 400-GeV machine for the same cost. By 1975 the maximum energy had been raised to 500 GeV, although the cost of electric power forbids extended periods of operation at that energy.

The system of accelerators at Fermilab follows what by now should be a familiar pattern. Protons are given an initial push in a Cockcroft-Walton generator (750 keV), then pass through a linac (200 MeV) and a booster synchrotron (8 GeV) before being inserted into the main-ring synchrotron. The main ring is a little smaller than that of the SPS; its circumference is 6.3 kilometers. It is made up of 774 bending magnets, each 20 feet long, and 180 smaller focusing magnets. A series of radio-frequency cavities add 2.8 MeV to the protons' energy on each revolution. Going from 8 GeV to 400 GeV therefore requires 140,000 turns, which are completed in about three seconds. Controlling the rapidly changing events in such a large and complicated machine requires a battery of three large digital computers and many smaller satellite computers.

The process whereby a prima ballerina accelerator becomes a mere member of the ballet company, serving as the injector for a larger device, is about to be repeated at Fermilab. A new ring of magnets and radio-frequency cavities is being installed under the existing main ring, so that there will be two accelera-



VERSATILITY OF THE MULTIPLE RINGS at Fermilab will enable the facility to operate either as a fixed-target accelerator or as a colliding-beam device. Protons will be delivered to external targets (*a*) either by the main ring (at energies of up to 500 GeV) or by the Tevatron (at 1 TeV). One scheme for colliding beams would employ the Tevatron to store counter-rotating protons and antiprotons. Protons would first be brought up to about 100 GeV in both synchrotrons; they would then be extracted from the main ring and employed

tors in the one tunnel. Protons will be brought up to an energy of about 150 GeV in the main ring, then transferred to the lower ring for further acceleration. The new ring is made up of superconducting magnets, which can reach a field strength twice that of the old magnets. The maximum energy of the accelerator is thereby doubled, to 1 TeV. In recognition of this milestone the new accelerator has been named the Tevatron.

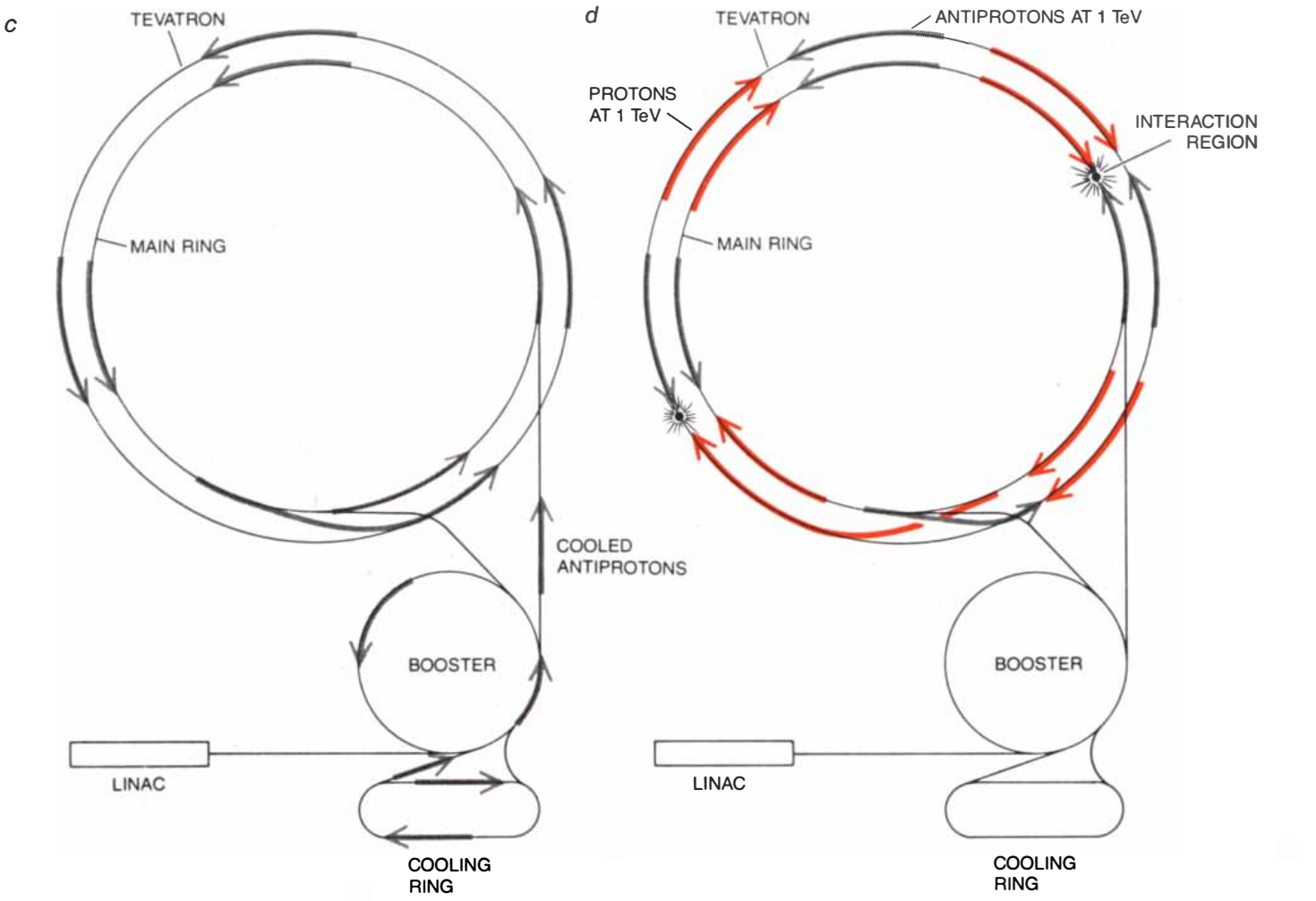
The use of superconducting magnets was contemplated when plans were first made for Fermilab, but the risk of an unproved technology seemed too great then. Building the Tevatron has turned out to be a challenging task even today. The most obvious problem is that of cooling almost 1,000 magnets, strung out over 6.3 kilometers of tunnel, down to 4.5 degrees Kelvin, the temperature where the special conductors of the magnet windings lose all resistance to the flow of electricity. In order to maintain that temperature a river of liquid helium will be pumped through the ring. Twenty-four small refrigeration plants will be spaced around the tunnel, and the central helium liquefier will be the largest in the world, with a capacity of 4,000 liters per hour.

The windings in the magnets are formed of a niobium-titanium alloy embedded in a copper matrix. Almost 19,000 miles of this cable will be required to complete the ring, qualifying Fermilab for another superlative: it is the largest consumer of superconducting materials in the world. At maximum field strength the superconductors will carry a current of 4,600 amperes, and when a magnet is "quenched," or loses the superconducting property, the energy stored in the field (about half a million joules per magnet) must be dissipated without destroying the windings.

A particularly taxing problem has been the need to maintain the uniformity of the magnetic field to an accuracy of better than one part in 1,000. Because the windings are immersed in their own field they are subjected to a reactive force of about a ton per linear inch. The coils cannot be allowed to move even a thousandth of an inch, however, because the movement would distort the field and might also give rise to too much frictional heat. The windings are immobilized by laminated collars of stainless steel. The alignment of the magnets is also complicated by thermal contraction when the ring is cooled to its operating temperature; a magnet six meters long contracts by about two centimeters.

There is now substantial confidence that these problems have been understood and solved. A string of 24 superconducting magnets has been installed in one segment of the tunnel, and a beam at 90 GeV has been diverted through it from the main ring. The magnets performed as expected. Completion of the full lower ring is now expected toward the end of 1981, and protons at 1 TeV should be delivered to the experimental areas in 1982. The rate of construction depends mainly on the rate of funding.

It is worthwhile pausing to consider just how much energy 1 TeV per proton is. In units more commonly applied to macroscopic objects, a 1-TeV particle has an energy of 1.6 ergs, which is roughly the kinetic energy of a flying mosquito. At full intensity the Tevatron will accelerate 5×10^{13} protons at a time, which will give the total beam an



to generate antiprotons (*b*), which have the same mass but the opposite electric charge. The antiprotons would be shunted through the booster to a "cooling" ring, where they would be formed into a well-collimated beam. When enough antiprotons had been accumulated, they could be returned through the booster to the Tevatron, where they would be injected counterclockwise, opposite to its usual sense of rotation (*c*). Protons and antiprotons would then be accelerated simultaneously to 1 TeV, yielding a center-of-mass energy of 2 TeV.

energy of eight million joules. That is comparable to the energy of a 100-pound artillery shell. If the beam should ever go out of control, it could melt the walls of the vacuum chamber and destroy the surrounding magnets; obviously such an accident must be avoided.

When the particles in a synchrotron have reached their full energy, they are nudged out of their orbit by a special magnet and deflected into an external beam line. Eventually they strike a target. Interactions of the protons with the target can be studied directly, and it is also possible to create beams of secondary particles knocked out of the target. At Fermilab, for example, there are separate areas for experiments with protons (the primary particles), with mesons (particles of intermediate mass, such as the pion), with neutrinos and muons, and with photons. Neutrino experiments have been exceptionally rewarding in recent years because the neutrino is subject only to the weak force; the properties of that force can therefore be observed without interference from other kinds of events.

The detectors and other apparatus required for the experiments are often built on a grand scale, comparable to that of the accelerators. The armor plating of two decommissioned battleships has been incorporated into the shielding of a detector at Fermilab. A detector now installed at CERN employs many tons of fine Carrara marble. These detectors are "counters": they sense the energy and direction and penetrating power of various particles and record them electronically for later analysis. Another kind of detector is the bubble chamber, where the passage of an electrically charged particle leaves a telltale track of small bubbles. From photographs of the tracks the sequence of events following a particle collision can be reconstructed. Fermilab has a bubble chamber 15 feet in diameter; CERN has three chambers, the largest about 12 feet in diameter.

The CERN and Fermilab machines are both proton accelerators, but electrons can also serve as the working medium of an accelerator. The physical principles are exactly the same, although the characteristics of the machine are somewhat different. The main reason for the difference is the greater importance in electron accelerators of the electromagnetic radiation that dissipates the energy of an accelerated particle.

It has been understood for more than a century that any accelerated electric charge must radiate electromagnetic waves. "Acceleration," in this context, refers not only to a change in the speed of a particle but also to any change in its direction. If it is made to follow a circular path, then it undergoes a continuous acceleration even if the speed is constant. The radiation emitted under these conditions is called synchrotron radiation, because it was in the synchrotron that it was first observed.

The synchrotron radiation appears as an intense beam of high-energy electromagnetic waves, with a continuous spread of wavelengths extending to the ultraviolet and X-ray regions of the spectrum. The energy drained away by this process must be made up by supplying additional radio-frequency power. Thus synchrotron radiation acts as a resistive force analogous to friction.

The energy emitted in the form of syn-

chrotron radiation varies inversely as the mass of the particle raised to the fourth power. Because the mass of the proton is 1,836 times that of the electron the problem of supplying the lost energy is greater for electrons by a factor of 10^{13}. At the beam energies that have been attained so far synchrotron radiation is not a significant consideration in the design of proton accelerators, whereas it is the principal restraint on the energy of electron accelerators.

One solution to the problem of synchrotron radiation is to build a linear accelerator, where there is no curvature and where the radiation resulting from the gentler straight-line acceleration of the electrons is negligible. It was for this reason that a linac was chosen for the SLAC electron accelerator. The principal drawback of this approach is again the cost of the many radio-frequency cavities and the oscillators needed to excite them. In a linac an electron passes through each cavity only once, which is a great extravagance. SLAC cost $115 million; Fermilab cost almost $250 million, but the cost per electron volt was more than 10 times less than it was at SLAC.

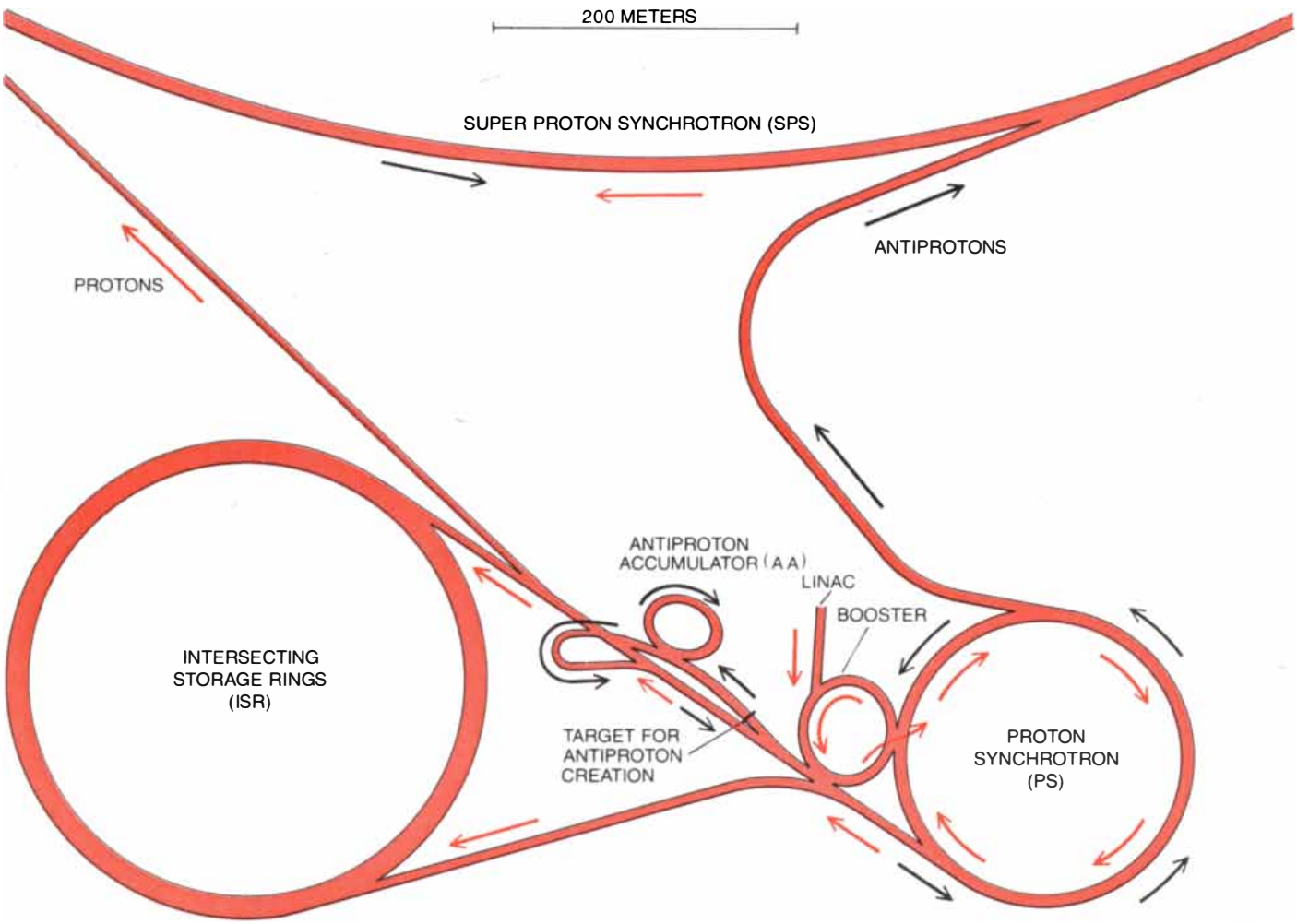
The alternative to a linac is building a synchrotron for electrons and accepting the cost of synchrotron radiation. A compromise can be reached between construction costs and the continuing expense of operating the radio-frequency power supplies. The energy radiated per revolution is inversely proportional to the radius of curvature of the particle track, and so the energy loss declines as the accelerator is made larger.

In 1965 a 10-GeV electron synchrotron was built at Cornell University in a tunnel dug under an athletic field. At the time the machine was notable for its size; the circumference is about 630 meters. It reached about half the energy of SLAC at about a tenth the cost.

Another electron synchrotron, with a beam energy of 6.3 GeV, was constructed in 1972 in Cambridge, Mass. It has since been dismantled. A third was built at Hamburg in West Germany; it is called DESY, standing for Deutsches Elektronen-Synchrotron. With a beam energy of 7 GeV, DESY has now become the basis for a diversified high-energy-physics laboratory.

For a given radius of curvature the energy lost to synchrotron radiation increases as the fourth power of the beam energy. The energy loss becomes the dominant consideration in the design of the accelerator at about 10 GeV, and it imposes an almost impenetrable barrier beyond a few hundred GeV. Of course, the feasibility of any instrument is a matter of judgment and is subject to change; 25 years ago the limit of electron synchrotrons was thought to be about 1 GeV.

It should be said in defense of synchrotron radiation that it is not entirely wasted energy. One effect of the radiation is to damp out small excursions of the electrons away from their mean trajectory, making a beam of electrons easier to control than a beam of protons. What is more important, the radiation itself has become a valuable tool for



CONVERSION OF A SYNCHROTRON is also planned at the European Organization for Nuclear Research (CERN). A proton-proton colliding-beam facility, the Intersecting Storage Rings (ISR) has been operating at CERN since 1971. It consists of two interlaced rings, with eight crossover zones where collisions can be observed at center-of-mass energies of up to 62 GeV. A project now under way will employ the much larger ring of the Super Proton Synchrotron (SPS) for proton-antiproton collisions. The source of protons for all the CERN devices is the Proton Synchrotron (PS), completed in 1959. Antiprotons will be created as secondary particles from the PS beam and collected in a cooling ring called the Antiproton Accumulator (AA). Protons and antiprotons will then be injected in opposite directions into SPS and brought up to 270 GeV each. Hence maximum center-of-mass energy will be 540 GeV. Only a small arc of SPS is shown.

biological and materials studies. It is the most intense source known of ultraviolet radiation and X rays. Facilities for exploiting the radiation have been set up at several high-energy-physics laboratories, and a number of small accelerators have been built explicitly as sources of synchrotron radiation.

In spite of the vexing problem of synchrotron radiation, electron accelerators will have an increasingly important part in future high-energy-physics programs. The reason is that the electron is a much simpler particle than the proton, and therefore makes a better probe of the structure of matter. When two protons (or other hadrons) collide, the outcome is complicated by the multiquark structure of the particles. The electron gives no evidence yet of having an internal structure, and so the results of electron interactions are more easily interpreted. This purity is achieved in double measure when an electron collides with another electron or with a positron. Such collisions can be arranged by building a device in which an accelerated beam impinges not on a fixed target but on another beam of particles.

If the entire 500 GeV per particle that is generated by the largest proton synchrotrons were released in a collision with a fixed target, many of the goals set for the next generation of particle accelerators would already have been achieved. The weakon, for example, is expected to have a mass near 100 GeV, and so it would have been seen by now. Not all the energy of an accelerated particle is made available for the creation of new particles, however, when a moving projectile collides with a fixed target. Instead a large fraction of the energy goes into setting in motion the system made up of the two particles.

Consider a 500-GeV proton striking a stationary proton, such as the nucleus of a hydrogen atom. If the projectile were stopped by the collision, then all the energy it had acquired would have to be dissipated. What actually happens is that the accelerated proton keeps moving and the target proton moves in the same direction along with it. Indeed, because of relativistic effects the rapidly moving accelerated proton has a mass 530 times that of the target proton. The accelerated particle is not stopped by the collision, just as a truck is not stopped when it collides with a bicycle. The two particles together retain an energy of 469 GeV, and only 31 GeV is made available for the creation of new particles.

The energy released in a particle collision can be calculated most easily by changing one's frame of reference. The collision is seen most realistically by an observer who is moving parallel to the beam and at the same speed as the center of mass of the two-particle system. Such an observer would see the beam proton and the target come together symmetrically with an energy of 15.5 GeV each, or a total energy of 31 GeV. That is the center-of-mass energy. According to the kinematics of relativity theory, the center-of-mass energy grows only in proportion to the square root of the beam energy, so that as accelerators become larger more of the energy invested is in this sense wasted. In order to reach a center-of-mass energy of 100 GeV with a fixed-target proton accelerator the beam energy would have to be more than 10,000 GeV.

The view of the collision from the center-of-mass frame of reference suggests a solution to this problem. If the two particles are both accelerated to the same energy but are moving in opposite directions, they can be made to collide head on. In the simplest analysis both could be stopped in their tracks. Then all their energy would be liberated. If the particles have the same rest mass and if they have been accelerated to the same energy, the center-of-mass energy is simply the sum of the two beam energies. In order to reach 100 GeV all that is needed is a pair of beams with energies of 50 GeV each.

Head-on collisions can be arranged by building storage rings, in which beams of particles circulate continuously. A storage ring resembles a synchrotron: it has an annular vacuum chamber, which is surrounded by bending and focusing magnets, and has at least one radio-frequency cavity. Usually the ring is employed not to increase the particles' energy but to maintain them in a constant orbit and at a constant energy. The radio-frequency cavity supplies only enough power to make up for the energy lost by synchrotron radiation.

One method of arranging head-on collisions is to build two storage rings that are tangent to each other. The particles collide at the point of contact. In another plan two rings are interlaced, each of them following an undulating path like the strands of a braided hoop; the beams collide at each intersection. The cleverest of the storage-ring designs is possible only when the beams being stored are made up of particles and their corresponding antiparticles, such as electrons and positrons. It is then possible to make do with a single ring. Suppose the magnetic fields and the radio-frequency cavities are adjusted to maintain electrons circulating clockwise at some fixed energy. Positrons, having the opposite electric charge, respond to magnetic and electric fields in a manner exactly opposite to that of electrons. The positrons can therefore be injected into the same ring, but in a counterclockwise direction, and the single set of magnets and radio-frequency cavities will maintain both kinds of particle in orbit. If the ring holds one bunch of electrons and one bunch of positrons, collisions will be observed at two points diametrically opposed.

Given the enormous energy advantage of a colliding-beam device, it may seem surprising that anyone would now contemplate building a fixed-target accelerator. Energy, however, is not the only pertinent measure of an accelerator's performance. Another important factor is the rate at which interactions are observed. A solid or liquid target has a far greater density of particles than an accelerated beam, with the result that the interaction rate is much higher in a fixed-target accelerator. In the colliding-beam device, most of the time the particles in the beams do not collide but pass through one another without interacting at all. The probability of interactions is proportional to a parameter called the luminosity, which essentially measures the brightness of the beam. The luminosity of fixed-target machines is greater than that of storage rings by a factor of about a million.

Interpreting the results of a physics experiment requires a statistically significant sample of events. Collecting such a sample can be difficult when the events of interest take place at a rate of only a few per day. For this reason storage rings may be expected to give the first glimpse of the commonest phenomena at high energy, but greater accuracy and the observation of rare events may follow only when the fixed-target accelerators reach comparable center-of-mass energies.

Another advantage of the fixed-target accelerator is versatility. A storage ring generally provides only one kind of collision, such as protons on protons or electrons on positrons. A fixed-target machine, on the other hand, can generate a variety of secondary beams, including neutrinos, muons, pions and other mesons, antiprotons and the massive particles called hyperons. Indeed, the small fraction of the beam energy that is made available in the center-of-mass system is an advantage for creating secondary beams; it is the large quantity of "wasted" energy that accelerates the secondary particles.

One further advantage of the fixed-target accelerator has already been described at length: the machine built to supply particles for research can later supply them to a larger accelerator, or to a storage ring.

The first electron-positron storage ring was begun in 1959 by Bruno Touschek and an enthusiastic group of Italian physicists at the National Laboratories of the National Committee for Nuclear Energy (C.N.E.N.) at Frascati near Rome. Called ADA, for Annello d'Accumulazione, the ring was later moved to the Orsay laboratory outside Paris. The energy of each beam was .25

GeV, yielding a center-of-mass energy of .5 GeV.

By the early 1970's half a dozen other storage rings were in operation. ADA was the direct progenitor of a larger ring at Orsay (a third has since been built there) and of a ring at Frascati that was named ADONE, for "Big ADA," with an energy of 1.5 GeV per beam. The first in a succession of pioneering storage-ring projects at Novosibirsk in the U.S.S.R. was completed in 1965. The Cambridge Electron Accelerator was modified at about that time to form a storage ring with an energy of 3.5 GeV per beam. The life span of the Cambridge machine was brief, but the machine revealed an anomaly in the total rate of electron-positron interactions that presaged the discovery of charm.

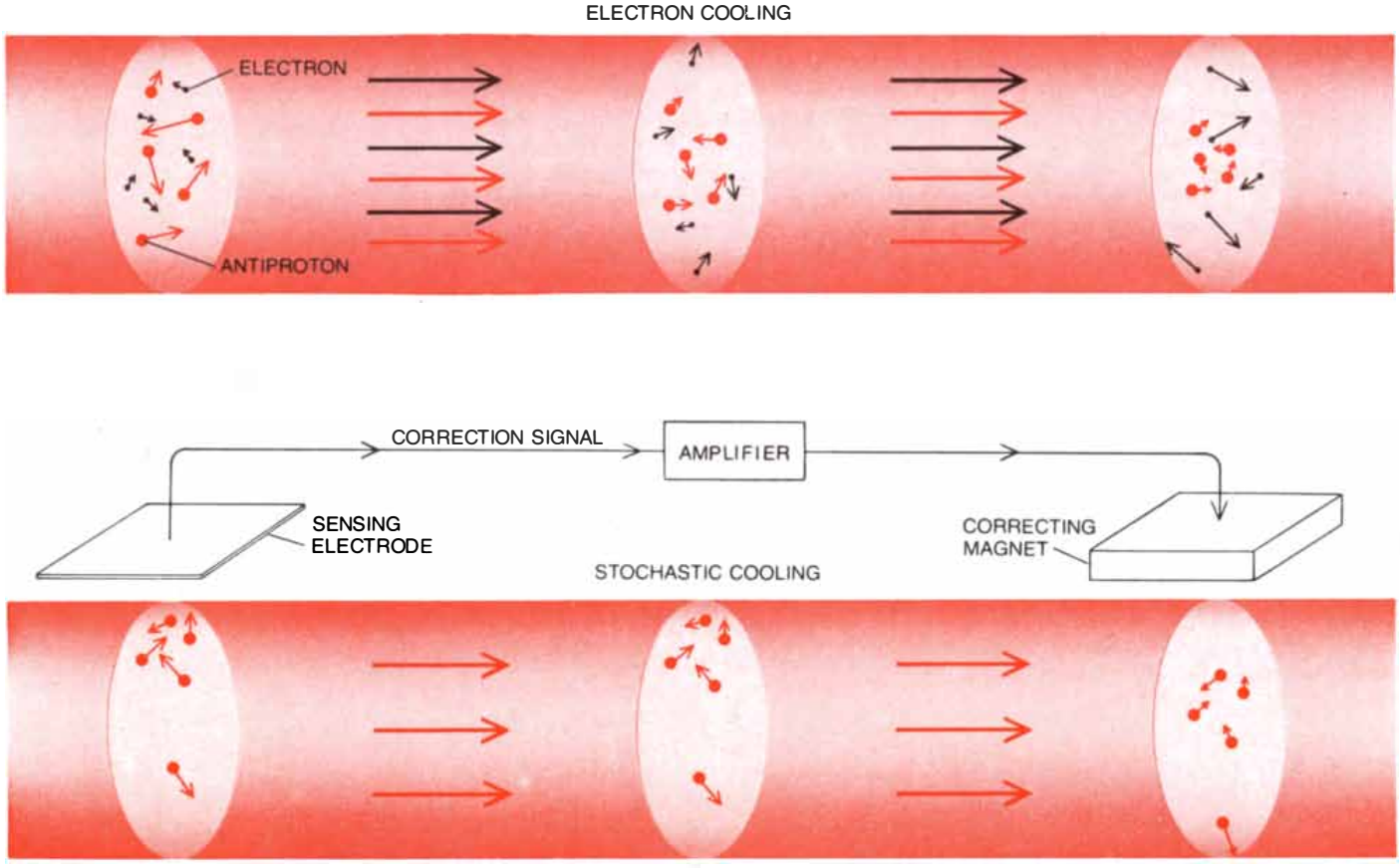
By 1974 storage rings at two laboratories had exceeded 4 GeV per beam. One of them, called SPEAR, is at SLAC and is filled with electrons and positrons from the linac there. The other, called DORIS, was built next to the DESY synchrotron in Hamburg, and receives its particles from that source.

It was with the SPEAR storage ring that the first charmed particle, the *J* or psi, was seen at 3.1 GeV. As I have mentioned, the same particle was discovered simultaneously in a fixed-target experiment at Brookhaven, but the signal at SPEAR was much clearer. Workers at DORIS were able to detect the new particle immediately when they tuned their instrument to the right energy. ADONE had just missed seeing the psi because it was designed for a maximum center-of-mass energy of exactly 3 GeV, but it was possible to nudge the machine over this limit and detect the particle. The subsequent exploration of the spectrum of charmed particles was carried out largely with these three storage rings.

The discovery of charm revitalized elementary-particle physics. Obviously it pleased the experimenters who had made the discovery, but it was also rewarding to theorists, who had predicted the existence of the charmed quark and had waited 10 years for it to be revealed. After such a success it is little wonder that plans were immediately drawn up for a new round of electron-positron storage rings. A proposal to build a larger ring at SLAC, in collaboration with the Lawrence Berkeley Laboratory, was soon approved; it inherited the name PEP. In Europe two quite similar projects were considered, one in Hamburg to be called PETRA and the other at the Rutherford Laboratory in England. The two proposals were brought before an international body, the European Committee for Future Accelerators; following the recommendation of the committee, construction began on the German project but the Rutherford plan was dropped.

The specifications of PETRA and PEP are remarkably similar. PETRA is designed to reach a maximum energy of 19 GeV per beam, PEP 18 GeV. In size (2.2 kilometers in circumference) and luminosity they are almost identical. Most likely the costs are also similar, although differences in accounting methods make the comparison difficult.

As soon as construction began it was apparent the two laboratories were in a race. PETRA has won: it began operating in 1978, whereas PEP is not expected to be in operation until early this year. The faster progress of PETRA can be attributed to the elegant but spare



BEAM-COOLING TECHNIQUES are essential to the success of proton-antiproton storage rings. The antiprotons are created with a comparatively wide range of speeds and directions and therefore cannot be injected directly into an accelerator. When the protons are viewed in their own frame of reference, they make up a hot gas, the particles of which have randomly directed velocities. One means of cooling the gas employs electrons to carry off the heat. In repeated collisions with the antiprotons the electrons come away with most of the randomly directed momentum. In another cooling method (stochastic cooling) the average position of the beam is sensed, and then a correction is applied to keep the beam centered in the vacuum tube. Some particles may be affected adversely by the correction, but a majority will always have their extraneous components of motion reduced. After many repetitions of this process the antiproton beam will be cooled. Only the transverse component of the particle velocities is shown; there is also variation in speed along the axis of beam.

design of Gustav Adolph Voss, to the fact that the German group had a head start and perhaps also to the fact that money was made available as it was needed. Indeed, the builders of PETRA were encouraged to spend quickly in order to stimulate business in the Hamburg area. PEP has been nursed along on funds doled out almost a dollar at a time, and it has also been slowed by heavy rains (which flooded the tunnel excavations) and by labor trouble.

The first experiments at PETRA suggest that both machines have an interesting future. Preliminary data provide encouraging support for quantum chromodynamics, the theory that has been devised to describe the interactions of quarks and gluons inside hadrons.

The race to exploit the simplicity of electron-positron collisions at higher energy is not being contested by PEP and PETRA alone; a dark horse has emerged. It is the Cornell Electron Storage Ring, or CESR, which has been squeezed into the same tunnel with the existing Cornell electron synchrotron. Under the direction of Boyce D. McDaniel of Cornell the project was completed in less than two years and at a cost of less than $20 million. The first electron-positron collisions were detected last summer, and experiments have now begun. CESR is unusual in being sponsored by the National Science Foundation; virtually all other high-energy-physics laboratories in the U.S. are funded by the Department of Energy.

An innovative method of filling the storage ring has been introduced at CESR by Maury Tigner of Cornell. Bunches of particles are passed back and forth between CESR and the electron synchrotron, which have different radii. By taking advantage of the slight delay in the outer ring, many bunches can be superposed in the storage ring, thereby improving the luminosity of the beams. The maximum energy is 8 GeV per beam, which may turn out to have been a fortunate choice. It is intermediate between the energies of SPEAR and DORIS and those of PEP and PETRA, and it falls in a range where a rich spectrum of new particles is expected to be found. Those particles are the ones related to the upsilon, incorporating the bottom quark.

Heady with success, the Cornell group has made an initial design for an electron-positron storage ring capable of 50 GeV per beam. With a center-of-mass energy of 100 GeV, a particularly exciting prospect is the copious production of the neutral weakon, or Z^0. At that energy the electrons and positrons would radiate 1.6 GeV of synchrotron radiation per turn; this energy loss would be made up by superconducting radio-frequency cavities with a total power consumption of 36 megawatts. The machine would be four kilometers in circumference, and Tigner estimates it would cost $120 million.

A still grander electron-positron device is already in an advanced stage of design. Proposed by physicists at CERN, it is called LEP, for large electron-positron storage ring. Several candidates for the next large-scale European high-energy-physics facility were submitted to the European Committee for Future Accelerators in 1978. LEP won the endorsement of the committee, and a detailed design study was therefore undertaken. No funds have yet been committed to the construction of the ring, but if money is made available for a large European project, LEP will presumably have priority.

In an initial stage the maximum energy of LEP might be about 86 GeV per beam, but that could later be raised to as much as 130 GeV by adding superconducting radio-frequency cavities. In order to minimize the energy loss through synchrotron radiation the ring would be made very large, with a circumference of 30 kilometers. As with the SPS, the tunnel would be excavated from underground. Indeed, it would extend under the Jura Mountains west of Geneva, and three of the experimental halls would be half a mile underground. By locating LEP adjacent to the SPS, facilities could be added later for collisions between electrons in one ring and protons in the other. Indeed, once the LEP tunnel is in place a ring of superconducting magnets for a proton synchrotron could be added to it. Protons injected by the SPS could be brought to an energy of between 3 and 6 TeV, although this potentiality has not yet been publicly mentioned by the proponents of the machine.

The estimated cost of the initial phase of LEP is 1,275 million Swiss francs (some $800 million), more than the combined cost of Fermilab and the CERN SPS. If construction were started soon, the ring might be operating by the late 1980's.

There is no reason for colliding-beam devices to be confined to electrons and positrons. In fact, a proton-proton collider has been operating at CERN since 1971. Called the Intersecting Storage Rings (ISR), it consists of two interlaced rings that store counter-rotating proton beams. The rings cross over at eight points around their circumference, meeting at a shallow angle. Detectors can be set up in seven of the interaction zones.

The protons stored in the ISR are supplied by the adjacent PS, at energies of up to 28 GeV. In the ISR they can be accelerated slightly, to an energy of 31 GeV per beam. The resulting center-of-mass energy, 62 GeV, is the highest available today. To reach the same center-of-mass energy with a fixed-target accelerator would require a beam energy of 2 TeV.

A device that is similar to the ISR in conception but much larger in scale is now under construction at Brookhaven. Called ISABELLE (for Intersecting Storage Accelerator plus *belle,* for beautiful), it will consist of two interlaced rings, but they will be about four times as large as those of the ISR. Moreover, both rings will be made up of superconducting magnets, with a much higher field strength than that of conventional magnets. The rings will operate at 400 GeV per beam, for a center-of-mass energy of 800 GeV. The equivalent fixed-target accelerator would have a beam energy of 340 TeV.

Protons for ISABELLE will come from the AGS. They will be injected at about 30 GeV; then each of the ISABELLE rings will be operated briefly as a synchrotron to bring the protons up to maximum energy. Thereafter the magnet currents will be held fixed, and the protons will circulate with constant energy. There are six intersections, and there is adequate space for experimental facilities at all of them.

The superconducting magnets differ in several details from those being built at Fermilab for the Tevatron. For example, the vacuum chamber threaded through the Tevatron superconducting magnets will be cooled to the liquid-helium temperature of the magnets, whereas the vacuum tube in ISABELLE will be warm. On the other hand, the comparatively massive iron yoke of the ISABELLE magnets will be cooled, whereas the Tevatron yoke will remain at room temperature. The cooling systems themselves will be quite different. The Tevatron will employ liquid helium, but ISABELLE will be cooled by the forced circulation of high-density helium raised to a pressure beyond the critical point, where the distinction between liquid and vapor disappears. Because there is so much iron in the ISABELLE magnets it is expected to take two weeks for them to be cooled to their working temperature of 3.8 degrees K.

The aperture of the ISABELLE magnets will be greater, allowing for a larger-bore vacuum chamber. That in turn will make possible more intense proton beams and higher luminosity. The aim is a luminosity 10 times greater than that of any previous colliding-beam device, and 25 times greater than the luminosity of the ISR. ISABELLE is expected to cost $275 million. The construction schedule calls for experiments to begin in 1986.

Before ISABELLE is finished even higher center-of-mass energies may have been achieved at Fermilab and at CERN, although the luminosities will

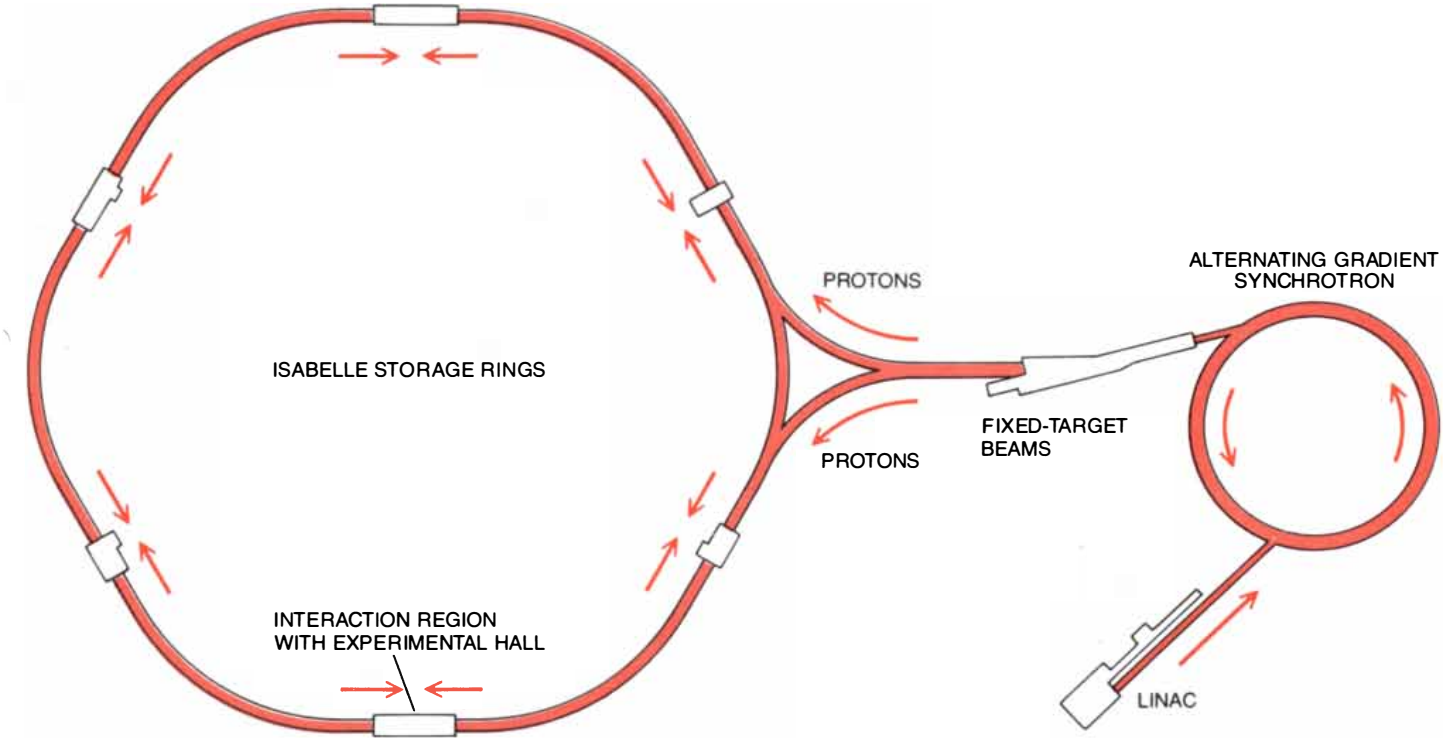
be much lower. The extreme energies would be obtained through improvised schemes for operating the large synchrotrons as storage rings. At Fermilab the presence of two rings in the same tunnel makes possible a variety of colliding-beam arrangements. One plan would be to store counter-rotating beams of protons in the two rings and to bring them into collision at one of the straight sections that are spaced around the rings. The Tevatron could be operated at its full 1-TeV energy as a storage ring, but the maximum steady magnet current would limit the main ring to about 250 GeV. The center-of-mass energy for such collisions is about 1 TeV. By pulsing the main ring up to its full energy a 40 percent improvement in center-of-mass energy could be obtained, although at considerable cost to the rate of interactions.

A proposal that would yield even greater energy and one that, for the present at least, is being seriously pursued, is to employ the Tevatron alone as a single-ring colliding-beam device for protons and antiprotons, operating on the same principle as the electron-positron rings. The difficulty of such a plan is in accumulating a sufficiently intense beam of antiprotons, which are not available in ordinary matter. Antiprotons are formed in collisions of protons with a fixed target, but they emerge with a wide range of speeds and directions and cannot in that state be injected into the narrow aperture of a synchrotron. Viewed in the antiprotons' own frame of reference the rough beam of them resulting from the collisions is a hot gas, the temperature of which must be reduced substantially before the particles can be employed in a colliding-beam experiment.

The prospects for building a proton-antiproton storage ring depend on two recently developed methods for artificially cooling a beam of antiprotons. One method, called electron cooling, was invented by the late Gersh I. Budker and his colleagues at Novosibirsk. A rough beam of antiprotons is first confined in a low-energy storage ring of very large aperture. Electrons are then passed through a long straight section of the ring in such a way that they move parallel to the average path of the antiprotons and at the same average speed. The electrons have a much lower temperature, and as a result of collisions they carry off the randomly directed components of the antiprotons' momentum. In effect the hot gas of antiprotons gives up its heat to the cold gas of electrons. At the end of the straight section the two kinds of particles are separated by a magnetic field. The electrons strike a collector, but the antiprotons continue to circulate and pass through the cooling region repeatedly.

Electron cooling was demonstrated by Budker, and larger cooling rings have been set up both at CERN and at Fermilab. The CERN workers were the first to repeat the Novosibirsk results; they found the rate of cooling was even faster than had been expected. The Fermilab plan calls for injecting both protons and cooled antiprotons into the Tevatron at an energy of about 150 GeV. A rather complicated procedure of shuttling particles among the main ring, the Tevatron, the booster and the cooling ring will be required in order to inject both beams. The particles and antiparticles will then be accelerated simultaneously to any energy up to the maximum of 1 TeV. Thus the center-of-mass energy can reach 2 TeV.

At CERN another method of antiproton cooling is being tested. Invented by Simon Van der Meer of CERN, it is called stochastic cooling. The antiprotons are again stored in a small ring of large aperture, this one equipped with a system of electronic sensors and a special orbit-correcting magnet. On each revolution the sensors detect the average position of the particles at some point along the ring. If the center of charge of a fluctuation of the particles in the beam has strayed from the axis of the vacuum chamber, a correction signal is calculated and dispatched to the correction magnet. (Because the antiprotons move at nearly the speed of light the signal can reach the magnet before the particles do only by taking a short cut across a chord of the circle.) The correction magnet restores the center of charge to its proper position. In the process some particles may well be



PROTON-PROTON COLLISIONS are the purpose of ISABELLE, a device now being built at the Brookhaven National Laboratory. The protons will be supplied by the Alternating Gradient Synchrotron, a 33-GeV accelerator built in 1961. They will be injected into interlaced rings made up of superconducting magnets and brought to 400 GeV per beam; hence center-of-mass energy will be 800 GeV. Particular care is being taken to ensure high luminosity, 10 times that of any previous storage ring. ISABELLE will be completed in 1986.

further deflected from their proper trajectories rather than returned to them, but the majority always move in the proper direction. The system of sensors and magnets acts like Maxwell's demon, cooling the beam by deft manipulations of the distribution of particle velocities.

At CERN the protons and antiprotons will be stored in the SPS. Because of limits on the steady-state current in the conventional magnets, the maximum sustained energy will be 270 GeV per beam. That still amounts to 540 GeV in the center-of-mass system. By pulsing the magnets a center-of-mass energy of 1,000 GeV could be reached, but at a much lower average rate of interactions.

The parallel efforts of CERN and Fermilab to establish proton-antiproton collisions will undoubtedly be seen as another race, and with some justice. One prize may well be the first observation of the weakon. CERN has committed more of its resources to the project and for now has a considerable lead, but the estimated completion dates, early in 1981 for CERN and one or two years later for Fermilab, are close enough for factors unforeseen to determine the outcome. In any case it must also be said that a spirit of generous cooperation prevails between the two laboratories. Both of the projects were largely inspired by the brilliant conceptions of the same man: Carlo Rubbia of Harvard University.

Neither of these makeshift storage rings is expected to have a high luminosity. The facilities for experimental apparatus will also be less than ideal. Furthermore, there will be fierce competition for running time with the very attractive fixed-target programs, which are incompatible with the colliding-beam experiments. The one irresistible allure of the beam-cooling projects is the opportunity to reach extremely high energy quickly and cheaply. What the CERN and Fermilab colliders will provide is a glimpse from Mount Pisgah of the promised land. Will it be ISABELLE and LEP that will actually take us there?

ACCELERATOR	DATE OF FIRST OPERATION	PARTICLES ACCELERATED	BEAM ENERGY (Gev)	CENTER-OF-MASS ENERGY (GeV)	INTENSITY (PARTICLES PER PULSE)
PS CERN, Geneva	1959	Protons	28	7.4	
AGS Brookhaven National Laboratory, New York	1961	Protons	33	8	8×10^{12}
SLAC (linear accelerator) Stanford University	1961	Electrons	22	6.5	
Cornell Electron Synchrontron Cornell University	1967	Electrons	12	4.9	
Serpukhov Proton Synchrotron Serpukhov, U.S.S.R.	1967	Protons	76	12	5×10^{12}
Fermilab main ring Batavia, Ill.	1972	Protons	500	30.7	2×10^{13}
DESY Hamburg	1974	Electrons	7	3.8	
KEK Japan	1975	Protons	12	5	
SPS CERN, Geneva	1976	Protons	500	30.7	10^{13}
Tevatron Fermilab	1982	Protons	1,000	43	5×10^{13}
Beijing Proton Synchrotron Beijing	1985	Protons	50	9.8	10^{13}
UNK Serpukhov, U.S.S.R.	Late 1980's	Protons	3,000	75	
Pentevac Fermilab	Un-scheduled	Protons	5,000	97	10^{14}
VBA No site selected	Un-scheduled	Protons	20,000	137 to 194	10^{13}–10^{15}

INVENTORY OF LARGE ACCELERATORS includes one linac (at SLAC); the rest are synchrotrons. Among the projects that are not completed, the Tevatron at Fermilab and the Beijing Proton Synchrotron in China are under construction. UNK, a planned 3-TeV proton synchrotron to be built at Serpukhov, near Moscow, has not been given final authorization. The Pentevac is a design exercise, describing the most powerful accelerator that could be built on the Fermilab site. The VBA (Very Big Accelerator) is under discussion as the successor to the accelerators and storage rings planned or being built now. Its maximum energy and its cost would be an order of magnitude greater than those of the largest accelerators operating today. Building the VBA would probably require pooled resources of U.S., U.S.S.R. and other nations.

The impression should not be given that the only interesting high-energy-physics laboratories are in the U.S. and Western Europe. In Japan a 12-GeV proton synchrotron, KEK, has been operating for several years and a 200-GeV synchrotron, TRISTAN, is planned. The Chinese have announced plans for an ambitious complex of accelerators to be built on a site 40 miles northwest of Beijing. Construction has begun on a 50-GeV proton synchrotron, which is expected to be completed by about 1985. The lovely site, which has a view of the Ming tombs and is near the Great Wall, is large enough to accommodate a 1.5-TeV superconducting proton synchrotron for which the 50-GeV machine could serve as an injector.

The scope of the accelerator laboratories in the U.S.S.R. is comparable to that of the American and European programs. A series of four electron-positron colliding-beam devices has been built at Novosibirsk, the latest one having a beam energy of 7 GeV. There is a 6-GeV electron synchrotron at Yerevan in Armenia, and there are small proton synchrotrons at Moscow (7 GeV) and at the Joint Institute for Nuclear Research in Dubna (10 GeV). The Serpukhov Institute for Nuclear Physics, 60 miles south of Moscow, has an accelerator that for several years was the most powerful in the world: it is a 76-GeV proton synchrotron. When a proposed expansion of the facilities there is completed, the Serpukhov laboratory may reclaim that distinction.

In the new project, which is called UNK, protons will be transferred from the 76-GeV accelerator to a new site six kilometers away. There the protons will be injected into a 400-GeV synchrotron, which will serve as the booster for a third synchrotron, to be built in the same tunnel, with a maximum energy of 3 TeV. The booster will be of conventional design, but the 3-TeV ring will employ superconducting magnets. The magnets are now being developed at Serpukhov and at the Saclay Nuclear Research Center outside Paris.

Multiple modes of operation are envi-

sioned. The 3-TeV protons and secondary beams derived from them will be available for fixed-target experiments. Collisions between the 3-TeV protons and the 400-GeV protons from the booster will also be possible, yielding a center-of-mass energy of somewhat more than 1 TeV. The possibility has also been discussed of adding a 20-GeV electron synchrotron, which would provide for electron-proton collisions. Beam cooling would supply antiprotons for collisions with a center-of-mass energy of 6 TeV. If a second large ring were constructed, the same energy could be attained in proton-proton collisions. Plans and specifications for UNK are now being drawn up; the machine will take some seven years to complete.

A dream that has not yet even begun to make its way through the bureaucracy is the Site-Filler or Pentevac accelerator that could be built at Fermilab. The idea began with the question: What is the most powerful accelerator that could be fit into the Fermilab site? The land available will accommodate an accelerator no larger than five kilometers in diameter, or 17 kilometers in circumference. The energy that could be developed in a proton synchrotron of that size depends strongly on future developments in the technology of superconducting magnets. The maximum field strength possible now in a magnet suitable for an accelerator is about 50 kilogauss. It seems reasonable to assume that in about 10 years with improved materials and improved methods of fabrication the limit will be raised to 85 kilogauss (although eventually much higher fields will be attained). In that case the accelerator could reach 5 TeV. A varied menu of experiments would then be made available. Included on the list would be fixed-target beams, proton-proton collisions at energies of up to 4.5 TeV and proton-antiproton collisions at 10 TeV. An electron accelerator could be added, further increasing the combinatorial richness of the laboratory.

It may be possible to defend the thesis that the world needs two synchrotrons as similar as the Fermilab main ring and the SPS, or two storage rings as similar as PEP and PETRA. It is certainly true that none of those instruments will sit idle for lack of worthy experiments to carry out. For the largest accelerators now contemplated, however, any possible benefits of duplication become moot. The worldwide budget for high-energy physics is not likely to support more than one proton synchrotron with a beam energy of 20 or 30 TeV. If such a machine is ever to be built, it cannot be as a national project, nor is it likely to be the undertaking of a regional consortium such as CERN. Worldwide sponsorship would be appropriate to such an endeavor; it might be necessary, and a dream of many physicists is that it would be possible.

With a curious blend of pragmatism and idealism physicists have long advocated and practiced international collaboration. CERN is an exemplary model, but there are other robust international endeavors, including some with a broader constituency. In 1961, as a result of the Atoms for Peace movement, an exchange of particle physicists was worked out between the U.S. and the U.S.S.R. for the discussion of undertakings in the mutual interest. The possibility of a "World Accelerator," to be supported by the pooled resources of many nations, was one of the topics considered. Those talks came to an end with the episode of the *U*-2 flight over the U.S.S.R., but the idea did not die.

Enthusiasm for the World Accelerator surfaced again in 1975, at an international meeting convened in New Orleans by Victor F. Weisskopf of the Massachusetts Institute of Technology. What was a strong sentiment at the New Orleans meeting has now been embodied in an official agency, the International Committee for Future Accelerators, which operates under the aegis of the International Union of Pure and Applied Physics. The committee is charged with coordinating the plans of the various national and regional laboratories and with encouraging collaboration among them. Many physicists think such joint ventures are already working well, and are suspicious of official meddling.

A second function of the committee has been received more enthusiastically: it is to provide a forum for the discussion of what is now called the VBA, for Very Big Accelerator. To that end two international workshops have been held, one at Fermilab in 1978 and the second at CERN a few months ago. Both meetings have been concerned largely with establishing the possibilities and limits of accelerator and particle-detector technology, but of course the discussions were held in the context of the physical phenomena to be expected at VBA energies.

It appears that building a 20-TeV proton synchrotron is feasible and would require no major departures from present practice. A large share of the increase in energy would be obtained by the simplest of strategies: increasing the radius of the ring. Very-high-energy colliding-beam devices, on the other hand, might benefit from new techniques. For electron-positron collisions, for example, energies beyond the LEP range might be reached most efficiently by a pair of linacs aimed muzzle to muzzle, or by a combination of linacs and electron synchrotrons. A means for improving the luminosity of colliding beams without accelerating more particles has been suggested; it consists in raising the particle density of the beams by squeezing them down to a microscopic cross section at the point of interaction. That presents a new phenomenon that was recognized by the first workshop sponsored by the International Committee for Future Accelerators: in a small beam the magnetic field becomes large, so that synchrotron radiation in the magnetic field itself becomes an important energy loss. This phenomenon is known as "beamstrahlung," a pun on bremsstrahlung, or braking radiation, the process in which charged particles emit electromagnetic radiation on being decelerated. Burton Richter and his colleagues at SLAC are seriously proposing that colliding 50-GeV linac beams could be a Z^0 factory—a challenge to the Cornell plans.

One thing that most interested physicists agree on is that a world effort should not be resorted to until it is clearly recognized as the only way to achieve the construction of a particular accelerator. It is significant that when serious discussion of a World Accelerator began in 1960, it was a 100-GeV proton synchrotron that was being considered. Now, because of technological advances, 10 times that energy can be achieved by a national effort alone. At the time the International Committee for Future Accelerators was organized a proton synchrotron of 10 TeV and an electron-positron storage ring with colliding beams of 100 GeV seemed appropriate projects for a worldwide effort. The proposed construction of LEP at CERN has preempted the latter possibility. The workshops sponsored by the committee arrived at the conclusion that it would be possible to build electron storage rings with a circumference of about 100 kilometers and about twice the energy of LEP. Above that energy colliding linac beams with energies of up to 350 GeV appear to be more feasible. For protons a synchrotron with a circumference of between 50 and 100 kilometers, with energies in the TeV range and with all the options of antiproton-proton collisions, would also be a viable candidate for the Very Big Accelerator.

If it should become necessary to build such large accelerators in order to study the inner structure of leptons and quarks, and it probably will, then the International Committee for Future Accelerators is providing the necessary international foundation on which such an effort can be built. One is reminded of a fanciful suggestion made 30 years ago by Enrico Fermi: to build an accelerator that would encircle the world. By that standard an accelerator such as the VBA would not be very big, but the common effort needed to build such a machine might nonetheless help to bind the world together.

DEVICE	DATE OF FIRST OPERATION	TYPE	PARTICLES STORED	ENERGY PER BEAM (GeV)	CENTER-OF-MASS ENERGY (GeV)	LUMINOSITY (CM^{-2} SEC^{-1})
ADA (dismantled) National Laboratories of C.N.E.N., Frascati	1963	Single ring	e^+e	.25	.5	
Princeton-Stanford Storage Rings (dismantled) Stanford University	1964	Tangent rings	e^-e^-	.56	1.1	10^{27}
VEPP-2 Novosibirsk, U.S.S.R.	1965	Tangent rings	e^+e^-	.04	.08	
ACO Orsay Laboratories, Paris	1966	Single ring	e^+e	.5	1	
ADONE National Laboratories of C.N.E.N., Frascati	1965	Single ring	e^+e^-	1.5	3	
ISR CERN, Geneva	1971	Interlaced rings	pp	31	62	4×10^{31}
CEA-Bypass (dismantled) Cambridge, Mass.	1967	Single ring	e^+e^-	3.5	7	2×10^{28}
SPEAR Stanford Linear Accelerator Center	1972	Single ring	e^+e^-	4.2	8.4	10^{32}
DORIS DESY, Hamburg	1974	Single ring	e^+e^-	4.5	9	10^{32}
VEPP-2M Novosibirsk, U.S.S.R.	1975	Single ring	e^+e^-	1.3	2.6	
DCI Orsay Laboratories, Paris	1975	Interlaced rings	e^+e^-	3.7	7.4	10^{32}
VEPP-3 Novosibirsk, U.S.S.R.	1977	Single ring	e^+e^-	3	6	10^{30}
VEPP-4 Novosibirsk, U.S.S.R.	1978	Single ring	e^+e^-	7	14	10^{31}
PETRA DESY, Hamburg	1978	Single ring	e^+e^-	19	38	10^{32}
CESR Cornell University	1979	Single ring	e^+e^-	8	16	10^{32}
ISR $p\bar{p}$ CERN, Geneva	1980	Interlaced rings	$p\bar{p}$	31	62	10^{29}
PEP Stanford Linear Accelerator Center	1980	Single ring	e^+e^-	18	36	10^{32}
SPS $p\bar{p}$ CERN, Geneva	1981	Single ring	$p\bar{p}$	270	540	10^{30}
Fermilab $p\bar{p}$ Batavia, Ill.	1982	Single ring	$p\bar{p}$	1,000	2,000	10^{30}
VAPP Novosibirsk, U.S.S.R.	?	Single ring	$p\bar{p}$	23	46	
ISABELLE Brookhaven National Laboratory, New York	1986	Interlaced rings	pp	400	800	10^{33}
LEP CERN, Geneva	Late 1980's	Single ring	e^+e^-	86	172	10^{32}
UNK Serpukhov, U.S.S.R.	Late 1980's	Single ring	$p\bar{p}$	3,000	6,000	
VBA No site selected	Un-scheduled	Single ring	$p\bar{p}$	20,000	40,000	10^{32}

PARTICLE STORAGE RINGS will have a major role in the high-energy-physics program of the next two decades. Two new electron-positron rings have recently begun operating; a third such machine will soon be finished. They are CESR, an 8-GeV ring at Cornell University, and PETRA and PEP, two devices with almost identical specifications, respectively at Hamburg and SLAC. About a dozen more colliding-beam devices are planned, including numerous proposals to operate synchrotrons in a storage-ring mode. The largest project for which a detailed design is available is LEP, the Large Electron-Positron storage ring to be built at CERN. The initial phase of LEP would have an energy of 86 GeV per beam, but that could be raised to 130 GeV with superconducting radio-frequency cavities.

The Cell Biology of Human Aging

When normal human cells grow in laboratory glassware, they eventually lose the ability to function and divide. The timing of this loss may well represent the limit of the human life span

by Leonard Hayflick

The human life span has not increased significantly for millenniums: for most people the biblical fourscore years and ten is still the upper limit. What has changed, at least in the developed countries, is the average expectation of life. Improvements in the conditions of life and advances in medicine, particularly in the control of infectious diseases, have simply allowed more people to reach the limit of what appears to be a fixed life span. Even if all the major causes of death today were eliminated, the length of human life would still be fixed at 90 or 100 years. The reason is that although death may result from cardiovascular disease or cancer, those conditions arise from the normal age-related degeneration of the arterial walls or of the immune system. The fundamental cause of death in the later years is the body's increased vulnerability to disease or accident attributable to the inexorable decline in its peak functional capacity from age 30 on. Bernard L. Strehler of the University of Southern California has estimated that the loss of functional capacity after age 30 is about .8 percent per year. Moreover, in 1825 the English actuary Benjamin Gompertz discovered that the likelihood of dying doubles every eight years after about age 30. What causes this natural decline? Is it possible that it could be halted or even reversed?

With the exception of certain types of cells (notably nerve cells and some muscle cells), the organism represented by an individual human being is not the same organism it was a few years earlier. Many of its cells have died and have been replaced by their descendants. The new cells do not, however, make the organism younger. Aging is expressed not in individual cells but in cell lineages. Consider the oldest living things: bristlecone pines and redwood trees. The living cells in those trees are no more than 30 years old; the great bulk of the tree is made up mostly of dead cells that are not necessary for survival. Since dead cells should not be included in determining age, by this reasoning the trees are no more than 30 years old: considerably younger than the oldest nerve cells in many human beings.

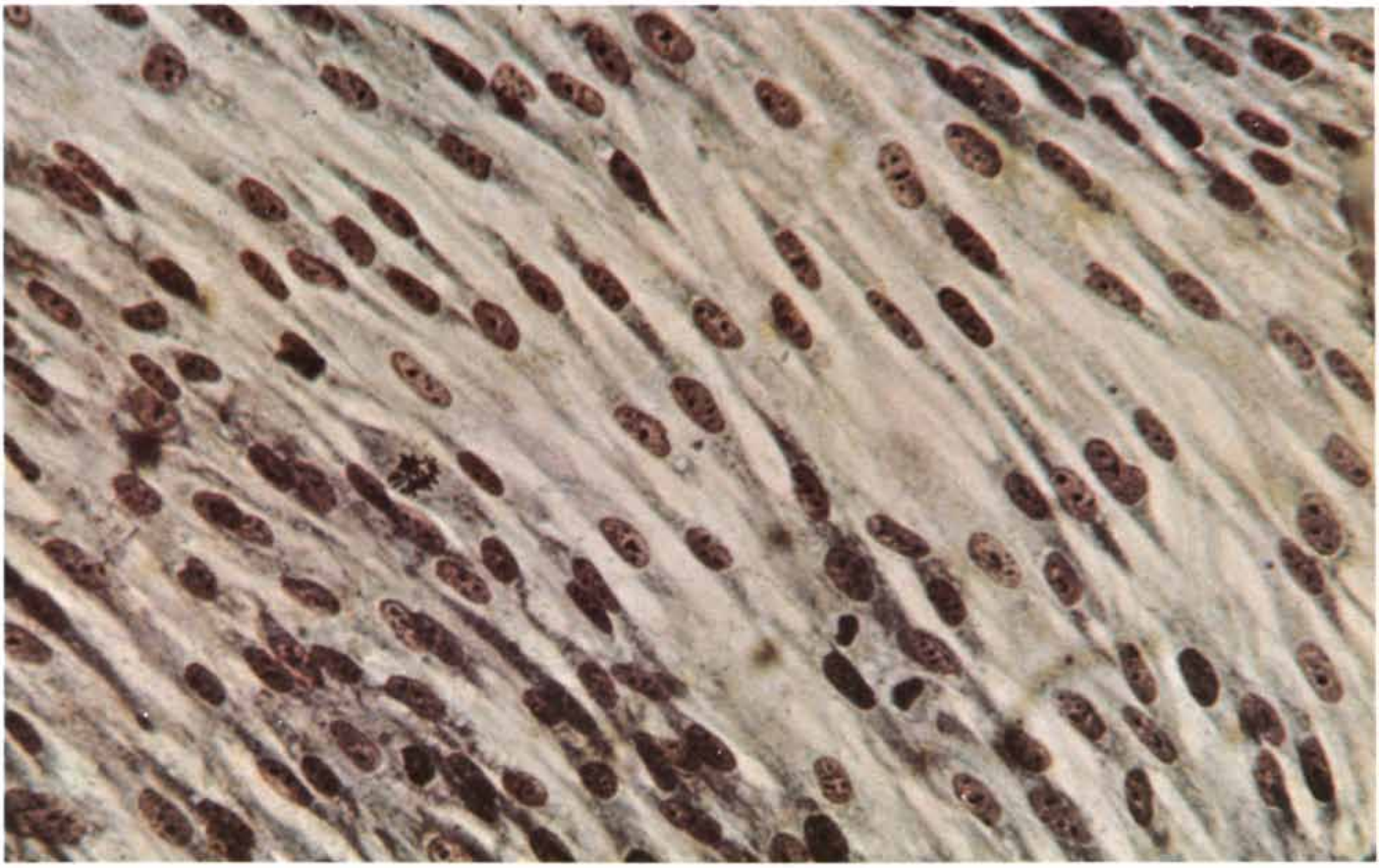
In my laboratory at the Children's Hospital Medical Center of Northern California we have been examining the process of aging at the cell level. When normal human fibroblasts (the structural cells of the body's soft tissues) are grown in laboratory glassware, over a period of months they divide many times and then slowly stop dividing and eventually die. Such findings suggest that the aging of a normal cell lineage is an innate property of cells.

Our cell-culture technique is a fairly simple one. We first expose small bits of embryonic lung tissue to digestive enzymes, which dissociate the tissue into millions of individual cells. The cells are separated from the enzymes by being spun in a centrifuge, are introduced into bottles containing a nutrient solution and are incubated at body temperature. The cells then proceed to divide regularly. After a week the population of cells attains a state of confluence: it covers the entire surface of the culture in a layer one cell thick. Once confluence has been reached normal cells are prevented from dividing further by the phenomenon of contact inhibition.

If additional cell divisions are wanted, the cells must be removed from the "mother culture" bottle and distributed in equal numbers to two other culture bottles, along with fresh nutrient medium. This procedure, called subcultivation, results in an approximate doubling of the number of cells removed from the mother culture by the time the daughter cultures reach confluence and stop dividing. (The fact that the population of cells doubles does not mean that each cell divides once; some cells do not divide at all and others divide more than once.) If four daughter cultures are prepared from a single mother culture, the total number of cells will double twice, although a longer incubation period is needed. Regardless of the dilution, however, the ability of normal human embryo fibroblasts to divide is limited to 50 population doublings over a period of between seven and nine months.

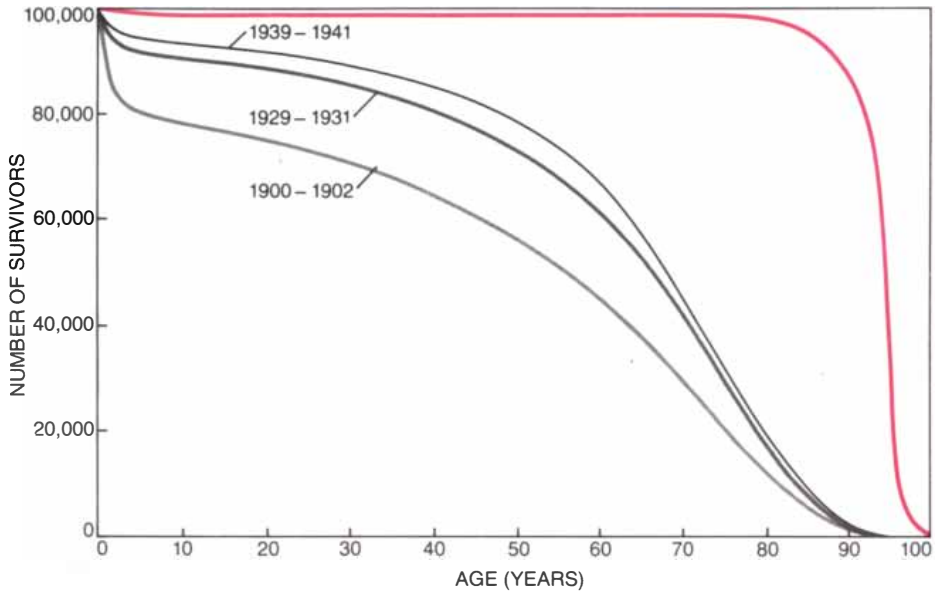
When normal fibroblasts reach the end of their ability to proliferate, they do not stop dividing suddenly. Rather the time necessary for the daughter cultures to reach confluence increases as the population approaches the limit of 50 doublings. The first sign of approaching senescence is the need for 10 days instead of a week for the culture to reach confluence. The ultimate result is a subcultivation in which confluence is never reached, regardless of how long the culture is incubated or how often the nutrient medium is changed. Finally the cells undergo a variety of degenerative changes and die.

At first we attributed the death of the cultures to experimental error; the division of cells in culture is known to be adversely affected by faulty preparation of the culture medium, by inadequate washing of the glassware, by contamination with viruses and by other technical oversights. Intensive studies suggested otherwise. The experiment that came closest to proving that the death of cultured normal cells is an innate cell phenomenon was done by Paul S. Moorhead and me. Taking advantage of the fact that normal human fibroblast strains derived from male and female embryos can be distinguished by comparing their sex chromosomes, we mixed female cells that were 10 population doublings old with an equal number of male cells that were 40 population doublings old. Unmixed cultures of each strain were retained as controls, and all three cultures were subcultivated in the usual way. After 20 population doublings we looked at the mixed culture and found only female cells present. The female control cells were also dividing but the male control cells had died. This result strongly suggested that the death of the older male cells was due to some innate property of the cells



"YOUNG" FIBROBLASTS (connective-tissue cells) obtained from human fetal tissue carpet the surface of a culture dish in this photomicrograph. The spindle-shaped cells were fixed and stained during a period of active proliferation. Reddish objects are cell nuclei.

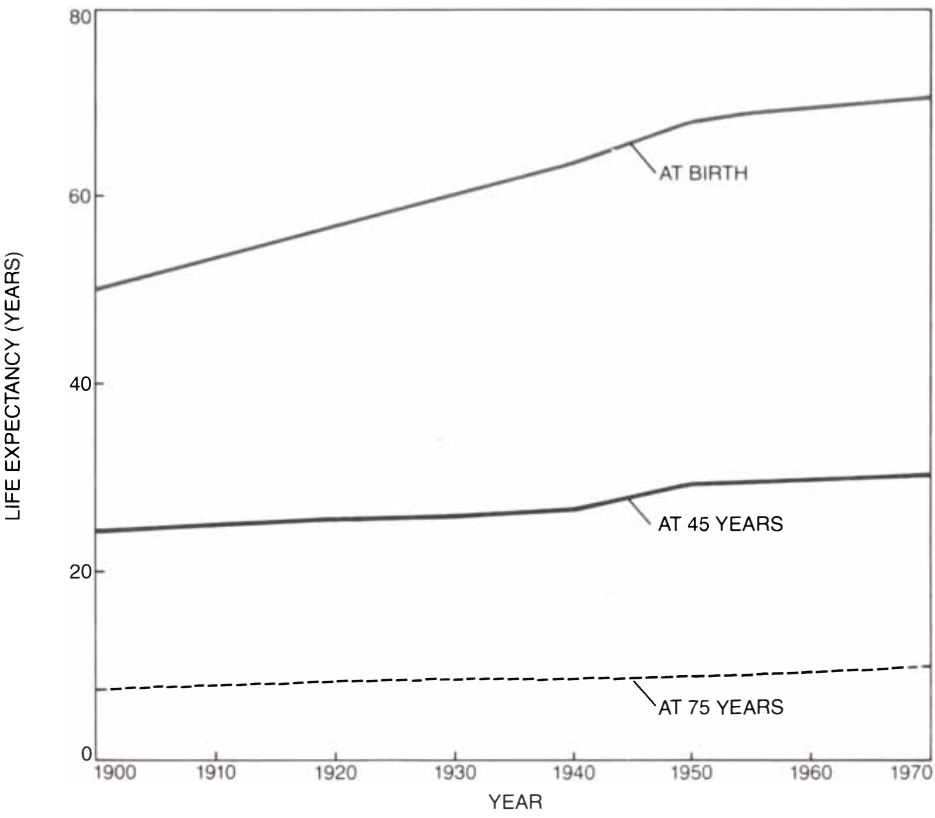
SENESCENT FIBROBLASTS were stained during the loss of division capacity that follows approximately 50 population doublings in culture. The cells undergo a variety of degenerative changes and then die. This observation suggests that aging is an innate property of normal living cells. Micrographs were made in the author's laboratory at the Children's Hospital Medical Center of Northern California.



HUMAN SURVIVAL CURVES for whites at three periods in U.S. history show that with the improvement of the conditions of life and advances in medicine fewer deaths occur in the earlier years of life. The elimination of all causes of death now attributed to disease and accident would result in hypothetical ultimate curve (*color*); individuals then would die only from the physiological deficits that lead to normal age changes. Data were collected by Alex Comfort.

themselves rather than to some technical error, since otherwise it would have been difficult to explain why only the female cells in the mixed culture survived.

The fact that normal fibroblast strains have a finite lifetime was also suggested by another observation. When fibroblasts derived from human embryos are stored at subzero temperatures in liquid nitrogen, they exhibit a remarkable "memory." If they are frozen at the 20th population doubling and later thawed, they will undergo 30 more doublings and then stop. If they are frozen at the 10th population doubling and thawed, they will undergo 40 more doublings. In other words, when the doublings before and after freezing are summed, the total number of population doublings is always 50. It is as if the cells have a built-in clock that counts divisions. One of our normal human-cell strains, designated *WI*-38, retained this memory after more than 13 years of preservation in liquid nitrogen!



LIFE EXPECTANCY for U.S. whites at birth and at ages 45 and 75 is depicted from 1900. Gain in life expectancy of about 24 years since 1900 has been mostly in younger age groups.

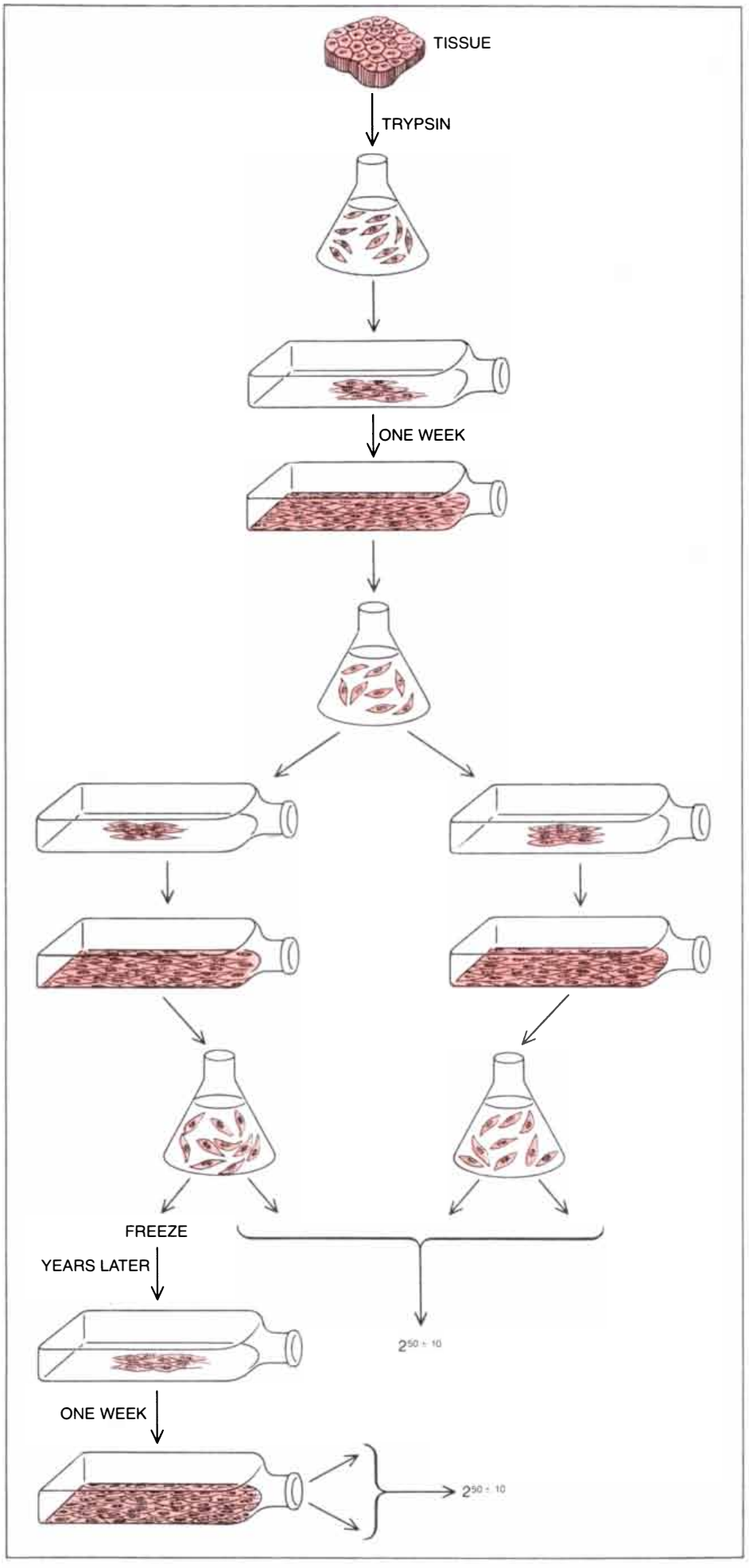
The finite lifetime of cultured human cells has now been observed in fibroblasts and in other normal cell types derived from many tissues, including the skin, the brain, the liver and smooth muscle. No exception has yet been found to the general rule that normal cells possess a limited capacity to divide. Cultured human-cell strains can, however, be transformed into an "immortal" cell line by being treated with the cancer-causing monkey virus SV40. About 600 immortal or transformed cell lines are known, many of which arose spontaneously in animal-cell cultures. The most famous human-cell line, called HeLa, was derived in 1952 from cultured tissue from the uterine cervix. It has been in continuous culture since then. The transformed cells are clearly abnormal in such characteristics as the number and shape of their chromosomes, their chemical properties and their staining characteristics. Moreover, most of the transformed cells give rise to tumors when they are injected into laboratory animals, whereas normal cells do not. One is therefore led to the paradoxical conclusion that for animal cells to achieve the potential for unlimited cell division they must acquire some or all of the properties of cancer cells.

Several subsequent discoveries suggested to us that the finite lifetime of cultured normal cells might represent a form of aging at the cellular level. We cultured fibroblasts from older human donors to determine whether older cells undergo fewer divisions in culture. Since our initial studies had been done with fibroblasts from the lung tissue of human embryos, we used lung fibroblasts obtained at autopsy from eight adults ranging in age from 20 to 87. We found that the older fibroblast populations doubled between 14 and 29 times in culture, with no clear correlation between the number of doublings and the age of the donor. Nevertheless, there was a significant reduction of the consistent 50 population doublings observed for human-embryo fibroblasts. More recently studies on cultured fibroblasts obtained from human skin biopsies by George M. Martin and his colleagues at the University of Washington School of Medicine, by Samuel Goldstein of McMaster University in Canada and by Edward Schneider and Youji Mitsui of the Gerontology Research Center in

Baltimore have suggested that the number of population doublings undergone by normal human fibroblasts in culture is inversely proportional to the age of the donor. Similar results have been obtained with cells derived from human liver by Y. Lebuilly in France and with muscle cells by Edwin L. Bierman of the University of Washington School of Medicine.

Aging is a property of all animals that reach a fixed size at maturity. (Some fishes continue to grow throughout their lifetime and do not age in the usual sense.) It is therefore clearly of interest to know whether the normal cells of animals other than man also become senescent in culture. The variations in life span among different species are far greater than those among individuals of the same species. A fruit fly is ancient at 40 days, a mouse at three years, a horse at 30, a man at 100 and some tortoises at 150. Although only a few species have been examined, the available evidence suggests that there may be a direct relation between the life span of a species and the capacity of its cells to divide in culture. Another interesting relation has been pointed out by George A. Sacher, Jr., of the Argonne National Laboratory. He finds a strong positive correlation between life span and the ratio of brain weight to body weight in 85 animal species ranging in size from the mouse to the elephant. This finding indicates that over the span of the evolution of vertebrate animals there may have been an important link between the increase of life span and the increase of brain size.

A second approach to the study of aging at the cellular level has been to transplant marked cells serially into inbred animals before the animal is old, thereby circumventing the death of the marked cells when the host animal dies. These in vivo experiments have complemented the tissue-culture ones. Results from several laboratories in which mammary tissue, skin and immature



CULTURE of normal human cells in the author's laboratory begins with the breakdown of tissue into individual cells with the digestive enzyme trypsin. The cells are then transferred to a flat bottle, where they multiply until they cover the surface of the culture vessel. Then they stop multiplying. The cell population is now divided into equal halves, which are recultured. This process can be repeated only about 50 times with human fetal cells; it cannot be repeated as often with cells from older donors. Cells grown in this way are stored by being frozen in liquid nitrogen. Thawed and cultured years later, the cells "remember" the population-doubling level at which they were frozen and continue to divide until the total number of possible doublings is reached.

blood cells were serially transplanted into rats indicate that the transplanted cells age and die as they do in culture. In some cases, however, the total calendar time that elapsed before the death of those cells was greater than the maximum life span of an animal of that species. This finding can be explained by the fact that the transplanted cells do not grow as rapidly as they would in culture but most of the time remain in a quiescent state. If they did grow as fast as the cultured cells, the transplant would in a few months reach a size many times that of the host.

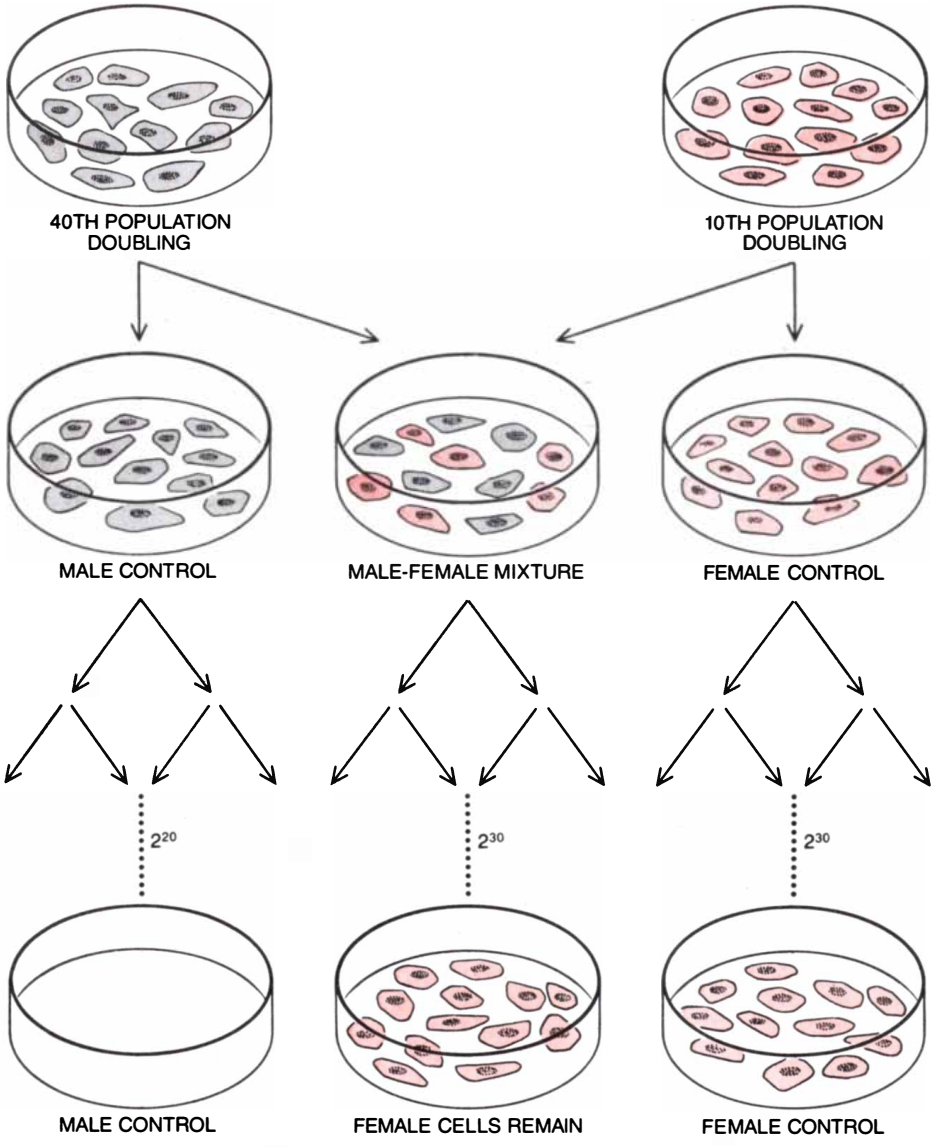
A similar phenomenon is observed when cultured normal cells are grown not at body temperature but at room temperature: they divide slowly and degenerate many months after the death of sister cultures grown at body temperature. Although calendar time may be extended this way, the total number of population doublings does not exceed that of the cells incubated at body temperature. In contrast, like the immortal transformed cell lines grown in tissue culture, cancer cells replicate indefinitely when they are transplanted from one inbred animal to another.

In an effort to locate and understand the mechanism controlling the finite replicative capacity of normal cells, Woodring Wright and I took a different approach. When cultured cells are treated with the drug cytochalasin *B,* they extrude their nucleus. The nuclei can be separated from the cells by centrifugation, so that it is possible to obtain millions of cells lacking a nucleus. Such cells are called cytoplasts, and they remain viable for days. In that time they can be fused with normal cells. In order to determine whether the clock that dictates a cell's replicative capacity is in the nucleus or outside it in the cytoplasm, we fused cytoplasts derived from young cells to normal old cells and vice versa. We found that the cytoplasm seemed to have little effect on the cell's rate of aging, suggesting the clock is probably in the nucleus. More recently techniques have been developed for inserting cell nuclei into cytoplasts. Audrey Muggleton-Harris and I reconstituted viable normal human cells from isolated young nuclei and old cytoplasts and from old nuclei and young cytoplasts. The results of preliminary experiments seem to indicate that the clock that determines the ability to continue cell division is indeed in the nucleus.

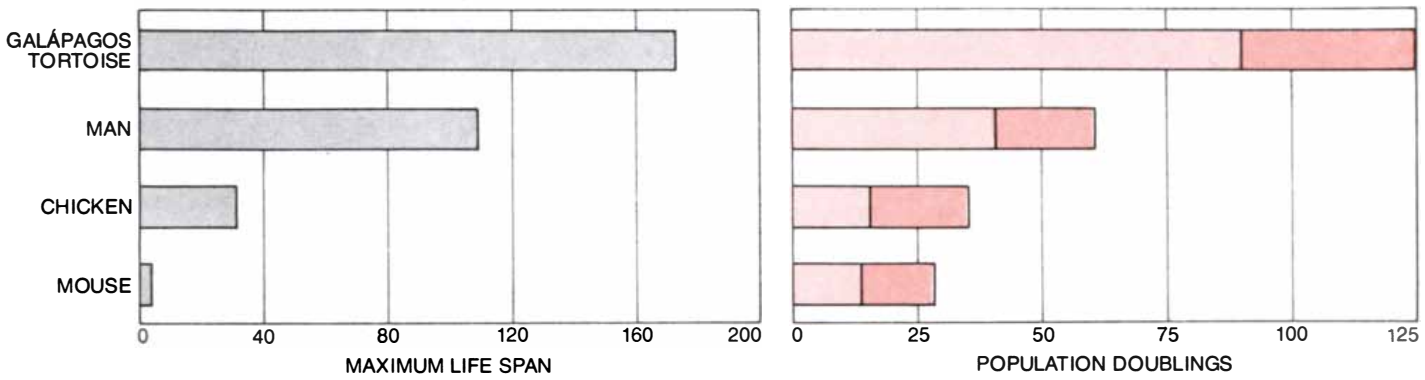
It is also possible to grow small bits of intact animal tissue directly on glass or plastic instead of dissociating them into individual cells. As early as the 1920's it was discovered that if a small scrap of tissue from a chick embryo is placed in culture, the time it takes for the first fibroblasts to migrate out from the edge of the tissue explant increases with the age of the embryo. This length of time is known as the latent period.

In the 1960's Milena Soukupová and Emma Holečková of Charles University in Czechoslovakia found a similar increase in the latent period with respect to the age of rats in explants from the heart, the liver and the kidney. In 1970 Roy L. Walford, Jr., and Harold Waters of the University of California at Los Angeles School of Medicine obtained 20 explants from living human donors ranging in age from a few minutes to 80 years and found that the latent period increased with age (although individual variations in the period tended to increase as well).

Now that there is ample evidence that cultured normal cells are not immortal but have a finite capacity to divide, it is important to assess the impact of this discovery on the understanding of the aging process. It now appears that some kind of clock mechanism exists in normal cells controlling their capacity to function and to reproduce. Not all the cells of the body, however, have the fibroblast's capacity for rapid replication. The tissues in which cells replicate rapidly are the skin, the blood-forming tissues and the lining of the gut. More specialized cells, such as nerve cells, endocrine cells, muscle cells, sensory cells and some cells of the immune system divide little if at all after maturity. Gerontologists tend to agree that the most important age changes occur not in the rapidly dividing cells but in the highly specialized cells. In fact, it seems likely that the animal ages and dies for other



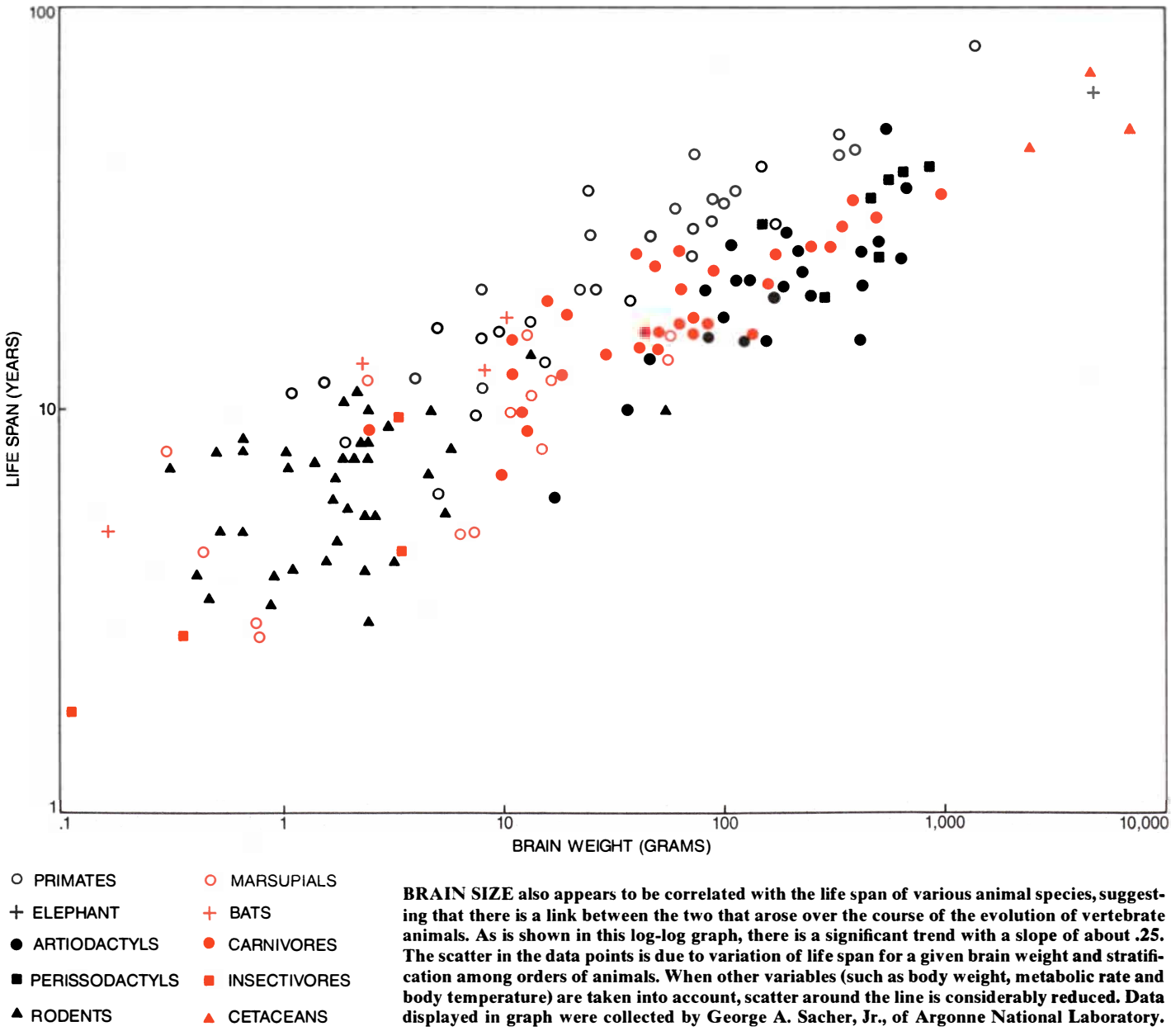
LIMIT OF REPLICATION of normal cells in culture was demonstrated by preparing a mixed culture of female and male fetal cells that were respectively "young" and "old." Unmixed cultures of each served as experimental controls. After 30 population doublings only the female cells remained in the mixed culture and only the female control cells were still dividing. The male cells had exhausted their population-doubling potential and had stopped dividing several weeks before the mixture was examined. It is unlikely that technical errors or contaminating viruses could explain why only the male cells in the mixed culture had died off. Most likely explanation is that loss of ability to divide in normal cells is due to activity of an intracellular clock.



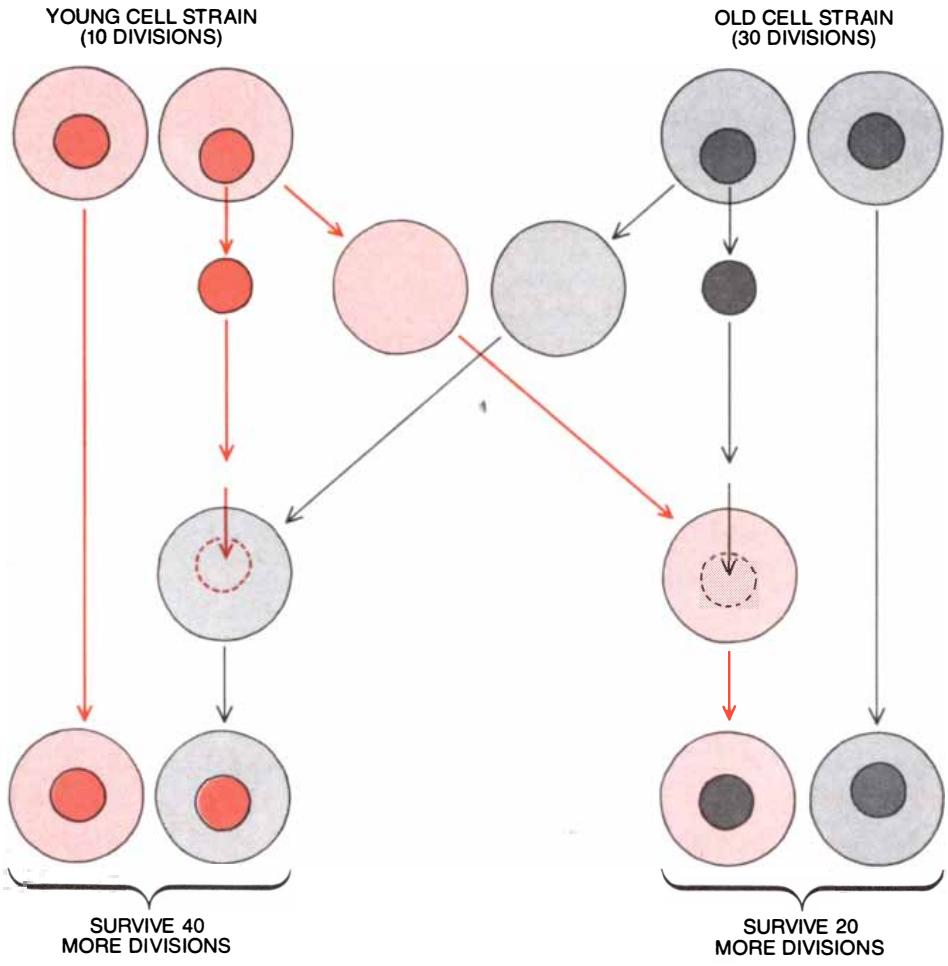
REPLICATIVE ABILITY of fibroblasts from the fetus (from the newborn animal in the case of the tortoise) and grown in culture is proportional to life span of species. Observation suggests that limited replicative ability of cultured normal cells may be correlate of aging.

reasons before the replicative end point of the fibroblasts and other rapidly dividing cells is ever reached.

Nevertheless, the mechanism that limits cell division in rapidly proliferating cells is probably the same one limiting the functional capabilities of specialized cells that divide more slowly if at all. Indeed, over the past decade a wide variety of physiological and morphological changes have been detected in cultured human fibroblasts well before they lose their ability to divide. Among them are changes in the utilization of foodstuff, in the expression and repair of the genetic material, in the metabolic pathways of the cell, in the activity of several crucial enzymes and in the morphology of the cell surface and of the organelles within the cell. Many of these changes are identical



BRAIN SIZE also appears to be correlated with the life span of various animal species, suggesting that there is a link between the two that arose over the course of the evolution of vertebrate animals. As is shown in this log-log graph, there is a significant trend with a slope of about .25. The scatter in the data points is due to variation of life span for a given brain weight and stratification among orders of animals. When other variables (such as body weight, metabolic rate and body temperature) are taken into account, scatter around the line is considerably reduced. Data displayed in graph were collected by George A. Sacher, Jr., of Argonne National Laboratory.



CELL-FUSION EXPERIMENT helped to determine the location of the clock that controls cellular aging. It was made possible by the development of techniques for removing nuclei from cells and reinserting them into other cells. The cells without nuclei, called cytoplasts, can survive for a short time. Here nuclei from young cells (10 population doublings) were inserted into the cytoplasts of old cells (30 population doublings) and vice versa. Preliminary results suggest that the nucleus rather than cytoplasm is the location of the clock that counts cell divisions.

with those that occur in human beings as they age. An understanding of why normal fibroblasts lose their ability to replicate may well provide insights into the age-related loss of other kinds of functional properties in more specialized cells.

What, then, is the underlying mechanism that causes age changes in most cells? Many gerontologists believe the answer can be found in the genetic message. They reason that if all the complexities of development from the fertilized egg to sexual maturation are orchestrated by the genetic apparatus, age changes are likely to be controlled by the genes as well. Three general hypotheses based on the properties of the information-bearing molecules in the cells (DNA and RNA) are now thought to be the most plausible explanations for aging.

The first hypothesis, formulated by Zhores Medvedev of the Medical Research Council in London and further developed by Leslie Orgel of the Salk Institute, proposes that over a period of time the information in the information-processing system represented by the transcription and translation of the genetic message in DNA into RNA and into enzymes and other protein molecules might be increasingly subject to error. Such errors would give rise to faulty enzyme molecules and lead to a decline in the functional abilities of the cell. The situation would be analogous to an error in the instructions of an automatic machine tool: the tool would turn out faulty parts that, when they were assembled into the final product, would reduce its efficiency or keep it from working altogether. Although this "error catastrophe" hypothesis has been tested in several laboratories by tracing protein synthesis in aging cells, the results have not yielded good evidence to support it.

The accumulation of errors in a biological system can be partially offset by known processes of repair, but the repair systems themselves do not perform perfectly or indefinitely. Ronald W. Hart and Richard B. Setlow of the Oak Ridge National Laboratory found that cultured skin fibroblasts from several mammalian species, ranging from the shrew to the elephant and man, are able to repair damage to their DNA caused by irradiation with ultraviolet in direct proportion to the lifetime of the species. For example, human beings live about twice as long as chimpanzees and the rate of DNA repair in human beings is also twice as high. More recently these workers and others have confirmed the previous studies and have also found that the capacity for DNA repair in cultured normal human cells decreases as they approach the limit of their replicative ability.

These processes might be compared to what happens to an automobile or a washing machine. Both machines have what engineers call a "mean time to failure." Without repairs they might function well for a few years, and with perfect repairs they might last indefinitely as long as spare parts are available. The lack of either perfect repairs or spare parts, however, means that the machine ultimately ends up on the scrap heap. Spare-part surgery for human beings is limited to a few organs, and the spare parts are themselves imperfect. Even if brain transplants were possible, the outcome would be futile. Immortality, after all, presupposes the continuity of self-identity, which someone else's brain would not provide.

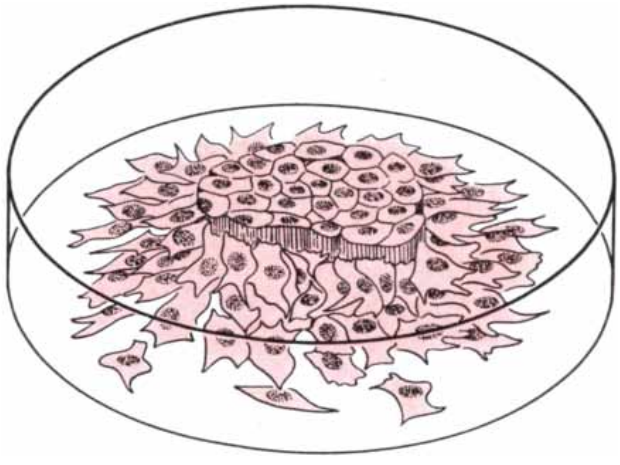
Medvedev has also put forward the second hypothesis about the genetic basis of aging. It is known that only about .4 percent of the information in the DNA of the cell nucleus is utilized by a given cell in its lifetime. Moreover, many of the genes along the DNA molecule are repeated in identical sequences, making the genetic message highly redundant. Medvedev has proposed that the repeated sequences are normally repressed but that if an active gene is extensively damaged, it is replaced by one of the identical reserve genes. The redundancy of the DNA might therefore provide insurance against the system's inherent vulnerability to random molecular accidents, lengthening the time before a sufficient number of errors could accumulate to confound the genetic message. Ultimately, however, all the repeated genes would be used up, errors would accumulate and physiological deficiencies leading to age changes would arise. This hypothesis would therefore predict that long-lived species should have DNA more redundant than the DNA of short-lived species.

The third genetic hypothesis of aging proposes that age changes are simply a continuation of the normal genetic signals regulating the development of an animal from the moment of its conception until its sexual maturation. There may even be "aging genes" that slow or

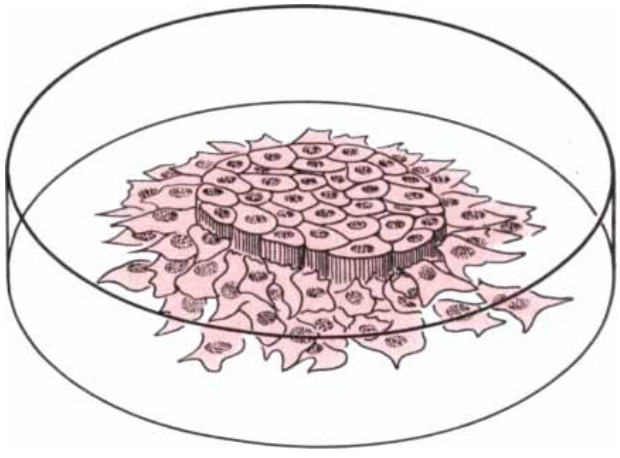
shut down biochemical pathways in a sequential manner and lead to the predictable expression of what are recognized as age changes. Graying of the hair, menopause and diminished athletic ability are a few of the hundreds of late-developing events associated with aging. None is regarded as a disease, but such normal events decrease cell function and increase vulnerability to disease. These genetically programmed events may differ in the time of their expression in different cell types. The root of aging might therefore result from deficiencies in a few key cell types whose rate of aging is the fastest and has the greatest effects.

The function of the hypothetical aging genes might be analogous to the normal functional decline and death of cells that take place on a massive scale during the development of the embryo. In vertebrate animals the modeling of a limb, for example, calls not only for the marshaling of millions of cells but also for the death and resorption of millions of others. The fate of these latter cells is determined by "death clocks" that operate on a precise schedule. It is conceivable that the same processes continue throughout life, operating at different rates in different tissues and ultimately leading to the normal age changes that increase susceptibility to disease. Advocates of such a predetermined genetic program for aging argue that the evolutionary success of a species depends only on the ability of its members to live long enough to procreate and raise their young; what happens after that is irrelevant to the survival of the species.

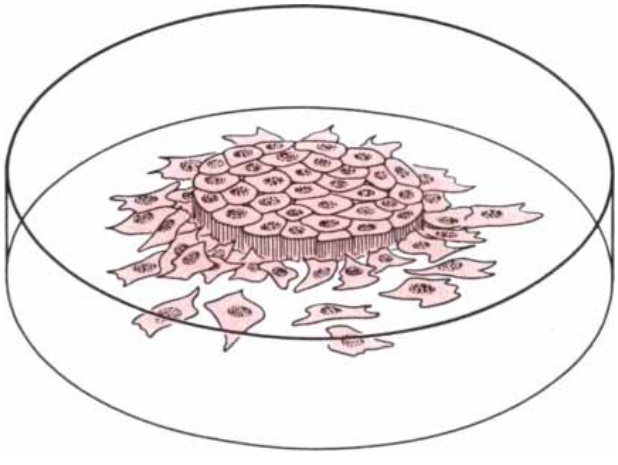
All three major hypotheses of aging are applicable equally to the demise of normal animal cells in tissue culture and to the aging of intact animals. They are not mutually exclusive and might well operate simultaneously. Two animal-cell lineages, however, are known to escape from the inevitability of aging or death. These are cancer cells and the germ cells: the egg and the sperm and their immediate precursors. It is interesting to speculate that cancer cells and germ cells might avoid senescence through a common mechanism. Indeed, the way genetic information is exchanged between cancer-causing viruses and animal cells might be similar to the way the genetic cards are reshuffled when the sperm fuses with the egg. The reshuffling of genetic information between two adjacent normal body cells (mediated by viruses, carcinogens or ionizing radiation) might give rise to a cancer cell. Such reshuffling of genetic information when egg and sperm cells are produced or fused could perhaps serve to reprogram or reset the cell's biological clock. By this mechanism even if the individual members of a species were programmed to die, the species would live. A human being, then, would be the germ cells' way of making more immortal germ cells.



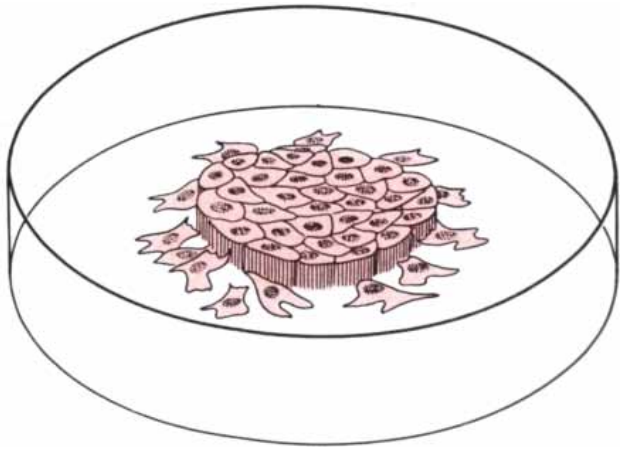
INFANCY



ADOLESCENCE



MATURITY



SENESCENCE

MIGRATION TIME of fibroblasts growing out from a tiny scrap of living tissue placed in culture varies inversely with age of the source. This has been ascertained for fibroblasts from chickens, rats and human beings and from several tissues, including heart, lung and skin.

World Uranium Resources

A case study in the estimation of mineral resources, supported by U.S. mining records, indicates that the supply of uranium will not be a limiting factor in the development of nuclear power

by Kenneth S. Deffeyes and Ian D. MacGregor

The task of estimating the world's mineral resources is a fairly recent undertaking in geology. Our own interest in the subject dates from 1972, when a group of analysts at the Massachusetts Institute of Technology published the book titled *The Limits to Growth*. The book was based on a study they had done that attempted to look into the future by establishing a quantitative model of the world economy and then extending the model forward in time with a computer. The model predicts disaster by about the year 2000 because of the depletion of resources. As geologists we were concerned with what was meant by the term resources in this context. A closer reading indicated that "resources" was actually being used in the same sense geologists use the word "reserves," that is, to denote the amount of material already discovered that is profitably recoverable under current economic conditions. To our surprise we could find no quantitative estimates of undiscovered mineral deposits or of known deposits that might become economic in the future. Before a study of the type exemplified by *The Limits to Growth* can ever be expected to yield a better forecast, estimates of the total availability of mineral resources must be attempted.

There is a particularly pressing need for such a resource estimate in the case of uranium. For example, the initial economic justification for the fast breeder reactor, which was being advanced at about the same time *The Limits to Growth* study appeared, assumed that the supply curve for uranium (the curve relating the price of the mineral to the amount recoverable at that price) followed a smoothly ascending trajectory [*see illustration on page 69*]. It seemed entirely possible at the time, however, that some modest price increase might cause virtually all the uranium deposits in a given geological category to be discovered and mined, and that a very large price rise would be necessary before another category would become exploitable; in that case the supply curve would first rise and then level off before beginning to rise again.

If such a curve were to describe the uranium-supply situation accurately, then raising the price to intermediate levels would amount to either a cash gift to the owners of the known deposits or an incentive to waste money, effort and energy in the search for nonexistent medium-grade deposits. (We know a few petroleum geologists who argue that oil prices above $7 per barrel are in just the kind of gap represented by the flat part of the colored supply curve in the illustration on page 69.) In spite of the obvious need for reliable resource estimates not only for nuclear fuel but also for fossil fuels, metals and other vital materials, there was no generally accepted method for making such estimates in 1972 and there is still none today. Indeed, for most minerals the raw geological data on which to base even a rudimentary resource-estimation method have yet to be assembled.

Of the few resource-estimation approaches that have been proposed in recent years most are narrowly specific for a given material. One model that was intended to be applicable to all resources was put forward in 1972 by Vincent E. McKelvey of the U.S. Geological Survey. He created a two-way grid of categories based on the degree of geological knowledge (known deposits, inferred deposits and probable deposits) and on current economics (economic, subeconomic and uneconomic). The only way of estimating the amount of material in each of McKelvey's categories is to ask a number of experts to guess at the amount of material in the category, then add up the estimates and divide the sum by the number of experts. Although the Geological Survey has gone on to try to measure the uncertainty in expert estimates for undiscovered oil, it is clearly difficult to base public policy on such subjective inputs.

A radically different approach to resource estimation was pioneered in the 1950's by M. King Hubbert of the Shell Development Company. He analyzed the history of oil production in the U.S. by fitting a symmetrical bell-shaped curve to the annual rate of domestic oil production. At the time of Hubbert's first estimate oil production was still at a comparatively early stage in the U.S., and he relied on expert opinions to define the total area under the bell-shaped curve. By the late 1960's, however, enough annual production figures had accumulated for him to extract an estimate of the total U.S. oil resource without relying on expert opinions. The verdict of history has been favorable to Hubbert's initial estimate: the recent course of domestic oil production follows his 1957 prediction quite closely [*see top illustration on page 70*].

During the 1960's Hubbert's methods were severely criticized because he utilized time as the independent variable rather than some measure of the effort spent on petroleum exploration. In 1967 he introduced a second approach, in which he compared the number of feet of exploratory borehole drilled in search of oil with the amount of oil discovered. What he found was an exponential decrease in the number of barrels of oil discovered per foot of borehole drilled as exploration proceeded. Not surprisingly, Hubbert's estimate of the total U.S. oil resource derived from the area under the exploratory-effort curve was remarkably similar to his estimate derived from the area under the bell-shaped annual-production curve.

Could Hubbert's methods serve as a basis for estimating other mineral resources, specifically uranium? There is no hope of applying his first approach, based on the bell-shaped annual-production curve, to uranium. The history of uranium mining is marked by "boom and bust" cycles, and there would be no plausible predictive value to such a time-dependent analysis. There is another bell-shaped curve, however, that is well known to geologists. In 1954 Louis H. Ahrens of the University of Cape Town noted that the abundances of trace elements in granites generally fol-

low a bell-shaped curve when the number of samples is plotted against the logarithm of the trace-element concentration [*see bottom illustration on page 70*]. On a logarithmic scale the curve fits a normal, or Gaussian, curve well enough to have led Ahrens to suggest that such a log-normal distribution is a fundamental law of geochemistry.

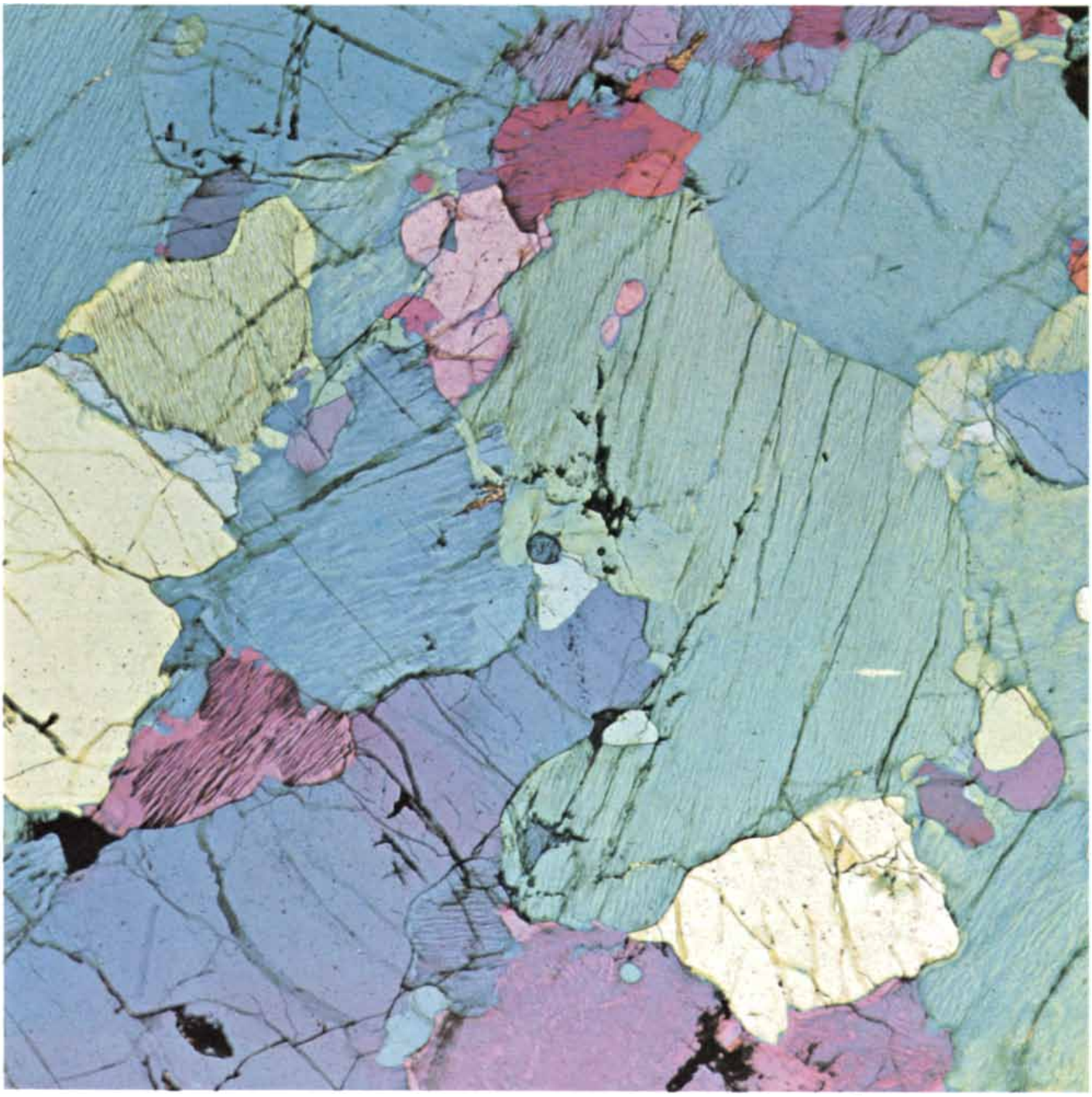
Although we doubt that log-normality is a law, we have taken it as a starting point for our investigation. Our approach has been to ask whether the distribution of uranium in the earth's crust can be reasonably approximated by a bell-shaped log-normal curve. In addition we have asked whether the uranium deposits actually mined appear to be a portion of the high-grade "tail," or ascending slope, of that distribution. This approach preserves what we feel are the two most important guiding principles of Hubbert's work, namely recognizing the geological framework that contains the deposits of interest and examining the industry's historical record of discovering those deposits. Our findings, published recently in the form of a book-length report prepared for the U.S. Department of Energy, suggest that for uranium the crustal-distribution model and the mining-history model can be brought together in a consistent

URANIUM IS TRANSPORTED from the earth's mantle to the upper part of the continental crust as a trace element in upwelling blocks of granite, such as the ones seen in this false-color ERTS 1 image of an arid region in Western Australia, viewed from an altitude of 570 miles. The large light-colored areas in the image are granitic domes, the vestiges of molten bodies of rock formed by heat generated as a by-product of the natural radioactivity of certain elements, including uranium. The rock formations were carried to the surface sometime in the Precambrian era (more than 600 million years ago). The subsequent weathering of the exposed granite has had the effect of releasing uranium into the surrounding sedimentary rocks, where under suitable conditions it can be concentrated with the aid of flowing water to form economically recoverable deposits. The chief alternative mechanism for transporting uranium to the surface is volcanism.

picture. In brief, we conclude that both sets of data can be described by a single log-normal curve, the smoothly ascending slope of which indicates approximately a 300-fold increase in the amount of uranium recoverable for each tenfold decrease in ore grade. This conclusion has important implications for the future availability of uranium.

There is no underlying geological theory requiring that the abundance in the earth's crust of trace elements such as uranium should follow a log-normal distribution. We have simply reasoned that rich deposits of uranium (and other trace elements) are formed only where several improbable circumstances come together. There has to be a source of the element, enough water to transport it, suitable subsurface conduits, complexing agents to carry the element in solution and other agents to precipitate the final minerals. Clearly if any one of these ingredients is lacking, no uranium deposit can be formed. Therefore the probability that an ore deposit will be formed at a given site is determined by multiplying the probabilities that each essential ingredient was present. Whenever probabilities add, the central-limit theorem of statistics holds that the eventual distribution approaches the bell-shaped normal distribution. Since the multiplication of probabilities corresponds to adding on a logarithmic scale, it is not surprising to see the familiar bell-shaped curve appear when elemental abundances are plotted on a logarithmic scale.

We hasten to point out that this is only an approximate argument; no rigorous

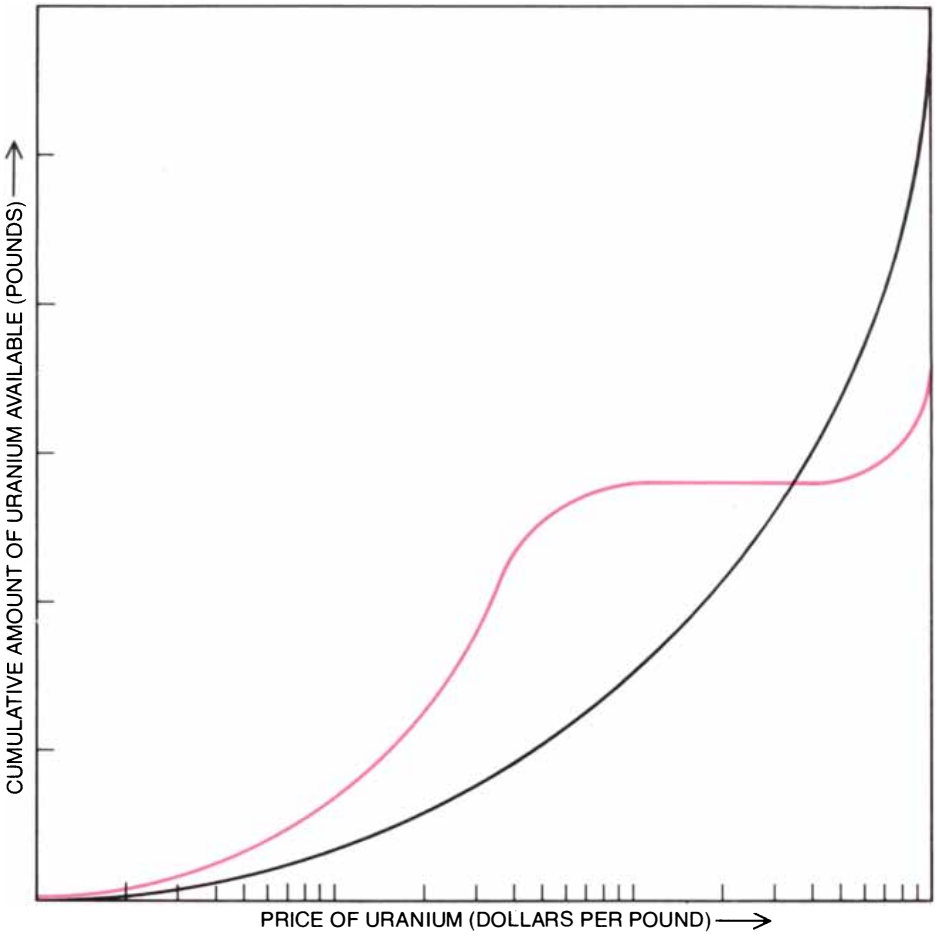


URANIUM IN GRANITE is usually concentrated in a few of the less abundant minerals rather than being evenly scattered throughout the major minerals. This transmitted-light micrograph (magnification about 50 diameters) shows a slice of granite only .03 millimeter thick. The colors result from the interference of the transmitted light waves, which are broken into two components traveling at different speeds as they pass through the crystalline materials of the sample and through a single-crystal quartz plate. Most of the uranium in the field of view is contained in a crystal of zircon, or zirconium silicate ($ZrSiO_4$), which appears as the small dark blue grain at the center.

statistical basis exists for expecting a log-normal distribution. If the curve showing the log-normal distribution were a full description of the minor elements of the earth's crust, then the recoverable ore deposits should be simply the high-grade tail of the log-normal distribution. A quick calculation shows that this is not always the case. For example, taking the best-fitting log-normal curve for Ahrens' plot of the distribution of lead in granite and extending the curve out to a lead concentration of 20,000 parts per million (the lower limit at present for economically recoverable lead ores), one finds that the total amount of lead ore in the world should be no greater than a single 50-gram specimen of galena (lead sulfide) in the mineral collection of the Smithsonian Institution! Of course there are minerals so rare that only a single specimen is known, but galena is not one of them; more than 190 million tons of lead has already been mined in this form. Accordingly Ahrens' curve for lead in granite has to be modified. There are two obvious ways the actual distribution of a trace element could differ from the log-normal one. Either the high-grade tail of the distribution contains more material than the theoretical log-normal curve suggests or the ore deposits form a separate and independent distribution.

The case for a separate distribution describing ore deposits has been stated by Brian J. Skinner of Yale University. He points out that in ore deposits the metals being mined exist as minerals in which the metal is an essential component of the crystal structure. In ordinary rock the trace amounts of the same metal can be accommodated by an atom-for-atom substitution in common minerals. For example, lead exists as galena in most lead ores, but in common rock-forming minerals such as the feldspars lead substitutes in small amounts for potassium. Skinner has pointed out that the overall crustal distribution could be bimodal, with one peak for the ore deposits and a second peak for the common rocks. The consequences for society could be severe if the high-grade peak of a trace element such as uranium were to be exhausted and mining had to shift suddenly to the much-lower-grade common rocks. A bimodal distribution of the type Skinner proposes would cause the kind of discontinuity represented by the flat portion of the colored supply curve for uranium in the illustration on this page.

Of all the metals, chromium seems to us to be the one most likely to exhibit the kind of bimodal distribution Skinner suggests. A single geological process, the crystallization of large basaltic magma bodies, has produced all the major chromium deposits in the world. The worldwide uniformity of chromium deposits may reflect the metal's comparatively simple geochemistry and the dominance of a single type of deposit.

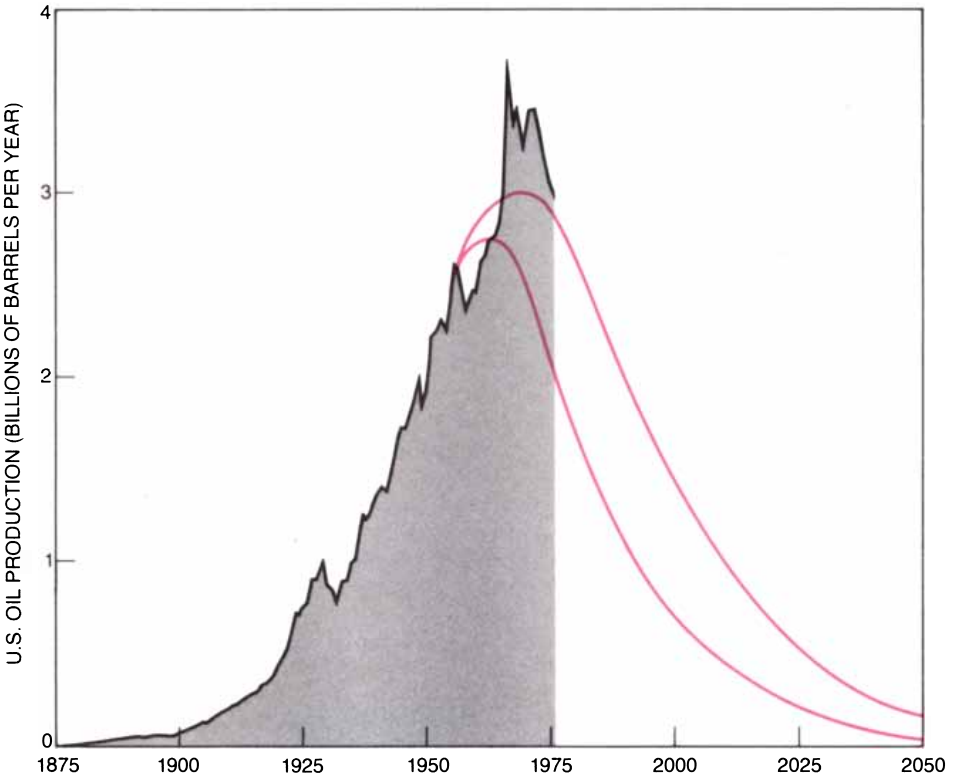


HYPOTHETICAL SUPPLY CURVES for uranium relate the projected price of the uranium (*logarithmic horizontal scale*) to the cumulative amount recoverable at that price (*arithmetic vertical scale*). One possibility is that the supply curve will simply follow a smoothly ascending trajectory (*black curve*); in other words, the higher the price, the greater the amount of uranium available. The other possibility is that some modest price increase will cause virtually all the uranium deposits in a given geological category to be discovered and mined; a very large price rise would then be necessary before another category would become exploitable. The result would be a supply curve that would rise and then level off before beginning to rise again (*colored curve*). The authors' findings suggest that the black curve is more likely to be correct.

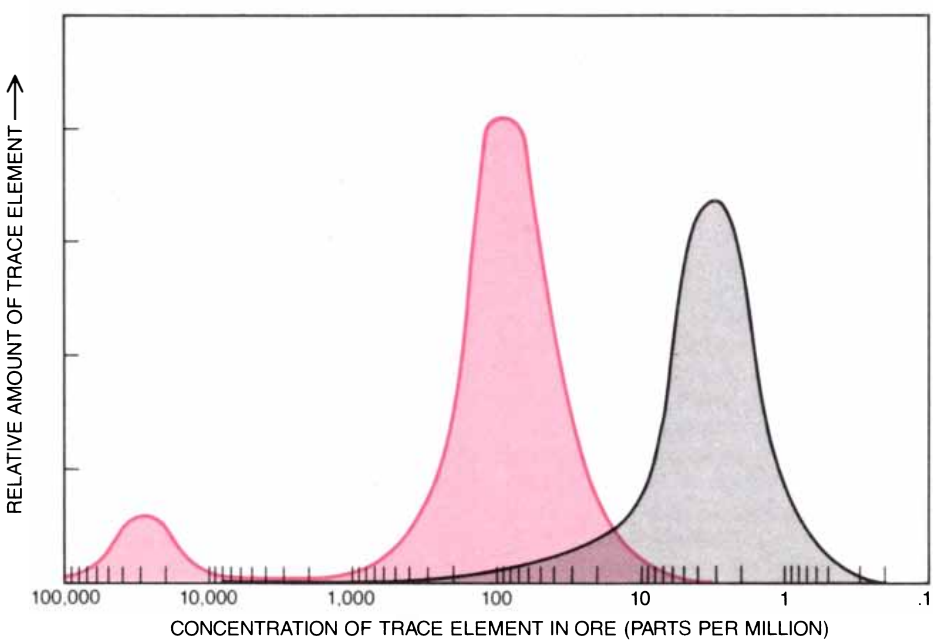
Uranium, in contrast, exhibits an enormously complex range of geochemical behavior and a wide variety of different kinds of economic deposit. Two valence states of uranium, one with a deficit of four electrons and the other with a deficit of six, occur in nature, and a variety of ions such as the uranyl carbonate complexes, $UO_2(CO_3)_2^{--}$ and $UO_2(CO_3)_3^{----}$, contribute to the diversity of uranium's behavior. Instead of a bimodal distribution such as the one that might be expected for chromium, uranium ought to have a multimodal distribution. Each chemical style and geological setting should have its own characteristic average uranium concentration. We have tested the simple hypothesis that the ensemble of all these diverse uranium deposits combines to form a simple log-normal distribution. To repeat, we have no strong theoretical reason to expect a log-normal curve; we have simply put a fresh blade in Occam's razor and examined the simplest hypothesis first.

Unfortunately it is not possible to determine the distribution of uranium deposits directly by analyzing for uranium in a large random sample of rocks. Even if there were a laboratory that could turn out one uranium analysis per second, it would take 30 years, running around the clock, to define the shape of the tails of the distribution out to such rare occurrences as economic uranium mines. An analytical scheme of this type is tantamount to prospecting for uranium by picking up rocks blindfolded until one accidentally discovers a substantial number of economic uranium mines. Since it is not possible to measure the distribution directly, one has to approach the problem in stages, examining first the uranium distribution in the major geochemical reservoirs and then focusing on the smaller units that have an unusually high uranium content.

Two dominant factors influence the partitioning of uranium into the major geochemical reservoirs: its large ionic radius and its production of heat by radioactivity. Under the conditions expected in the earth's mantle uranium exists in the quadruply ionized valence



SUCCESSFUL RESOURCE ESTIMATE was made in 1957 by the geologist M. King Hubbert, who fitted a bell-shaped curve to the history of U.S. oil production, relying on expert opinions to define the area under the curve. (The two colored curves represent the upper and lower limits of his estimate.) The recent course of domestic oil production follows Hubbert's prediction closely. Area under curve in this case corresponds to total amount of resource available.

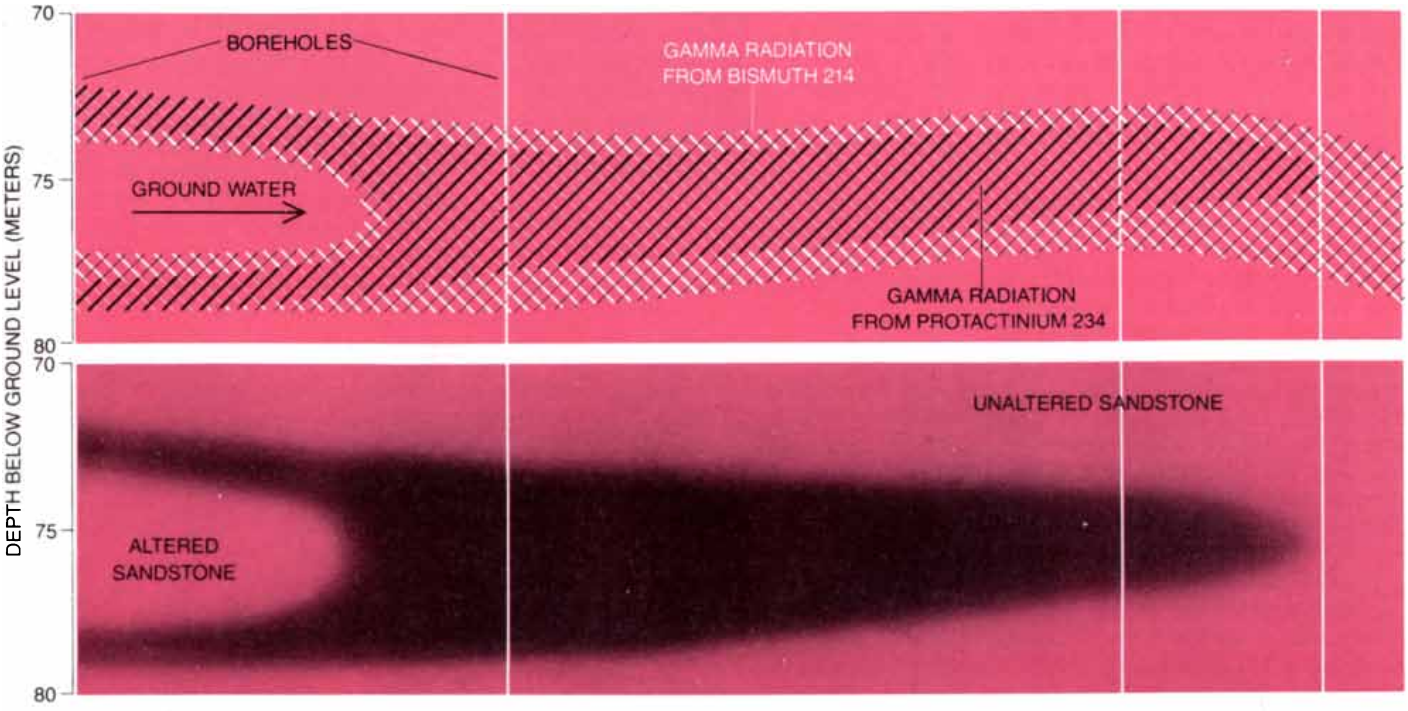


DISTRIBUTION OF TRACE ELEMENTS in granite generally follows a bell-shaped curve when the amount found in various samples is plotted against the logarithm of the trace-element concentration. The central question asked at the outset of the authors' study was whether the distribution of uranium in the earth's crust could be described by such a bell-shaped log-normal curve. Their interpretation of the evidence suggests that it can, and the black curve is their approximation of the overall shape of the log-normal distribution curve for uranium. (The curve is not perfectly symmetrical: its high-grade "tail," or ascending slope, has been adjusted to indicate the availability of somewhat more high-grade ore than the theoretical log-normal curve would suggest.) An alternative model of the overall crustal distributions of the trace elements, put forward by Brian J. Skinner of Yale University, portrays them as following a bimodal curve, with one peak for the ore deposits and a second peak for the common rocks. Chromium, for example, may have a bimodal distribution, represented here by the colored curve.

state, and its ionic radius in crystals is comparable to the radius of ionized sodium. Like sodium ions and other large ions, uranium ions selectively enter any partial melt that is generated from the mantle. Therefore uranium ions are readily transferred out of the mantle. (It happens that the other two major radioactive heat producers, thorium and potassium, also have large ionic radii and are chemically fractionated out of the mantle. The world would have been a very different place if any of the smaller common ions such as magnesium had been radioactive, because the heat source would then have remained deep in the mantle, where it would have given rise to much stronger convection currents and disrupted the crust to a greater extent.) The extremely low uranium content of iron meteorites suggests, by analogy, that the earth's core is also deficient in uranium.

Within the earth's crust there are a few crystallographic niches where uranium can make an approximate fit. There is, however, an additional constraint on the crustal distribution of uranium. If the uranium content of a block of rock within the crust generates heat faster than the heat can be carried away by thermal conduction, the block will melt (partially or completely) and come toward the surface as molten magma. To be more precise, the time required for the block to reach its melting point is proportional to its specific heat and is inversely proportional to its internal rate of radioactive heat production. The distance to which heat can diffuse is given by the square root of the product of the time and the thermal diffusivity. Any body of rock that is deep enough and has a sufficiently high uranium content will melt. During mountain building, for instance in a collision between two continents, the components of the newly formed crust are reassembled without regard to their uranium content. In time self-heating can melt large bodies of rock, forming the characteristic post-tectonic granites of the earth's mountain belts. The upward transport of heat-producing elements by the mobilization of granite is a major mechanism in transferring uranium into the upper part of the continental crust. Afterward erosion exposes these granites to weathering, and the uranium in them finds its way into sedimentary rocks.

There is a second way large amounts of uranium-bearing granitic material are transferred to sedimentary rocks. Some melts with granitic compositions reach the surface in volcanic eruptions. The more water-rich of these melts typically form a froth of millimeter-size steam bubbles with glass walls. Small amounts of material are preserved in this stage as the lightweight rhyolitic rock called pumice, but for most of the



URANIUM DEPOSIT TRAVELS SLOWLY through sandstone in the direction of the flowing ground water, as these two cross sections show. An ore deposit of this type, called a roll-front deposit, has a characteristic crescent-shaped reaction zone, where the uranium ions are precipitated out of the water, usually in the form of the mineral uraninite (UO_2). The actual movement of the deposit can be demonstrated by subsurface gamma-ray measurements, made with the aid of a detector that is able to distinguish the gamma radiation emitted by protactinium 234, the first "daughter" isotope of uranium 238, from the much more abundant gamma rays emitted by bismuth 214, a later radioactive-decay product (*top*). The fact that the area where the gamma rays from the bismuth predominate (*heavy white hatching*) is generally downstream from the area where the gamma rays from the protactinium predominate (*heavy black hatching*) is taken as evidence that the uranium deposit itself (*gray shading at bottom*) is being transferred gradually downstream by moving ground water.

froth the continued expansion of the steam shatters the bubble walls, and a cloud of steam and fine glass shards is propelled high into the atmosphere. These are the Plinian eruptions, named for Pliny the Younger's description of the eruption of Mount Vesuvius in A.D. 79, which buried Pompeii and Herculaneum. In the 1950's several geologists noted that the total amount of volcanic glass of granitic composition transferred to the surface by these enormous volcanic explosions may be equal in volume to the great bodies of granite within the mountain belts. The products of the eruption are often propelled upward through the entire troposphere, and the fallout of volcanic ash and other fragments, collectively referred to as tephra, can extend downwind for hundreds of kilometers [see "Tephra," by Laurence R. Kittleman; SCIENTIFIC AMERICAN, December, 1979].

Once the atmospheric dispersion and the further transport by rivers have brought the volcanic fragments to sites where sediments are accumulating, the glass, which is less stable than a crystalline assemblage of the same composition, gradually breaks down. The release of chemicals from the breakdown of volcanic glass accounts for the extraordinary fertility of volcanic soils. At the same time a small amount of uranium is released.

Uranium derived either from the surface weathering of granite or from the breakdown of volcanic ash can be transferred through the accumulating sediments by ground water, and at least some uranium ore deposits are found where uranium is precipitated out of moving ground water. There are several other types of uranium deposit, but in one way or another they all owe their origin to the progressive upward transfer of uranium into the top part of the continental crust and also to one or more final steps of concentration within the crust.

The uranium content of the major units of the earth is supplemented by some smaller units that are highly enriched in uranium. The most uranium-rich of all these units are metal-rich veins deposited in fractures by the movement of hot water (with the heat possibly coming from the radioactivity of the uranium in the adjacent rocks). Vein deposits such as those at Joachimsthal in Czechoslovakia, Great Bear Lake in Canada and Katanga in what is now Zaire were the main source of uranium before World War II. In addition to veins deposited by hot water there are some high-grade uranium deposits in pegmatites, which are unusually coarse-grained igneous rocks formed from melts rich in dissolved gases.

Several recently discovered deposits of uranium ore in Saskatchewan and northern Australia have been found at the base of Precambrian sandstone layers that overlie even older Precambrian granites and metamorphic rocks. Possibly as a result of their great age, these deposits, which are a newly recognized category of deposits of the unconformity type, have subsequently been rearranged into veins and other fillings of open spaces, and they have been heated to temperatures of a few hundred degrees Celsius. During this later phase of their history much of the evidence that would have unequivocally demonstrated their origin has been obliterated. As a result there is some disagreement among geologists on the question of why the uranium in these deposits is associated with the unconformities.

The next category of uranium ore (in descending order of uranium content) consists of the deposits of the sandstone type. These ores typically consist of uranium minerals deposited interstitially between the sand grains in sandstones. The bizarre aspect of such deposits is that many of them are moving actively, if slowly, through the sandstone when the faster-moving uranium miners come along and capture them. The motion of the uranium deposit is quite similar to the motion of a front in a laboratory chromatography column. If oxygenated water from the surface is introduced into a sandstone that contains small amounts of reducing agents (such as fos-

sil organic matter, petroleum or sulfides), a zone of active chemical reaction will develop between the reducing agents and the oxygen in the water. As oxygenated water continues to be introduced the reaction zone moves in the direction of the water flow.

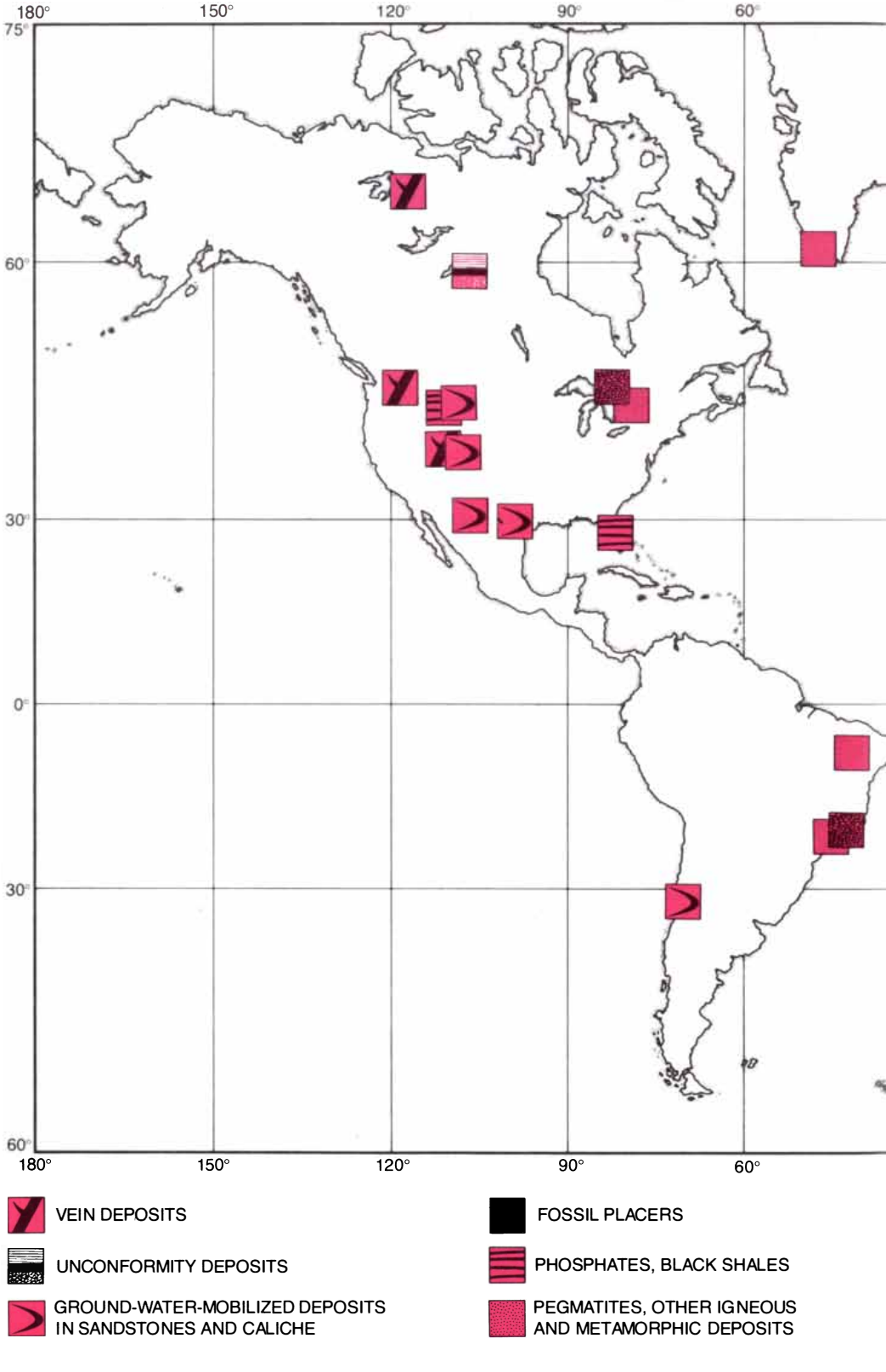
Oxidizing even 1 percent of the organic carbon in a sandstone requires all the dissolved oxygen in 500 pore volumes of water; therefore the reaction zone moves along much slower than the water. As a result the reaction zone becomes a boundary between oxygenated water and water from which the oxygen has been removed. Upstream from the reaction zone uranium is soluble as the hexavalent uranyl carbonate complex, and uranium picked up from the surface weathering of granite, from the alteration of volcanic ash or from other sources can move toward the reaction zone. Within the part of the zone in which oxygen is being removed from the ground water uranium is changed from the hexavalent state to the quadrivalent one. The quadrivalent uranium is insoluble and precipitates out, usually in the form of the mineral uraninite (UO_2). Because of the characteristic shape of these actively moving uranium deposits they are called roll-front deposits.

The fact that deposits of this type are actively moving today has been directly demonstrated by subsurface measurements made with a gamma-ray spectrometer lowered down a series of boreholes. A group at the firm of Princeton Gamma-Tech has packaged a germanium-crystal detector, its cryogenic heat sink and a multichannel analyzer into a probe two inches in diameter and about 12 feet long. The probe, which has been used for uranium prospecting by the U.S. Geological Survey, among others, can separate the gamma radiation emitted by protactinium 234, the first "daughter" isotope of uranium 238, from the abundant gamma rays emitted by bismuth 214, a later product of the radioactive-decay chain.

From data obtained this way one can plot a cross section showing both the gamma radiation from the protactinium near the uranium and the gamma radiation from the bismuth, which has been transferred downstream by the moving ground water as the several daughter isotopes sequentially decayed [*see illustration on preceding page*]. The new technique is an important advance; until now the lack of correspondence between the radioactive bismuth (which dominates in most total-radioactivity surveys) and the uranium has been a bane to the prospector, because the most radioactive rocks are not necessarily the most uranium-rich.

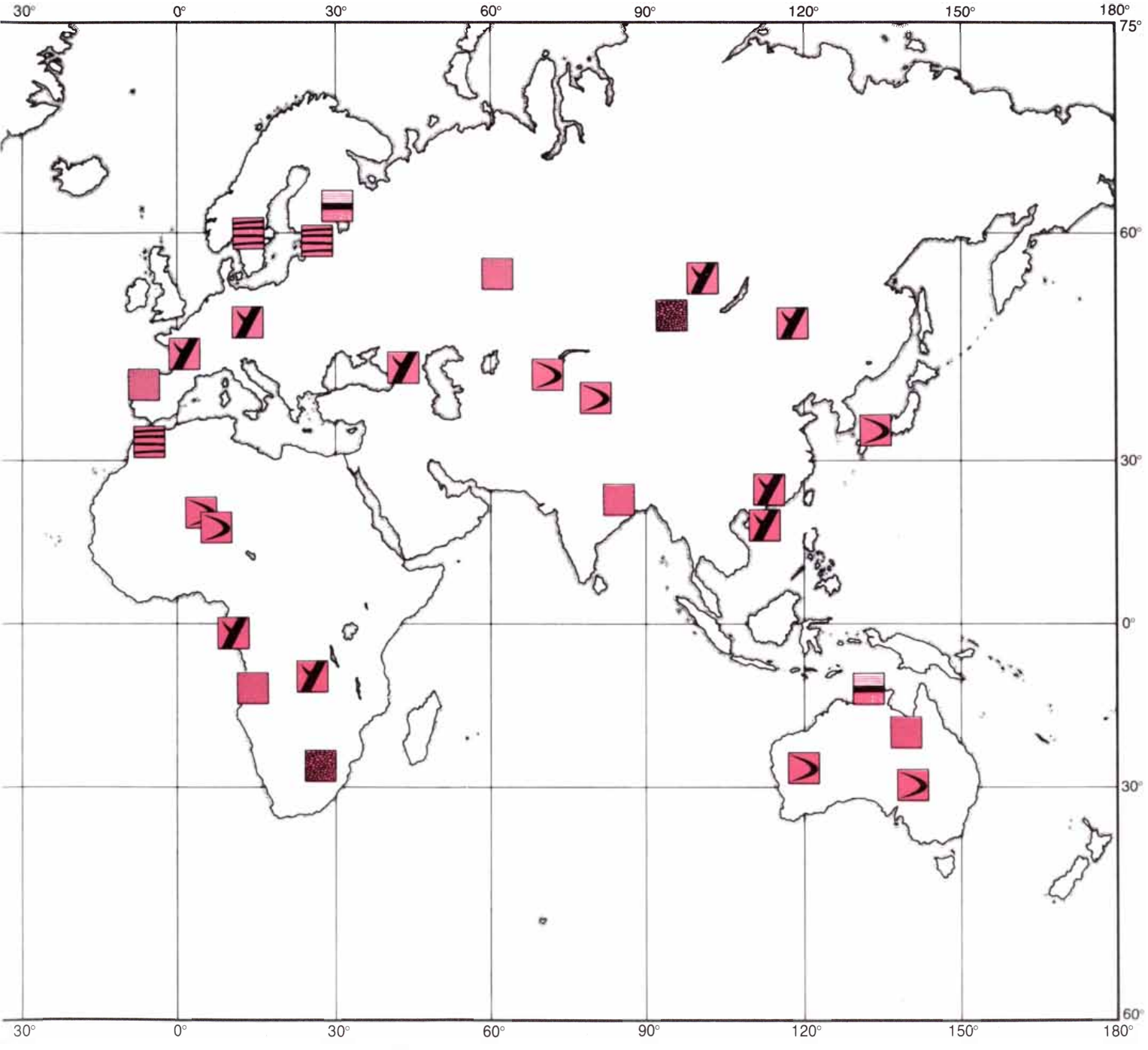
Uranium deposits of the sandstone type in Wyoming, New Mexico, Utah and Colorado have dominated U.S. uranium production. Until recently there were very few significant deposits of uranium known in sandstone outside the U.S. Several uranium mines that are being opened in Niger seem to have the same origin as the sandstone deposits in the western U.S., although only sketchy descriptions of them are currently available.

Uranium deposits of lower grade than the typical sandstone deposits are not now economic to mine, unless a valuable by-product or coproduct can be extracted along with the uranium. By far the most important example is the coproduction of gold and uranium from the Witwatersrand deposits in South Africa. During the first 60 years of the exploitation of the Witwatersrand gold deposits the cost of mining and extraction was borne primarily by the gold that was recovered. The existence of uranium in the gold-bearing rock was long

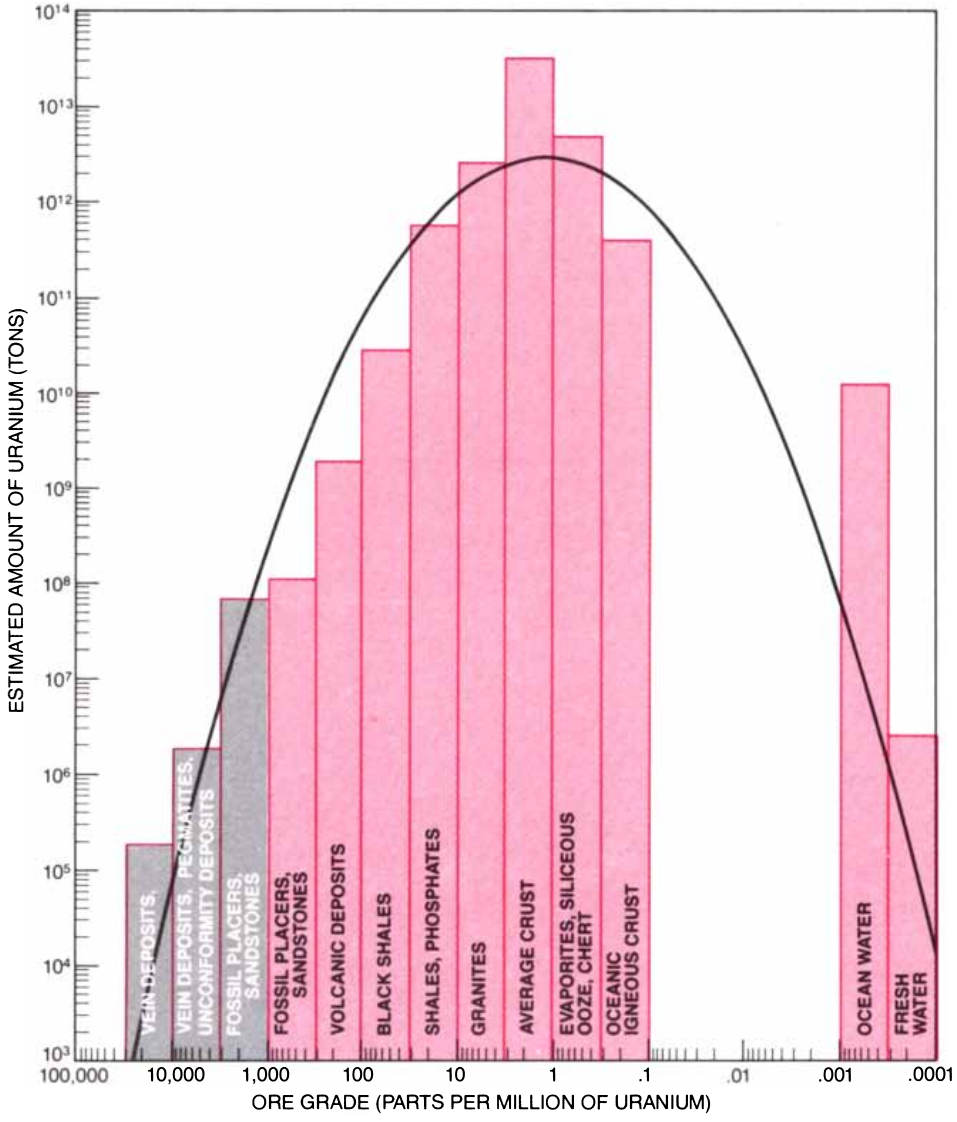


known; in fact, the radioactivity of the uranium was used as an exploration guide for selecting the best gold ores. After World War II some extraction of uranium was begun by sending the gold ore through a uranium-recovery cycle after the gold was extracted. In the course of the uranium extraction some additional unrecovered gold was liberated, and so the process was reversed, the uranium being recovered first and the gold recovered second. Economic changes also supervened; some mines were opened that would have been uneconomic for gold alone but were viable when operated for the coproduction of uranium and gold.

The Witwatersrand deposits, along with similar deposits at Elliot Lake in Canada, are fossil placers. Present-day placer deposits of gold in rivers are familiar, but in the earth's present oxidizing atmosphere uranium minerals are quickly oxidized to the soluble hexavalent complexes. If the earth's early atmosphere was in equilibrium with the igneous rocks of the crust, however, the conditions would have been reducing instead of oxidizing. Before the atmosphere was converted by photosynthesis into an oxidizing state, uraninite could travel in rivers as stable heavy grains and could be concentrated along with



KNOWN DEPOSITS OF URANIUM are identified on this world map by a set of symbols suggestive of the various major types of deposit. The key at the left lists the six most important categories of uranium ore in descending order of their uranium content. The deposits shown are economically recoverable now or in the foreseeable future either in straight mining operations or in association with the recovery of some valuable by-product or coproduct. Examples of the latter type of association include the fossil placer deposits at Witwatersrand in South Africa and Elliot Lake in Canada, where both gold and uranium are recovered, and the phosphate deposits in various parts of the world, where uranium is liberated in the process of converting phosphate rock into soluble phosphate fertilizer. The newly discovered deposit of the unconformity type at Jabiluka in northern Australia apparently contains more uranium than all the uranium that was produced in U.S. between 1948 and 1970. Lower-grade deposits, such as abundant Chattanooga shales in the eastern U.S., are not included.



DISTRIBUTION OF URANIUM in the earth's crust among the major geological reservoirs of the element is plotted here on a log-log chart. The bars representing the various categories of ore deposits (in descending order of their uranium content) define a log-normal global-abundance curve, which in this instance takes the form of a parabola opening downward (*black curve*). Ascending slope of parabola is roughly 2.5 to 1, which corresponds to a 300-fold increase in the estimated amount of recoverable uranium for every tenfold decrease in the ore grade. The three gray bars at the left indicate the deposits now being mined for uranium alone.

the grains of gold. No known deposits of the Witwatersrand type are younger than 1,800 million years; most geologists accept this finding as evidence for the time at which the atmosphere became at least partially oxidizing.

Another possible by-product association for uranium comes from the mining of phosphate rock for fertilizer. Phosphate rock typically carries from 10 to 300 parts per million of uranium, lower than the minimum concentration now being mined as straight uranium ore. The process of converting the phosphate rock into soluble phosphate fertilizer, however, liberates most of the uranium. If for no other reason than environmental concerns, removal of the uranium would be desirable. Once the advantages of recovering the uranium are established, the value of the rock is the sum of the values of the uranium and the phosphate, and mining the most valuable uranium-phosphate rock can be considered. Iris Y. P. Borg of the Lawrence Livermore Laboratory showed in 1974 that by-product uranium from phosphate mining alone might supply the entire demand of U.S. fission power plants for uranium.

There are additional subeconomic uranium deposits, including the extensive sedimentary-basin fills in Nevada and adjacent states, which contain abundant volcanic ash that has released uranium into the system. Marine black shales usually contain more than the average crustal abundance of uranium, and some black shales, such as the Chattanooga shale, get into the range between 10 and 100 parts per million of uranium. The most uranium-rich of the granites contain roughly 10 parts per million of uranium.

In the aggregate all these deposits define a global abundance curve for uranium [*see illustration at left*]. Of course, there is the possibility that we have grossly underestimated the amount of uranium in one of the categories. Since airborne radiation counters became a common prospecting tool some 25 years ago, however, millions of miles of flight lines have been surveyed. Therefore the odds are small that some major category has been systematically underestimated. Because the vertical scale of this particular global uranium-abundance chart is logarithmic, the expression for the expected normal curve is a parabola opening downward. As the illustration shows, a parabola can be drawn that gives a reasonable approximation of the data. Therefore the simplest hypothesis, a single log-normal distribution, explains most of the data, and we see no reason at this time to entertain more complicated hypotheses. Of particular interest is the slope of the parabola as it goes through the concentrations that are now being mined. The slope on the log-log plot is about 2.5 to 1, which corresponds to a 300-fold increase in recoverable uranium for each tenfold decrease in ore grade. This trend would seem to guarantee an ever increasing uranium supply as the uranium prices justify mining lower grades of ore. Before accepting this conclusion, however, we examined carefully the history of uranium mining in the U.S. to see whether the relation between ore supply and ore grade could be confirmed.

For most metals other than uranium there has been a long historical progression from the mining of high-grade deposits toward lower-grade ores, greater depths and higher costs. In contrast, some of the lowest-grade uranium ores ever mined were mined early in the intensive program immediately after World War II. Because subsidies were made available in the U.S. to encourage the mining of even the lower-grade ores, the time progression of ore grades is chaotic. By ignoring time and concentrating on the amount of uranium recovered in each ore grade, however, we have constructed an artificial history that develops as if the highest-grade uranium ores had been mined first.

The only purchaser of uranium in the U.S. initially was the Atomic Energy Commission, and even after there was a private market for uranium the A.E.C. and its successor agencies continued to keep records of the amounts of uranium mined. Access to the records is difficult because mining companies made the data available to the Government with the understanding that data would be released only in statistical categories that did not identify individual mines or companies. Throughout our work we had to write computer programs for tabulating the data into categories and submit each program to a Department of Energy employee who would run the

program and examine the output before letting us see the result. The frustration we experienced in getting the programs debugged and putting the data in satisfactory form was enormous. As a partial way of sparing future investigators the same difficulties we have appended a 400-page atlas of graphs to our recent Department of Energy report.

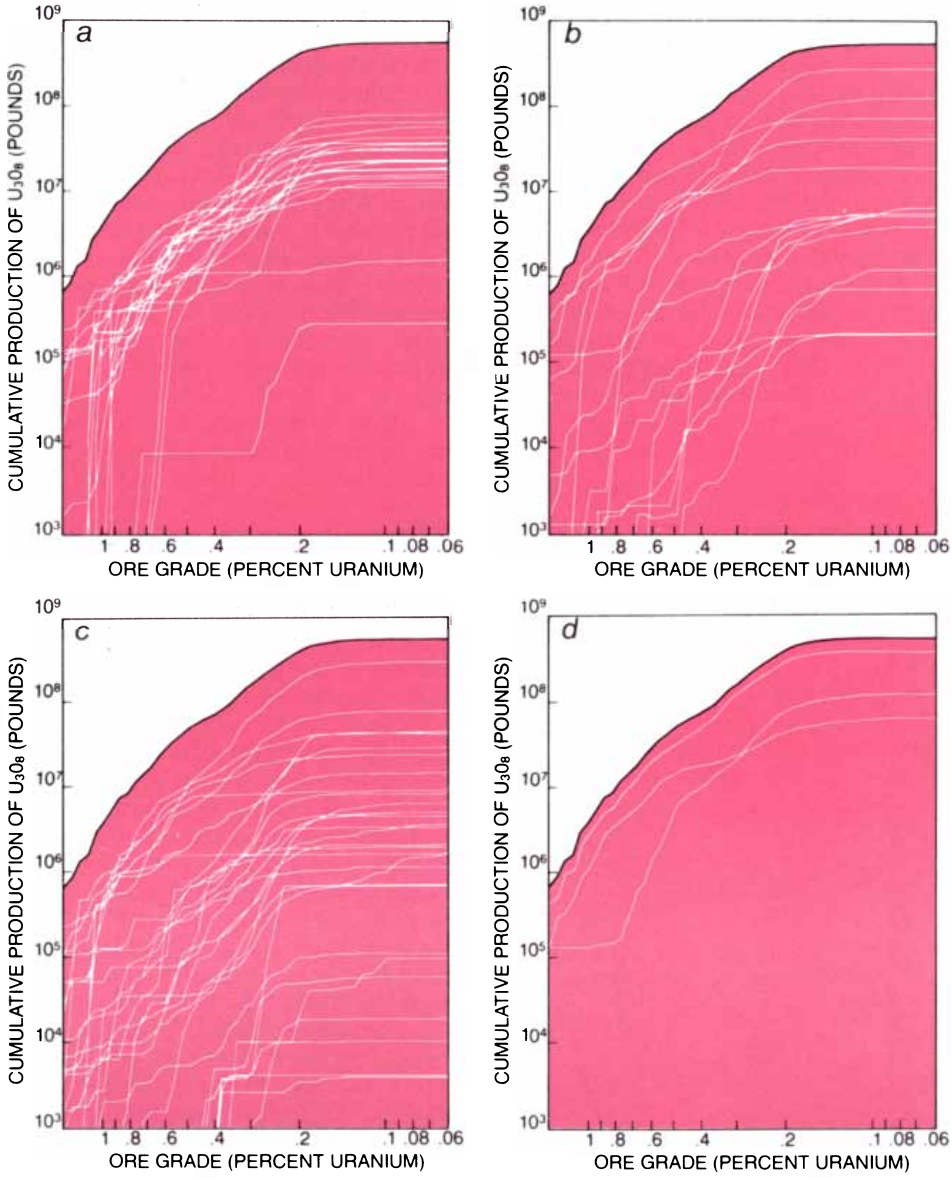
Even if there were no human errors in recording the mining history, there is one way the reported ore grades do not faithfully represent the ore in the ground. There is a tendency in metal mining in general and in uranium mining in particular to balance out blocks of unusually high-grade ore by mining simultaneously some low-grade ore to maintain a constant product. Some uranium mines maintain "pads" near the mine mouth with stocks of low- and high-grade ore for blending with the current production to keep an optimal feed going to the processing plant. For any one mine the ores in the ground are thereby averaged into a single ore grade, which may continue unchanged for years. Fortunately most uranium mines have their own processing plants optimized for the average ore in that mine; hence there is no averaging imposed between mines. Because of the large number of mines in the U.S. we do not think large anomalies appear in the data as a result of the blending of ores.

The curve of U.S. uranium ore grades mined with respect to the amount of uranium recovered is shown by the black curve at the top of each of the four graphs at the right. The slope of the curve of mined uranium deposits, which is remarkably similar to the slope of the curve for the earth's crust in the same composition range, is about 2.5 to 1 (on a log-log scale), again implying a 300-fold increase in available uranium for a tenfold decrease in ore grade. We have tested the firmness of this conclusion by subdividing the data into subsets reflecting the depth of the mine, the thickness of the ore layer, the mineralogy of the uranium, the geology of the host rock and the geographical region. In every case the slopes of the curves relating ore grade to uranium content were substantially the same, except for those subsets including only a few mines. We are convinced that the agreement between the crustal-abundance data, taken over a wide range of geological occurrences, and the U.S. mining-history data establishes an expected trend for future uranium exploration and mining. Furthermore, the curve suggests that a large amount of uranium will be recovered in the range between 300 and 800 parts per million and that deposits as lean as the Chattanooga shale will not be mined in this century.

The slope of the curve relating the ore grade to the amount of ore mined for uranium can be compared with the slopes of similar curves for other metals [*see illustration on next page*]. Such a comparison shows that the curve we postulate for uranium is steeper than the curve for copper and flatter than the curves for chromium and molybdenum. Therefore the uranium curve does not seem to be unusual.

There is a striking parallel between the characteristics of uranium and silver. Both metals have roughly the same average crustal abundance (one part per million for uranium and .1 part per million for silver); both have had (until recently) roughly comparable prices; the leanest minable ore grades for both are enriched about 2,000 times over the average crustal abundance, and both are recovered by using a complexing agent to dissolve the metal selectively into a dilute solution and then stripping the metal from the solution. The main difference is that unlike the short history of uranium the mining history of silver in the U.S. started in about 1860 and the last major silver discoveries were made in about 1915. Since then most domestic silver production has been the by-product of copper mining.

There is a second method for estimating the amount of uranium that is likely to be discovered at each ore grade. This method is derived from Hubbert's extrapolation of the amount

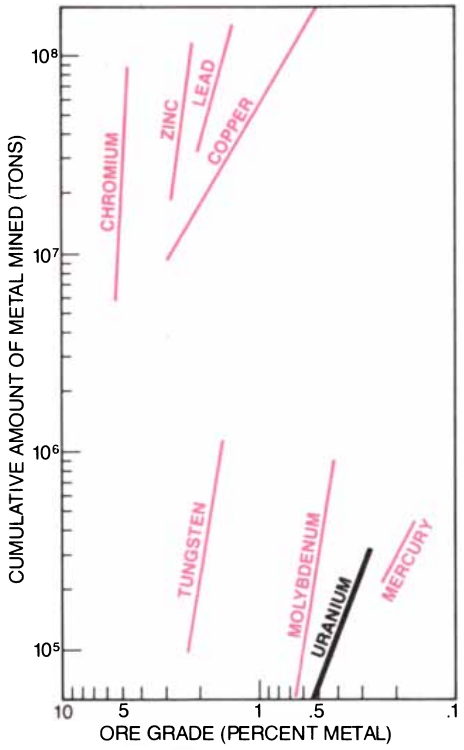


COMPUTER-GENERATED CURVES are based on reported records of the amount of uranium recovered (in the form of "yellowcake," or U_3O_8) from a large number of mines in the U.S. as a function of the grade of the recovered ore. The overall history of uranium mining in the U.S. is given by the black curve at the top of these four graphs. The white curves were obtained by subdividing the data into subsets depending on the depth of the mine (*a*), the geology of the host rock (*b*), the mineralogy of the uranium (*c*) and the geographical region of the U.S. in which the mines were situated (*d*). In every case the slopes of the curves relating ore grade to uranium content are substantially the same: about 2.5 to 1 on a log-log scale, again implying a 300-fold increase in the amount of uranium available for each tenfold decrease in ore grade. These summary curves, which are representative of a 400-page atlas prepared by the authors for the Department of Energy, suggest a large amount of uranium could be recovered in U.S. in the range between 300 and 800 parts per million (between .03 and .08 percent uranium).

of oil found per foot of borehole drilled. Michael A. Lieberman of the University of California at Berkeley has plotted the amount of uranium found per foot of exploratory borehole drilled against the total number of feet of exploratory borehole drilled. He has gone on to estimate the total size of the resource by extrapolating the resulting curve out to an infinite amount of exploratory drilling. Since Lieberman's curve exhibits an approximately exponential decline, the extrapolation is mathematically very simple. Nevertheless, there are several difficulties with the details of Lieberman's analysis. The data Lieberman had available were categorized by current mining costs: the known amounts of uranium ore minable at $8 per pound of "yellowcake" (U_3O_8), at $10 per pound, at $15 per pound and at $30 per pound. Lieberman was aware that in some years more uranium disappeared from the $8-per-pound category as a result of inflation than disappeared as a result of mining. To avoid this difficulty we have carried out calculations both in constant-value dollars corrected for inflation and in current-value dollars.

A second problem with Lieberman's analysis was the underreporting of ore in the lower grades. Lieberman used the ratio of low-grade (high-cost) ore to high-grade (low-cost) ore to assess the total availability of uranium from the lower ore grades. In our study we were surprised to find that for one year, only three mines had bothered to report their ore reserves in the lowest-grade category ($30 per pound). This omission not only seriously underestimated the amount of low-grade ore available but also signaled another underestimation. Taking into account only the amount of low-grade ore found in and around high-grade deposits ignores all those terrains that have never produced high-grade uranium ores but that are known to hold low-grade deposits. For both of these reasons Lieberman obtained estimates of recoverable uranium considerably smaller than the current Department of Energy estimates.

We have repeated Lieberman's analysis using as various measures of exploration effort not only the number of feet of borehole drilled but also the number of individual boreholes, the total exploration expenditures and the number of acres of land leased for uranium exploration. When these estimates are carried out for constant and current dollars, an entire range of estimates results, most of which lie between Lieberman's estimate and the Department of Energy estimates. Our feeling is that the scatter among the various estimates is a rough measure of the uncertainty of the entire procedure, and we would be skeptical of any attempt to select a "correct" estimate out of the scattered answers.



SLOPE OF THE CURVE relating the grade of the ore to the amount of ore mined for uranium (*black line*) turns out to be not very different from the slopes of similar curves for the mining of other metals (*colored lines*). For example, the curve postulated by the authors for uranium is steeper than the curve for copper but less steep than the curve for chromium.

The prospect of an increasing supply of uranium as the ore grade is lowered seems considerably more optimistic than the comparable picture for petroleum. By utilizing the lower-grade components of the deposits of the sandstone type, by recovering the uranium that could be the by-product of phosphate mining and by searching for other ores in the range between 600 and 2,000 parts per million of uranium, the U.S. could supply even an expanded nuclear-power industry from domestic uranium sources without having the cost of uranium mining become an important part of the cost of energy. Of course, that may not happen. As has happened so often in the past, supplies from outside the U.S. could now come on the market at lower cost than domestic production would require. In particular the ore grades in some of the deposits of the unconformity type in Canada and Australia are at the level of 10,000 parts per million in shallow deposits, which could be mined by open-pit methods. One deposit alone, at Jabiluka in northern Australia, has already been found from core drilling to contain more uranium than was produced in the U.S. in the entire period from 1948 through 1970. If large amounts of Canadian and Australian uranium were offered on the world market at prices modestly above the cost of mining and processing, the U.S. would again face the dilemma of whether to subsidize an active program of domestic exploration and production or to buy the imported minerals and hold the lower-grade domestic ores as a reserve for the future.

A few years ago some opponents of nuclear power argued that estimates of undiscovered uranium were so low that they precluded the construction of additional power reactors. Now that much larger estimates of worldwide uranium availability are being made, what is the likely impact on the future of uranium-generated power? The present generation of light-water reactors was designed at a time when uranium ore was cheap, less than $8 per pound of yellowcake, compared with the present price of $40 per pound. Because the interest on the capital cost of the reactor was a much larger component of the cost of the electricity than the uranium was, the reactors were designed to minimize capital costs even if they were inefficient uranium burners. (One of our colleagues calls them uranium guzzlers.) Redesigning the next generation of reactors to be more uranium-efficient, even if reprocessing is forbidden in an effort to inhibit weapons proliferation, could extend the useful life of known uranium deposits by a factor of about five. If the estimates of uranium resources and the prospects of more efficient uranium utilization turn out to be correct, the need for the liquid-metal fast breeder reactor could be postponed until well after the year 2000. On the other hand, there may be other breeder reactors on the horizon. A symbiosis between the neutron-rich fusion reactor and the traditionally neutron-starved fission reactor may turn out to be more effective than either one of them alone.

(In the meantime there is another substantial untapped resource of fission fuel that should not be neglected. It has recently been estimated that the selective "mining" of some of the highly enriched uranium and plutonium currently stockpiled by the U.S. in the form of obsolete nuclear weapons could yield enough fuel to satisfy the demand of a significant fraction of the nation's nuclear power reactors.)

Useful as it is to estimate the future availability of uranium at this stage, we are convinced that geologists and geochemists will have to make even better estimates in the future not only for uranium but also for a wide range of other mineral resources. For some metals, such as copper, the U.S. is rapidly going to lower ore grades, whereas for others, such as lead, the ore grades have changed very little. Only by estimating in advance the likely tonnages of material available at each ore grade can geologists help an industrialized society to evolve without the kinds of shocks that abruptly increasing fossil-fuel prices have brought in the past few years.

EXCEL! EXCEL!

Many reach for the Olympic gold but only a few get it. And that's the reason for having the Olympics. We don't *have* to support the Games. We just want to. So we do. The Committee has made us the Official Photographic Consultant to the 1980 World Olympics. That means that every day of the Winter and Summer Games, from 8 a.m. to 2 a.m., our people and our equipment will be processing film for accredited press.* No charge. Our contribution to world peace regardless of where the Games are held... our tribute to youthful excellence. Photography in the hands of superbly skilled professionals has the assignment to give the world's peoples the feeling of being there in the best possible spot seeing everything clearly at the decisive moment. Between the hands on the camera and the eyes of the world, Kodak can make itself useful. We consider the Olympic Games vitally important. We would. We're just that kind of people. We feel kinship with others who value excellence, wherever in the world they live and for whomever they cheer. Incidentally, even if you are not a professional photographer, we want your pictures of sports events to look as though you were on your way to becoming one. Drop a postcard to Dept. 55W, Kodak, Rochester, N.Y. 14650 asking for Kodak Publication No. A99-2, "Sports Photography—Tips and Techniques."

KODAK... OFFICIAL PHOTOGRAPHIC CONSULTANT TO THE 1980 WINTER AND SUMMER OLYMPIC GAMES

Kodak

*And for medical staff calling for x-rays when tough luck has struck.

SCIENCE AND THE CITIZEN

Day in Court

Although in recent years I.Q. tests have been increasingly challenged as a measure of human abilities, they have continued to be widely given. In California the scores on I.Q. tests have provided school authorities with a basis for assigning children to special classes for the "educable mentally retarded." In these classes some 25 percent of the children are black, even though black children constitute only 10 percent of the total school population in California. A Federal judge, acting in a lawsuit brought against the state's department of education on behalf of black children alleged to have been wrongly placed in the classes, has now enjoined the state from employing I.Q. tests in deciding who should be put in them. He ruled that the tests "are racially and culturally biased, have a discriminatory impact against black children and have not been validated for the purpose of essentially permanent placements of black children into educationally dead-end, isolated and stigmatizing classes for the so-called educable mentally retarded."

The suit was originally filed in 1971 against the San Francisco school system. It challenged the use of the tests as violating the Federal and state constitutions by denying black children the equal protection of the law. In 1974 the court expanded the class action to include all black schoolchildren in California "who have been or may in the future be classified as mentally retarded on the basis of I.Q. tests."

The recent decision was issued by Robert F. Peckham, chief judge of the U.S. District Court for the Northern District of California. "We must recognize at the outset," he wrote, "that the history of the I.Q. test, and of special education classes built on I.Q. testing, is not the history of neutral scientific discoveries translated into educational reform. It is, at least in the early years, a history of racial prejudice, of Social Darwinism and of the use of the scientific 'mystique' to legitimate such prejudices." Acknowledging that black children generally score lower than white children on the standardized tests, Peckham considered and rejected the argument that the reason could be a genetic difference in intelligence. The reason, he said, is that the tests "were standardized and developed on an all-white population, and naturally their scientific validity is questionable for culturally different groups."

Peckham also considered the argument that success in the U.S. culture depends on the acquisition of skills the I.Q. tests can measure. The logic of the argument, he said, does not fit the facts. "To give up so easily on black children is simply not justified."

The judge noted that a different educational philosophy, partly reflected in a master plan for education being developed in California, is gaining adherents. Instead of labeling a child and putting him in a special class, "it aims to diagnose learning problems, to take the child as he or she is and to encourage remedial action through an individualized educational program." It seems likely that the California decision, invoking the Federal constitution, will affect school systems elsewhere in the country that assign children to special classes on the basis of scores on I.Q. tests.

New Math

Advocates of greater U.S. military spending have relied increasingly in recent months on two economic indicators that are often cited in conjunction to project an alarming image of the country's future defense capability. According to one mode of analysis, the portion of the total Federal budget devoted to military programs has been steadily declining for a number of years, largely at the expense of rising social-welfare costs. At the same time, the argument holds, the military expenditures of the U.S.S.R. have been increasing sharply and now surpass comparable U.S. expenditures by a wide margin. The statistical basis of both assertions has been called into question.

In its present form the U.S. military budget for the fiscal year 1980 (counting only direct Department of Defense line items) stands at approximately $122 billion, a gain of more than $30 billion (or roughly 34 percent) over the corresponding amount for 1976, the year President Carter was elected. Even when U.S. military expenditures are measured in constant, or inflation-adjusted, dollars, they have increased every year since 1976. (The issue now before Congress is whether the fiscal-year 1981 military budget should be raised by 3 or 5 percent over and above the rate of inflation, to a new high in current dollars of nearly $140 billion.)

Nevertheless, it has been stated publicly by Secretary of Defense Harold Brown, among others, that in the coming year "U.S. defense will... still receive the lowest fraction of our resources in 40 years." President Carter too has been quoted as saying that "over the last number of years, including since I have been in office even, the percentage of our total budget... that goes into defense has been decreasing."

The apparent contradiction of a budgetary trend that is simultaneously increasing and decreasing is explained by the substitution of relative terms for absolute ones: Any increasing series can be made to appear as if it is decreasing simply by expressing it as a fraction or percentage of some other series that is increasing faster. In the case of the increasing U.S. military budget the appearance of a decreasing trend is usually achieved by comparing it either to the rise in the total Federal budget or to the growth of the country's gross national product (G.N.P.).

The questionable validity of this accounting procedure has been pointed out by a number of individuals and groups opposed to excessive military spending, including the Center for Defense Information, a nongovernmental research organization based in Washington. Writing in *The Defense Monitor,* the center's monthly newsletter, James J. Treires, the principal economic analyst on the staff, gives several reasons for concluding that the declining-ratio method of accounting for the cost of the U.S. military establishment "conceals more than it reveals." With reference to a frequently reproduced chart prepared by the Department of Defense in which two curves representing Federal outlays, one labeled "DOD military" and the other "nondefense," are compared for the years 1950 through 1980, Treires comments: "First, it counts veterans' benefits as a 'nondefense' item. While these rapidly growing expenditures do not buy any current military capability, they are a direct result of past military efforts, and like the pensions of retired military personnel, which *are* included in the military budget, are simply deferred military costs.

"Second, and perhaps more important, the Pentagon's chart itself is based on the 'unified budget,' which adds together the social-insurance programs funded through employee contributions and the programs [such as] 'national defense' that are supported entirely out of Federal income and other general taxes. The 'DOD military' line represents an expenditure financed entirely by the general taxpayer and in direct 'competition' only with other programs similarly funded. The 'nondefense' line includes the huge social security, unemployment insurance, medicare and highway and airport programs.

"It is also inaccurate to include interest on the national debt, a $57-billion item in the FY 1980 budget, as a 'nondefense' expenditure. About two-thirds of this debt was incurred through military spending."

An "objective chart" of recent trends in Federal spending, Treires suggests, by eliminating all trust-fund programs and other irrelevant comparisons, would

Jameson Irish

"Let's give them ~~Scotch~~."

If they like fine Scotch, they'll love light, imported Jameson Irish.

Let them try a glass of Jameson Irish Whiskey the way they would their favorite Scotch.

They'll notice how much it tastes like fine Scotch—only lighter and more delicate.

Chances are, as dedicated Scotch drinkers, they will instantly appreciate this flavor difference.

Though it may take a little time for them to get used to saying, "Jameson Irish on the rocks, please."

Jameson. World's largest-selling Irish Whiskey.

80 PROOF • CALVERT DIST. CO., N.Y.C.

"show, as the President's budget does not, just where our income taxes have been going." Such a chart, he says, would reveal that U.S. military spending has continued its steady upward trend in both current and constant dollars throughout the past decade and "still constitutes the largest single expenditure from general revenues."

The other economic indicator factored into many "worst case" analyses of the relative status of U.S. military expenditures is the annual estimate of Russian military expenditures made by the Central Intelligence Agency. According to Franklyn D. Holzman, professor of economics at Tufts University and an associate at Harvard University's Russian Research Center, "there are many sources of possible exaggeration in the CIA estimates." For one thing, he explains in a recent article in *The New York Times,* "the CIA's published comparisons are always in dollars. Prices expressed in dollars exaggerate Soviet expenditures.... This is because the Soviet armed forces have twice the personnel of America's but add only a little more new equipment each year, and because, in the words of the Director of Central Intelligence, Adm. Stansfield Turner, 'In the United States manpower is relatively more expensive than hardware [whereas] in the Soviet Union military hardware is much more expensive than manpower.'" By valuing the manpower costs of the Russian armed forces at the wage scales of the U.S. armed forces, the CIA analysts arrive at an official dollar estimate of the military budget of the U.S.S.R. that is some 40 percent higher than the U.S. figure.

Conversely, Holzman continues, "a ruble comparison... exaggerates American expenditures. This is because our armed forces have more equipment per person than the Soviet forces and because equipment is relatively high-priced in the Soviet Union." In response to Congressional questioning the CIA some time ago prepared an unofficial comparison in rubles that put the military expenditures of the U.S.S.R. for 1977 at 25 percent more than those of the U.S. Although the downward adjustment in the estimated level of military spending by the U.S.S.R. may have satisfied Congressional interrogators, Holzman reports, "it did not satisfy economists used to such... comparisons. Experience has shown that ruble-dollar differentials typically exceed 50 percent. Clearly, then, if the Soviet Union outspends the United States in dollars by 40 percent, one would expect the United States to equal or outspend the Russians in rubles. The CIA's figures, therefore, are highly suspect."

The main reason the CIA gives for not publishing an official ruble estimate is that much of the equipment deployed by the U.S. is beyond the current capability of Russian technology and therefore cannot be given an exact ruble price. Instead the CIA's unofficial ruble comparison values U.S. high-technology equipment at ruble prices that are "applicable to the closest substitute goods that can be produced in both economies," a procedure the CIA concedes is merely a "rough approximation." Holzman observes: "No wonder American defense expenditures priced in rubles are estimated at less than the Russians' defense expenditures. If a properly high ruble price tag could be put on our high technology, the American defense package would certainly cost the Russians more to produce than their own. It might well be that they cannot produce our defense package at any cost." His article concludes: "The major fallacy in the CIA procedure is that the very dimension of the arms race in which America has the greatest advantage—advanced technology—and which makes most of the difference between military superiority and inferiority is enormously undervalued.... All things considered, the CIA's categorical conclusion that the Soviet Union is outspending the United States on defense may well be a figment of its faulty methodology."

One indication that the official dollar estimates of Russian military expenditures put out by the CIA are found to be unsatisfactory even within the Administration is contained in the Arms Control and Disarmament Agency's annual report *World Military Expenditures and Arms Transfers,* which notes in an introductory essay in the current edition that "alternative methods of calculating Soviet military expenditures in dollars are being considered for future editions of this publication." Why are such misleading comparisons given wide circulation in the first place? "Perhaps," remarks David T. Johnson, research director of the Center for Defense Information, "it is because the strictly military rationale for many military programs is obscure or dubious that high levels of military expenditure are frequently defended on other grounds."

Fascinatin' Algorithm

The ideal way to plan the production of an automobile plant, schedule workers on an assembly line or formulate a military strategy is to take all the relevant factors into account. Hence in the complex modern world some of the most frequently encountered computing tasks are those that seek to find the optimum value of a function whose variables are subject to a large number of constraints. If these limiting factors can be expressed as linear, or first-degree, equations, such tasks are the province of linear programming. Linear-programming problems arise continually in fields such as economic modeling and business planning, where deciding how to allocate manpower and resources for minimum cost or maximum profit typically involves hundreds or thousands of variables and constraints. On even the fastest modern computers the cost of performing such complicated computations can be enormous, and so there is considerable excitement at present among pure and applied mathematicians concerning a new way to approach problems of this kind. L. G. Khachian, a mathematician at the computer center of the Academy of Sciences of the U.S.S.R., has devised a new algorithm for solving linear-programming problems that, in principle at least, is significantly more efficient than any procedure previously known.

In order to appreciate the importance of Khachian's discovery it is necessary to understand how mathematicians classify different methods of solving computing problems. The efficiency of a solution procedure is determined according to how in the worst possible case the number of computational steps (and thus the time) required to implement the procedure increases as a function of the size, or complexity, n of the problem to which it is to be applied. In particular, a solution procedure is considered to be efficient if the number of steps required to implement it grows as a polynomial function of n (say n^2). According to this theoretical criterion, an example of an inefficient procedure is one where the number of steps grows as an exponential function of n (say 2^n). An exponential function increases far more rapidly than a polynomial one, so rapidly that when a problem is only moderately large, a solution requiring exponential time may take millions of years to implement on the fastest computer. This distinction is important in terms of linear programming because until Khachian's algorithm came to light the best procedure known for solving linear-programming problems, a procedure known as the simplex algorithm, could not be guaranteed to provide a solution in polynomial time.

More precisely, the simplex algorithm, which was devised in 1947 by George B. Dantzig of Stanford University, is based on a geometrical interpretation of linear programming in which each of the constraining equations in a problem is visualized as a plane in multidimensional space. It can be shown that the intersection of all the constraining planes forms a polyhedron, or multidimensional multifaceted box, whose vertexes correspond to all the potential optimum solutions of a function whose variables are subject to those constraints. To minimize or maximize the function, then, the simplex algorithm moves from vertex to vertex, calculating the function and successively choosing better solutions until an optimum one is finally attained.

In principle it may be necessary to check most or all of the vertexes of

For 1980, Mercedes-Benz introduces the most relentlessly efficient automobiles in its history.

Stringent efficiency is no sudden demand to the engineers of Mercedes-Benz. For 94 years they have built cars with little else in mind.

And for 1980 they have inched the standards up another notch. Without downsizing bodies or engines, without slashing weight or cutting corners in quality or safety or comfort, they have created the most relentlessly efficient Mercedes-Benz automobiles yet.

Diesels—and more power to them

In any ordinary year, it would be major news that Mercedes-Benz engineers had boosted the performance of their Diesel-powered cars.

And boost it they did—to a healthy degree. The muscular 300 SD Turbodiesel is even more muscular. The five-cylinder 300 TD Station Wagon, 300 CD Coupe and 300 D Sedan move more briskly. The 240 D Sedan enjoys new punch.

But 1980 is no ordinary year. The best news is that the legendary Diesel fuel efficiency remains legendary.

The pleasant shock of efficiency

All six Mercedes-Benz gasoline models remain on the leading edge of automotive technology for 1980.

In these cars for 1980, fuel efficiency gains some makers might be pleased to achieve in two, three or five years have been achieved in one. Advances ranging from a 14.3* percent increase in fuel mileage for the 280 CE Coupe, 280 E and 280 SE Sedans, to a 33.3 percent gain in the 450 SEL Sedan, the 450 SL Roadster and the 450 SLC Coupe. Compare this to other cars. Your mileage may differ depending on speed, weather conditions and trip length.

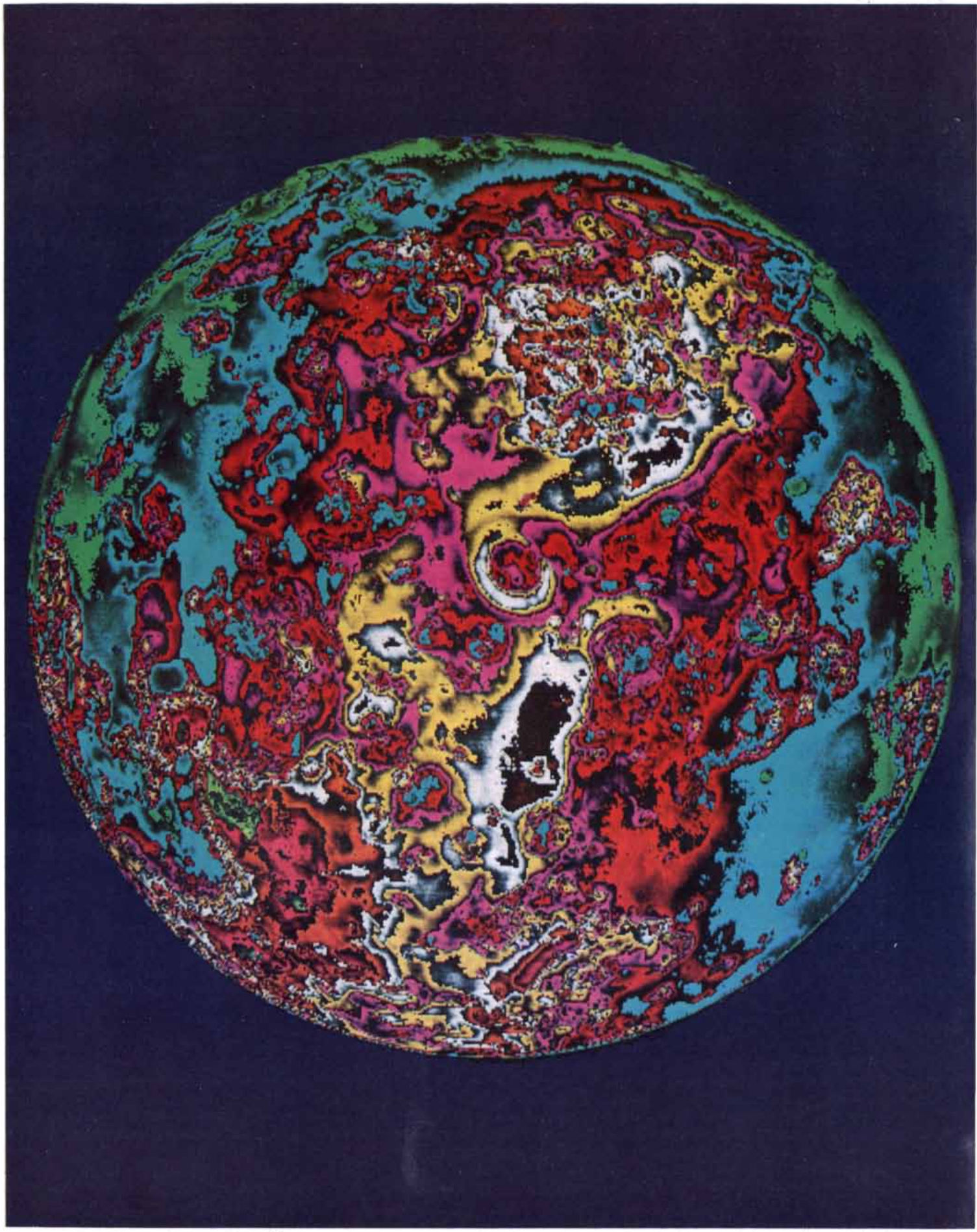
In 1980 as in 1886

Every car maker today speaks of its products as being "right for the times." Mercedes-Benz is no exception.

But there is one difference. Making its cars more efficient does not loom only as "the challenge of the eighties" at Mercedes-Benz. It has *always* been the challenge.

Engineered like no other car in the world

*California estimates vary.



Actual-size reproduction of a Polacolor 8 x 10 hardcopy of Io, moon of Jupiter, taken by Voyager satellite. An example of remote sensing. Photo courtesy of JPL/NASA.

Now, direct from your computer-big instant pictures.

Now your computer can immediately deliver sharp pictures in brilliant colors on Polaroid 8 x 10 Land film.

This is an ideal form of hardcopy. It adds impact to your graphic data by presenting it in a large, easily readable form. And it serves a wide variety of informational needs in business graphics and image processing—with applications in mapping, earth studies from satellites, medical diagnosis, computer-aided design and animation.

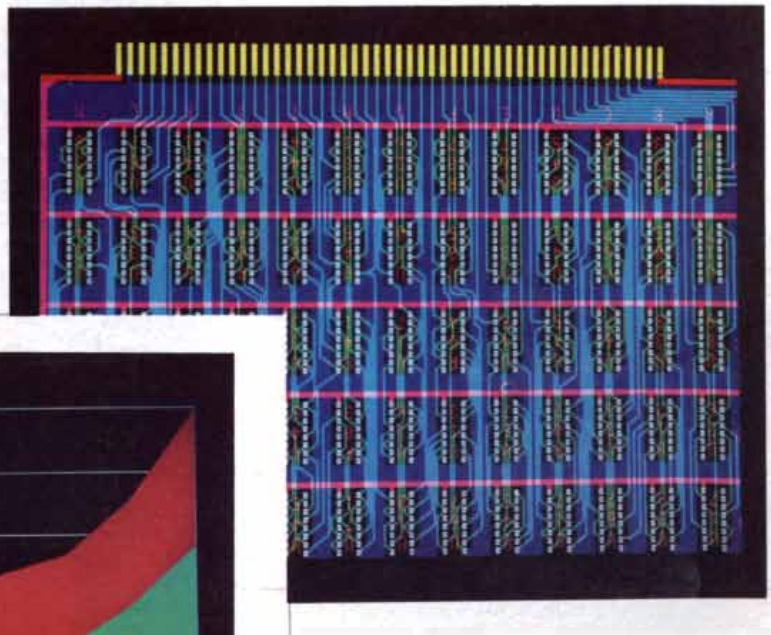
The pictures shown here were made by new computer peripherals—color cameras. They are manufactured by Dunn Instruments and Matrix Corp. Since the cameras use standard video (RS 170) signals, you can conveniently interface them with computers or color-raster based terminals.

The Polaroid 8 x 10 print produced by the camera is the highest-quality instant record available. Since you receive your hardcopy immediately, you can be sure you're getting exactly the record you want, exactly the way you've seen it displayed on your monitor. And you can distribute it at once if the data is urgently needed. The pictures are distortion-free, and their bright, saturated colors yield superb results.

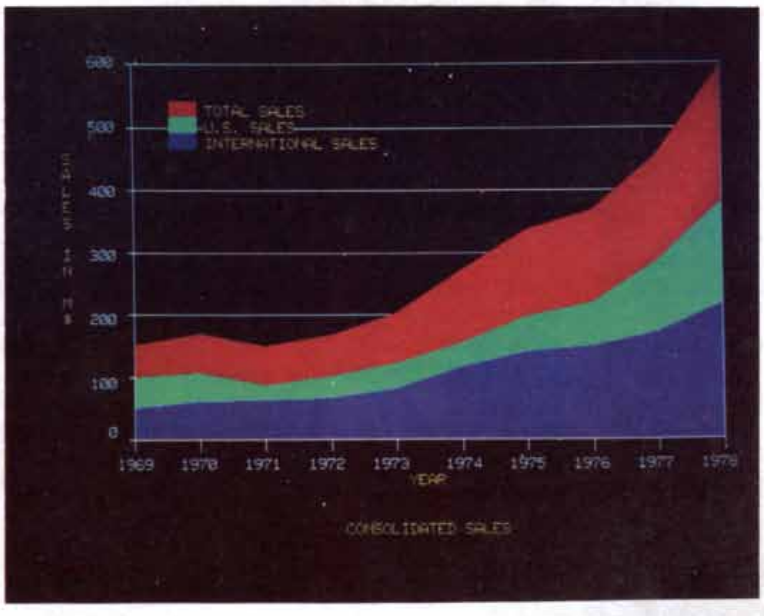
The print can also record continuous-tone images from digital color TV signals in such applications as remote sensing, animation and geological sounding. Its resolution surpasses the resolution of present color digital TV systems. As the technology of digital TV advances, these large Polaroid prints will be able to record even greater detail and more information.

The film's colors are stable, so the picture stays fresh and bright. The photo can be used as original art for high-quality printing. And because of its 8 x 10 size, it can be inserted as a complete page in a report.

For more information or a demonstration, write to: Polaroid Corporation, Dept. A428, 575 Technology Square, Cambridge, Mass. 02139. Or call toll-free from anywhere within the Continental U.S.: 800-225-1618. (In Massachusetts call collect: 617-547-5177.)



Computer-aided design of IC clip from Ramtek color terminal (Reduced from Polacolor 8 x 10 print.)



Typical management information graphic from Tektronix color terminal. (Reduced from Polacolor 8 x 10 print.)

Polaroid
Instant 8x10 color film

the polyhedron (usually an exponential function of the number of variables and constraints in the problem) before an optimum vertex, or solution, is discovered. (Actually mathematicians have constructed pathological examples of linear-programming problems where that is the case.) In most instances, however, only a "reasonable" number of vertexes need to be checked before an optimum one is found. In other words, although the simplex algorithm cannot guarantee a solution to a particular linear-programming problem in polynomial time, in practice the algorithm solves most problems quite efficiently.

Indeed, the simplex algorithm performs so well that mathematicians have long sought a refinement of the algorithm or a new procedure that would be efficient not only in practice but also in principle. No such algorithm was known, however, until October, 1978, when Khachian submitted to *Doklady,* the proceedings of the Academy of Sciences of the U.S.S.R., a note outlining a new algorithm for solving linear-programming problems that was guaranteed to find a solution in polynomial time. Because Khachian is apparently completely unknown in the U.S. it took several months for word of his findings to reach mathematicians here. Finally in the spring of last year translations of the note began circulating among computer scientists in the U.S. and Western Europe. Because the note did not include any proofs, however, there was no way of knowing whether the assertions in it were valid until midsummer, when László Lovász of the University of Szeged and Peter Gács of the University of Rochester sat down to reconstruct the details of Khachian's argument.

The work of Lovász and Gács (who in the course of retracing Khachian's steps were able to simplify his procedure) has given mathematicians a clearer understanding of the new algorithm. Like the simplex method it depends on a geometrical interpretation of linear programming, but one in which each problem is recast to create a larger set of constraints and variables in a geometrical space of higher dimension. Once again every constraint is visualized as defining a plane, but now an ellipsoid, or multidimensional ellipse, is defined with respect to the set of constraining planes. The algorithm proceeds by constructing a series of successively smaller ellipsoids whose centers converge on an optimum solution. The proofs supplied by Lovász and Gács confirm Khachian's assertion that if an optimum solution exists, it will be found by this procedure in polynomial time.

The new algorithm proves that linear programming can be classified as a polynomial-time problem, an important theoretical advance. The practical implications of Khachian's work have yet, however, to be determined. In its present form the algorithm could not be applied effectively to any real linear-programming task. Moreover, the simplex algorithm, whose uses have been refined for more than 30 years, is extremely successful in most programming situations. Some mathematicans believe Khachian's results will eventually yield even more efficient practical techniques for solving linear-programming problems, but a great deal of study and experimentation will be required in order to decide the issue.

Khachian's work may prove to be applicable, however, to some other types of problems with certain computational similarities to linear programming. In particular, the practical strengths and theoretical limitations of the simplex algorithm had prompted mathematicians to conjecture that linear-programming problems might belong to a special class of problems falling somewhere in between the class of "easy" problems (those with polynomial-time solutions) and the class of very difficult computing problems that are called nondeterministic polynomial-time complete, or NP-complete. It is generally believed that such problems require exponentially increasing amounts of time, and in addition it has been established that if any one of them can be shown to have a polynomial-time solution, they all have such a solution. Khachian's algorithm demonstrates that in theoretical terms linear-programming problems are easy, not NP-complete or even close to NP-complete, which raises the possibility that other problems thought to fall between the two classes can also be placed in one class or the other. Khachian's approach may offer an indication of how to tackle these problems and may even give some insight into the nature of the highly intractable NP-complete problems themselves.

Cloud Burst

For 10 years spacecraft equipped to detect gamma rays, notably the Vela satellites designed to monitor clandestine nuclear explosions, have recorded an average of nine or 10 mysterious bursts of gamma rays per year from regions outside the solar system but presumably within the galaxy. None has ever been associated with anything visible in an astronomical photograph that could reasonably serve as the source of such a burst. A burst recorded on March 5, 1979, however, was different from all the others. Its intensity was 10 times higher than any of the 90-odd bursts previously observed, and it was recorded simultaneously by 11 sensors aboard nine spacecraft, six American and three Russian, that are part of a recently formed interplanetary gamma-ray-burst network.

Although each of the gamma-ray sensors is by itself omnidirectional, the arrival time of the gamma-ray wave front at each of the spacecraft is accurately recorded, and from that information the direction of the source can be calculated by triangulation. Because the distance separating some of the detectors was more than 150 million kilometers the direction of the March 5 event could be fixed in a region of space only two minutes of arc across. Within that region a candidate source was found for the first time in the 10-year history of such events: a supernova remnant known as N49 in the Large Cloud of Magellan, a small galaxy close to our own (about 180,000 light-years away). Thus the inferred source is roughly 100 times more distant than the still unknown sources within our own galaxy where all previous gamma-ray bursts were believed to originate.

Since the March 5 event was 10 times stronger than any previous burst and seems to have originated from a point 100 times more distant, its energy output presumably exceeded that of previous bursts by a factor of 100,000 or even a million. If it is assumed that the energy was radiated equally in all directions, one can compute that for a tenth of a second gamma rays issued from the source at the rate of 10^{45} ergs per second, equivalent for that brief instant to the output of visible radiation from 10 entire galaxies the size of our own.

The second remarkable fact about the March 5 event is that the original giant burst was followed by a series of smaller bursts of diminishing intensity, 22 of which were recorded before the memories of the most capacious recording instruments were saturated. The smaller bursts were spaced precisely eight seconds apart. It is as if the original cataclysmic report "rang" for several minutes. Although no one has yet proposed a satisfactory model to explain the event, it seems likely that the source was a rotating beam of photons rather than a series of explosions that fired photons in all directions. If the beam model is correct, the rate of energy release would of course be much lower than 10^{45} ergs per second.

The nine spacecraft that recorded the event were: *Helios-2* (in interplanetary orbit around the sun), the Pioneer-Venus orbiter (in orbit around Venus), the International Sun-Earth *Explorer-3* (in orbit around the sun), three Vela satellites (in orbit around the earth) and three Russian vehicles, *Venera-11* and *Venera-12* (both in orbit around the sun after passing Venus) and *Prognoz-7* (in orbit around the earth). The initial report, soon to appear in *The Astrophysical Journal,* was submitted by an international group from the Goddard Space Flight Center of the National Aeronautics and Space Administration, the Los Alamos Scientific Laboratory, the Institute for Space Research in Moscow and the Centre d'Étude Spatiale des Rayonne-

Because you enjoy going first class.

This season, give the Scotch that's always in good taste. Passport. Made of Scotland's finest whiskies and honored in 126 countries around the world. It's always more rewarding to give a first-class gift.

Passport Scotch.

ments in Toulouse, which built the instruments on the Russian spacecraft.

Opportunistic Agent

The mysterious Legionnaires' disease epidemic of 1976 was eventually shown to have been caused by a previously unrecognized bacterium, *Legionella pneumophila,* with unusual staining and culturing characteristics, that turned out to have been responsible for other outbreaks of pneumonia before and after the one in Philadelphia that brought it to world attention. Last year the authors of a report on the disease and its agent remarked that for about a third of all the cases of pneumonia recorded in the U.S. each year the cause is not known, and they speculated that some of the unknown agents might be bacteria related to *L. pneumophila* (see "Legionellosis," by David W. Fraser and Joseph E. McDade; SCIENTIFIC AMERICAN, October, 1979). Now it appears this may well be the case. Another "new" pneumonia has been reported, and its agent has been shown to be a bacterium that has some features in common with the agent of legionellosis.

The new bacterium was first reported in *The Lancet* last July by A. William Pasculle, Richard L. Myerowitz and Charles R. Rinaldo, Jr., of the University of Pittsburgh School of Medicine and Presbyterian–University Hospital. A team of physicians, microbiologists and pathologists had been seeking to identify agents responsible for "opportunistic" pneumonias seen in patients undergoing therapy by drugs, such as corticosteroid hormones, that depress the cellular immune response. The team isolated from the lung tissue of two kidney-transplant patients with pneumonia a microorganism that had an unusual combination of characteristics: in staining tests it was both gram-negative and "acid-fast." Most acid-fast bacteria are gram-positive and give rise to lung-tissue changes rather different from the pathology observed in the two patients.

The newly identified agent was similar to *L. pneumophila* in that it had a structure characteristic of gram-negative bacteria, caused an acute pneumonia, could be isolated only in embryonated eggs and guinea pigs, could not be cultured on the standard bacteriologic culture mediums and could be visualized in lung tissue by the Dieterle stain. The new agent differed from *L. pneumophila* in some ultrastructural details and serologically (it failed to react with antiserums directed against the Legionnaires' disease bacterium), and it could not be cultured even on some modified mediums developed for culturing *L. pneumophila.* It was apparently a previously unrecognized bacterium, and it was designated the "Pittsburgh pneumonia agent" (PPA).

Now a total of 13 cases of PPA pneumonia have been reviewed in *The New England Journal of Medicine,* eight cases by Myerowitz and Pasculle and their colleagues at Pittsburgh and five cases by Beverly H. Rogers and her colleagues at the University of Virginia Hospital in Charlottesville. Ten of the 13 patients had received kidney transplants and were receiving steroids and other immunosuppressive drugs to combat graft rejection; the others had been under similar therapy for different reasons. In each case an acute, purulent bronchopneumonia developed soon after the administration of high doses of corticosteroids. The white blood cells in the alveoli of the lung were largely polymorphonuclear leukocytes rather than macrophages; granulomatous inflammation, which is usually seen in infections by acid-fast bacteria (and the presence of which is usually the only reason for doing an acid-fast staining test), was not present. Seven of the 13 patients died; as Morton N. Swartz of the Massachusetts General Hospital points out in a *New England Journal* editorial, the apparent mortality could be exaggerated because some cases were identified only by reviewing autopsy records and material.

The diagnosis of PPA pneumonia was established by isolation of the bacteria from the lung, by visualizing the bacteria in lung tissue and, in a number of cases, by indirect fluorescent-antibody tests in which the patients' blood serum was shown to react with PPA. The fact that the infection developed during hospitalization or just after discharge suggests that the bacterium was encountered in the hospital, but it is also possible that a latent infection was activated by the suppression of the immune response. Swartz points out that just as legionellosis was found to have antedated the Philadelphia outbreak, so a review of tissue samples from patients—and particularly patients who have received immunosuppressive drugs—with undiagnosed pneumonias may show that PPA pneumonia "has been around for some time." There are indications that the discovery of *Legionella pneumophila* was the first recognition of a new family of bacteria and that PPA is one of the relations.

Quailing before a Quail

Since ancient times it has been known in Europe and North Africa that the European migratory quail, *Coturnix,* was good eating but that eating it sometimes made the eater deathly ill. More recently it has been established that the cause of illness is a unique kind of food poisoning now named coturnism. The symptoms include difficulty in breathing, impaired speech, nausea, weakness and loss of feeling in the legs and partial paralysis; death may result. The reason the explicit connection between these symptoms and eating quail took so long to establish is that only certain of the birds are toxic to human beings at certain times of the year and in certain parts of their migratory range, so that the symptoms were often attributed to other causes.

What makes some of the birds poisonous some of the time? Two students of the question, Louis E. Grivetti, a nutritionist at the University of California at Davis, and his research associate, Bruce W. Kennedy, have been able to eliminate one long-lived suspicion. This was that the quail became toxic after they had eaten quantities of seeds rich in poisonous alkaloids, particularly those of the poison hemlock *Conium,* a member of the carrot family. When European quail were given a choice of food, the investigators found, they did not readily select hemlock seeds. Moreover, during the birds' migration from equatorial Africa to northern Europe (February through April) and their return south (August through October) hemlock is rarely in fruit along the flyways. And even when the quail were fed quantities of hemlock seed in the laboratory, they did not concentrate the alkaloids but excreted them, usually within 24 hours.

In a report for the British journal *Ecology of Food and Nutrition* Grivetti and Kennedy suggest that the stress of migratory flight may induce a buildup of some unidentified toxic metabolite in the bird's body. For example, the European quail's westernmost flyway brings it north to Spain and France after a long and tiring passage over the Sahara and the Mediterranean. Case histories of coturnism indicate that quail flying north on this route are sometimes toxic and quail flying south are never toxic. On the other hand, the easternmost flyway of the quail brings them north down the Nile valley and across the Levant and Turkey into the western U.S.S.R., and case histories of coturnism in the Greek islands and the area north of the Black Sea indicate that none of the birds making the northward passage appear to be toxic, whereas some of the southbound migrants are toxic on their return to the eastern Mediterranean. Why the early part of this southbound migration might be comparable in stress to the northbound passage across the Sahara and the Mediterranean remains unexplained.

Another variable is the way the birds come to market. For example, in Egypt the quail are netted and cooped for some time before they are sold for the table. The delay would allow ample time for any stress-induced toxin to be eliminated from the bird's body before the bird was eaten. In other areas, where private and market hunters bring fresh-killed birds to the table, no such elimination can take place. To test the stress hypothesis further Grivetti and Kennedy intend to collect fresh-killed northbound and southbound quail for laboratory analysis.

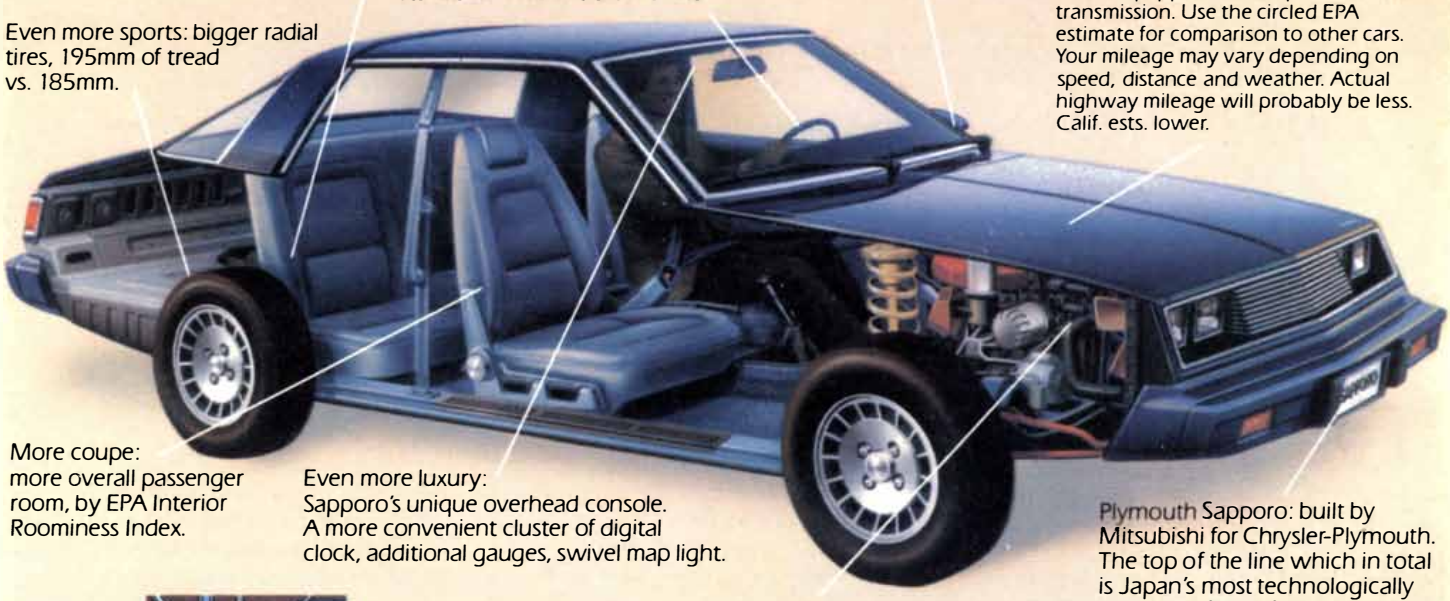
More sports: optional 4-wheel power disc brakes—power front disc/rear drum system is all Celica offers.

More luxury: power steering standard. $165 extra on Celica.

More luxury: dual electric remote-control racing mirrors, not available on Celica.

(21) EPA est. MPG/32 highway MPG
When equipped with 5-speed manual transmission. Use the circled EPA estimate for comparison to other cars. Your mileage may vary depending on speed, distance and weather. Actual highway mileage will probably be less. Calif. ests. lower.

Even more sports: bigger radial tires, 195mm of tread vs. 185mm.



More coupe: more overall passenger room, by EPA Interior Roominess Index.

Even more luxury: Sapporo's unique overhead console. A more convenient cluster of digital clock, additional gauges, swivel map light.

Plymouth Sapporo: built by Mitsubishi for Chrysler-Plymouth. The top of the line which in total is Japan's most technologically advanced fleet of passenger cars.

More sports: Sapporo's 2600cc MCA-JET engine. Sportier performance than Celica's 2200cc. Clean air and terrific mileage, too, because a unique jet valve maximizes combustion.

Sports and luxury: Silent Shaft, another Mitsubishi exclusive. Counterbalancing shafts reduce engine vibration for a smooth ride.

PLYMOUTH SAPPORO IS A LUXURY SPORTS COUPE FROM JAPAN, JUST LIKE CELICA EXCEPT SAPPORO HAS MORE LUXURY, MORE SPORTS AND MORE COUPE.

CHRYSLER
Plymouth

Plymouth imports by Mitsubishi. Japan's most technologically advanced passenger cars.

The Galilean Moons of Jupiter

In viewing the four largest Jovian satellites last year the Voyager spacecraft increased to nine the number of earthlike bodies that can be closely compared to gain an understanding of how they evolved

by Laurence A. Soderblom

One of the most spectacular scientific adventures of all time began on March 5 of last year. In a period of some 30 hours the spacecraft *Voyager 1* flew past the giant planet Jupiter and returned closeup pictures of three of the planet's four largest moons: Io, Ganymede and Callisto. First observed by Galileo in 1610, the four moons are commonly called Galilean. Detailed pictures of the fourth Galilean satellite, Europa, were made later (on July 9) by *Voyager 2,* which also explored the hemispheres of Ganymede and Callisto that had not been visible to its sister craft. Perhaps the most dramatic discovery made by *Voyager 1* was that on Io volcanic eruptions were in progress. *Voyager 2* therefore devoted nearly 10 hours to observing Io in a "volcano watch" inspired by the earlier pictures.

Both spacecraft also had distant views of Amalthea, an asteroidlike body discovered only 88 years ago and until recently thought to be Jupiter's innermost satellite. Voyager images have now disclosed, however, that the planet has another tiny satellite near the outer edge of a faint ring resembling one of Saturn's rings. This 14th satellite of Jupiter, which is only a few dozen kilometers in diameter and is about halfway between the surface of the planet and the orbit of Amalthea, was discovered by David Jewitt and G. Edward Danielson of the California Institute of Technology. It has temporarily been designated 1979J1.

Amalthea and the four Galilean moons (and probably 1979J1 as well) travel in circular orbits lying in Jupiter's equatorial plane and therefore constitute the "regular" satellite system of the planet. The remaining eight confirmed Jovian moons are much smaller and move in irregular orbits scattered well beyond those of the inner six. From close examination of the several hundred pictures of the Galilean satellites and Amalthea returned by the two Voyagers one can infer a great deal about their histories, relative ages and the nature of the geological processes by which they have evolved.

The use of the term satellites to describe the Galilean moons understates their importance to students of planetary science. Io, Europa, Ganymede and Callisto belong to the family of objects designated terrestrial, a family that includes Mercury, Venus, the earth, the earth's moon and Mars. The Galilean satellites are similar to the bodies of the inner solar system in both size and composition, so that both groups of objects should have evolved by comparable processes and on similar time scales. As a result of the Voyager missions the number of earthlike objects with which one can test theoretical models of planetary evolution has doubled. One can now appreciate that the planets of the inner solar system occupy only a small part of the spectrum of characteristics and evolutionary possibilities open to such objects.

It was known from studies done with telescopes on the earth that the Galilean satellites exhibit certain regular trends, such as decreasing density and increasing size with distance from Jupiter, trends remarkably like those in the solar system as a whole. Of Amalthea little can be told from the earth except that it is no more than a few hundred kilometers in diameter, dark and red. Io, the next satellite out from Jupiter, was known to have about the same size and density (3.5 grams per cubic centimeter) as the earth's moon. Unlike the other three Galilean satellites, Io shows no trace of water in its infrared reflection spectrum. Because of its brilliant orange red appearance and the steep drop in its reflection spectrum toward the ultraviolet it was suspected of having a surface rich in sulfur.

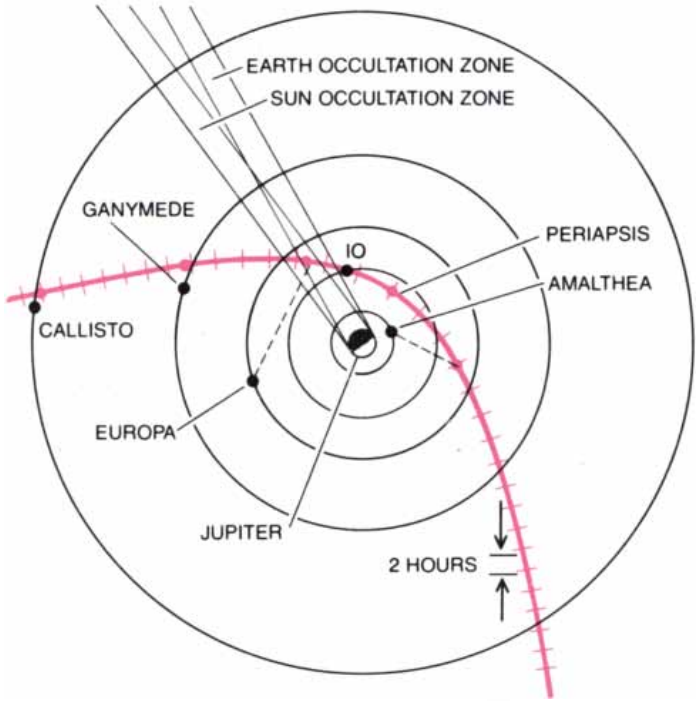
Europa, second out from Jupiter of the large Galilean satellites, is also about the size of the earth's moon but is far brighter: it reflects nearly 70 percent of the sunlight striking it compared with 7 percent for the moon (and 35 percent for the earth). Spectra made at near-infrared wavelengths indicated large amounts of water ice on the satellite's surface. Estimates of Europa's density (about three grams per cubic centimeter) suggested that it might have a shell of ice and liquid water as much as 100 kilometers thick.

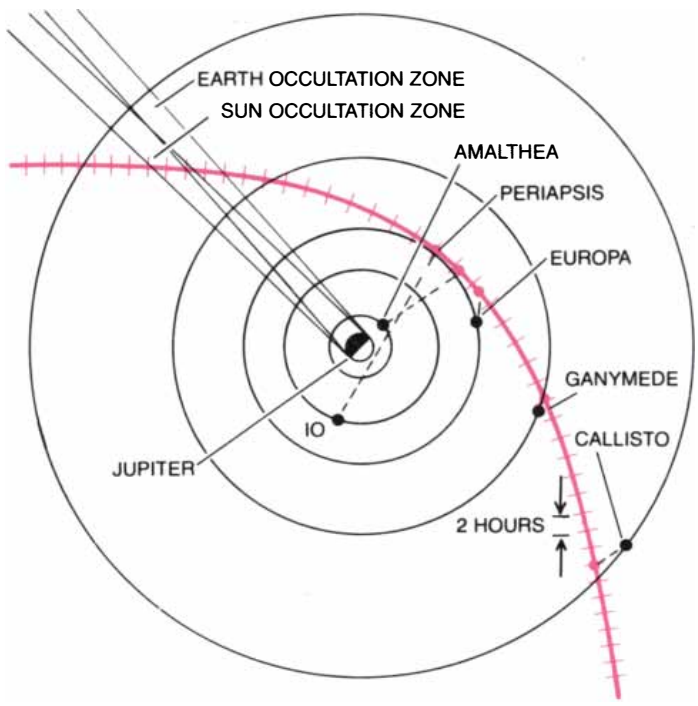
The third and fourth members of the Galilean satellite system, Ganymede and Callisto, are rather similar. Both were known to be about the diameter of the planet Mercury and to have a density of about two grams per cubic centimeter. (Mercury and the earth are about 5.5 grams per cubic centimeter.) One could therefore speculate that Ganymede and Callisto are even richer in water than Europa is, assuming that all three consist chiefly of water and typical silicates (silicon oxides). Ganymede, however, reflects 50 percent of the sun's light compared with only 20 percent for Callisto and shows more evidence of water or ice on its surface.

The Voyager Trajectories

As Amalthea and the four Galilean satellites move in their orbits they keep one face constantly toward Jupiter, just as the earth's moon keeps one face toward the earth. Their orbital periods

CALLISTO, the outermost Galilean moon, was photographed on March 5, 1979, at a distance of between 337,000 and 364,000 kilometers by the television cameras aboard the spacecraft *Voyager 1*. The other three Galilean moons, in the order of their distance from Jupiter, are Io, Europa and Ganymede. Jupiter has 14 moons in all, including a tiny new one discovered by *Voyager 2* and temporarily designated 1979J1. The new moon circles Jupiter within the orbit of another small moon, Amalthea, which in turn circles the planet inside the orbits of the Galilean satellites. The orbits of the remaining eight Jovian moons lie outside those of the Galilean satellites. Callisto is about the size of Mercury. It circles Jupiter once every 16.69 days at a distance of 1.8 million kilometers. In this mosaic the resolution, equivalent to the width of a pair of television scan lines, is seven kilometers. The circular feature near the left limb, or edge, is about 600 kilometers across. Concentric rings extend outward for another 1,500 kilometers.





TRAJECTORIES OF THE TWO VOYAGERS were selected to optimize their coverage of the Galilean satellites. *Voyager 1* (*left*) went within 277,000 kilometers of Jupiter on March 5. *Voyager 2* (*right*) did not go nearly as close to the planet, passing it at a distance of 650,000 kilometers on July 9. In these diagrams each satellite is at the position it occupied when spacecraft made its closest approach.

range from 12 hours for Amalthea to 16.69 days for Callisto. The Voyager trajectories took advantage of this synchronous rotation of the Galilean satellites and their rapidly changing positions to optimize pictorial coverage of their surfaces. As *Voyager 1* approached Io the spacecraft took pictures of the satellite's outward-facing and trailing hemispheres, flew under its south pole and proceeded to close encounters with Ganymede and Callisto, viewing their Jupiter-facing hemispheres and flying past them at high latitudes so that their north poles could be examined.

Voyager 2 encountered Callisto and Ganymede before it reached Jupiter and hence was able to photograph the opposite, outward-facing hemispheres and to view the south-polar region of Ganymede. As a result 80 percent of the surface of Ganymede and Callisto were covered with a resolution of about five kilometers or better. (Resolution is defined as the width of two television scan lines.) *Voyager 2* also photographed about a fourth of the surface of Europa with a similar resolution. Since Amalthea and Io revolve around Jupiter rapidly compared with the time each spacecraft was near the planet, they could be photographed at all longitudes with intermediate resolution (about 20 kilometers per line pair). Io's short orbital period (1.77 days) also made it possible for the two Voyagers to inspect the satellite repeatedly at many longitudes during the few days near each encounter. This provided a rather complete inventory and characterization of the plumes that marked the volcanic eruptions on the satellite.

The most exciting aspect of the satellites of Jupiter is what can be learned from comparing them. A great variety of geological processes and evolutionary rates are dramatically recorded in their appearance. I shall now describe Amalthea and the four Galilean satellites, in order of their distance from Jupiter, as revealed by the Voyager missions.

Amalthea

Amalthea is in a class by itself. Although it is smaller than the Galilean satellites by a factor of about 10, it is about 10 times larger than Mars's tiny

SATELLITE	DIAMETER (KILOMETERS)	MEAN DISTANCE FROM JUPITER (KILOMETERS)	ORBITAL PERIOD (DAYS)	BULK DENSITY (GRAMS PER CUBIC CENTIMETER)	MASS (MOON = 1)	CLOSEST APPROACH (KILOMETERS)		BEST RESOLUTION (KILOMETERS PER LINE PAIR)	
						VOYAGER 1	*VOYAGER 2*	*VOYAGER 1*	*VOYAGER 2*
AMALTHEA	155 × 270 (±8)	109,900	.49	?	?	420,100	558,270	7.8	11
IO	3,638 (±10)	350,200	1.77	3.53	1.21	18,640	1,127,920	1	21
EUROPA	3,126 (±10)	599,500	3.55	3.03	0.66	732,270	204,030	33	4
GANYMEDE	5,276 (±10)	998,600	7.16	1.93	2.03	112,030	59,530	2	1
CALLISTO	4,848 (±10)	1,808,600	16.69	1.79	1.45	123,950	212,510	2.3	4

PHYSICAL AND ORBITAL CHARACTERISTICS of Amalthea and the Galilean satellites are tabulated with the closest approach of the two Voyagers and the resolution of the best images. Each spacecraft bore two Vidicon cameras with lenses of different focal length: 200 and 1,500 millimeters. Color images were produced by making sequential exposures through orange, green, blue, violet and ultraviolet filters. Each image consists of 800 scanning lines with 800 pixels (image units) on each line. The two Voyagers returned 35,000 pictures in all. All the pictures in this article were prepared at the Jet Propulsion Laboratory of the California Institute of Technology.

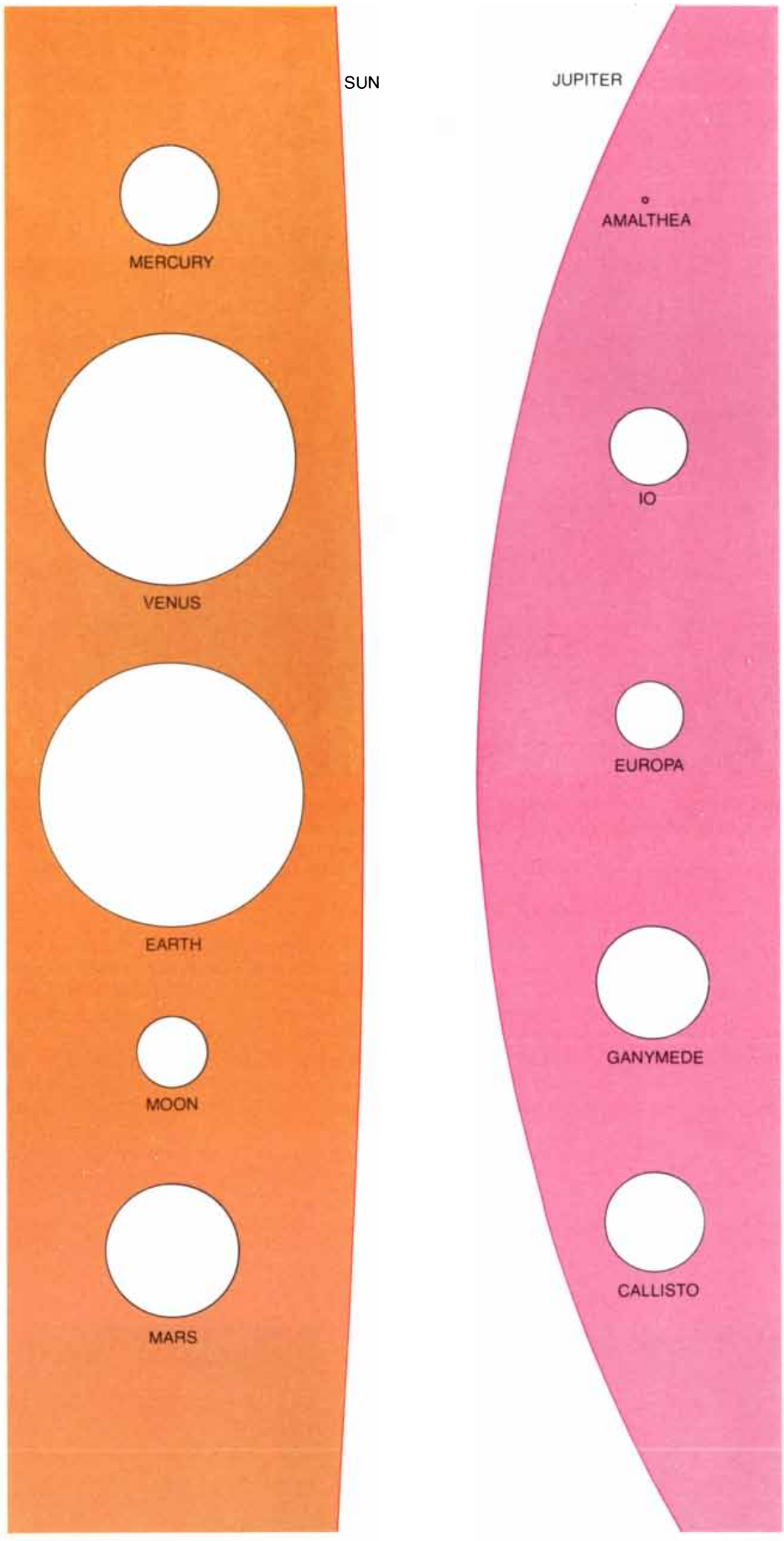
satellites Phobos and Deimos, which have been intensively examined by spacecraft missions to Mars. Amalthea is therefore the first in a class of intermediate-size objects, comparable to many asteroids, to be explored at close range. Of irregular ellipsoidal shape, its long axis (about 270 kilometers long) is pointed at Jupiter and its short axis (155 kilometers long) is at right angles to the plane of its orbit. Voyager imaging data show that it reflects about 50 percent more red light than violet light, confirming observations made from the earth. Although its overall albedo, or reflectivity, is about 5 percent, a few patches are as much as three times brighter [*see bottom illustration on next page*].

Rudolf A. Hanel of the Goddard Space Flight Center of the National Aeronautics and Space Administration and his colleagues on the Voyager infrared experiment discovered that Amalthea is warmer than it would be if it were simply absorbing and reradiating solar radiation and radiation reflected from Jupiter. The excess heat may be generated by electric currents traveling along the lines of force in Jupiter's magnetic field or by the bombardment of particles trapped in the planet's radiation belt. The irregular shape of the satellite implies a substantial internal rigidity and a low content of volatile substances.

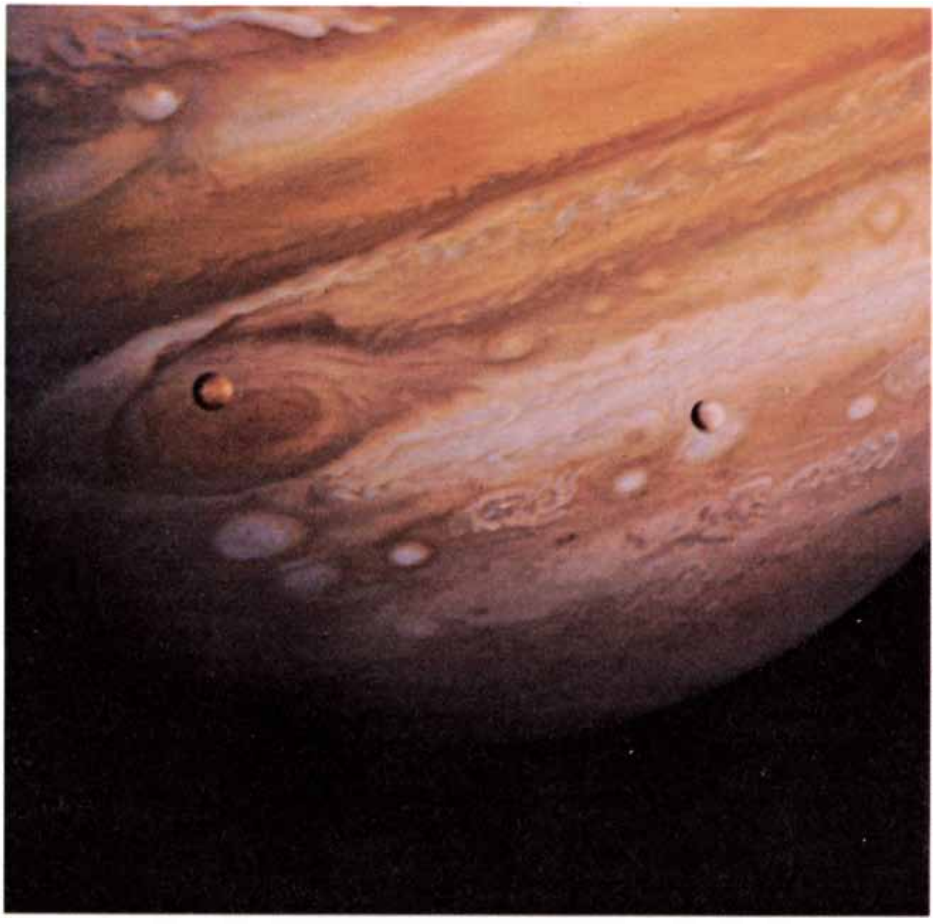
Io

Io, the innermost Galilean satellite, was generally expected to have an ancient cratered surface much like that of the earth's moon. There were, however, a few dissenters from this view. In a striking example of scientific prediction, published just three days before the close approach of *Voyager 1* to Io, Stanton J. Peale of the University of California at Santa Barbara and Patrick M. Cassen and Ray T. Reynolds of the Ames Research Center of NASA pointed out that since the satellite is subjected to resonant gravitational forces exerted by its sister satellites, mainly Europa, its orbit would be distorted out of the circular. Io would move in and out slightly in Jupiter's powerful gravitational field, with the result that it would be repeatedly flexed by tidal forces and an enormous amount of frictional heat would be dumped into its interior. And since the heat would ultimately have to be dissipated through the surface of the satellite, Peale, Cassen and Reynolds speculated, "widespread and recurrent surface volcanism would occur."

The prediction was borne out, but not immediately. The first images of Io with detectable surface markings showed small dark spots, some of which were surrounded by faint rings. The first impression was that these were the expected impact craters. The impression soon had to be revised when pictures of higher resolution showed no impact craters



TERRESTRIAL BODIES that have now been studied at close range in space missions are Mercury, Venus, the earth, the earth's moon, Mars and the four Galilean moons of Jupiter. These bodies are designated terrestrial because of their similarity in size and composition to the earth. The five terrestrial bodies of the inner solar system are drawn to the same scale at the left in the order of their distance from the sun. The shallow arc around them represents the limb of the sun. The Galilean moons are drawn to the same scale against the limb of Jupiter and in order of their distance from the planet. At this scale Amalthea is little more than a dot.



JUPITER, IO AND EUROPA were photographed by *Voyager 1* on February 13, 1979, nearly a month before the spacecraft made its closest approach to the planet. The distance was 21.6 million kilometers; the resolution is 390 kilometers per pair of television scan lines. Io appears against the disk of Jupiter, directly above the Great Red Spot. Europa is at the right.

at all! At maximum resolution (600 meters per line pair) a crater as small as one kilometer in diameter should have been visible. Assuming that on Io the rate of formation of impact craters larger than one kilometer is not drastically different from rates typical for the inner solar system, some process on Io's surface must be obliterating the impact craters in a period of something less than a million years.

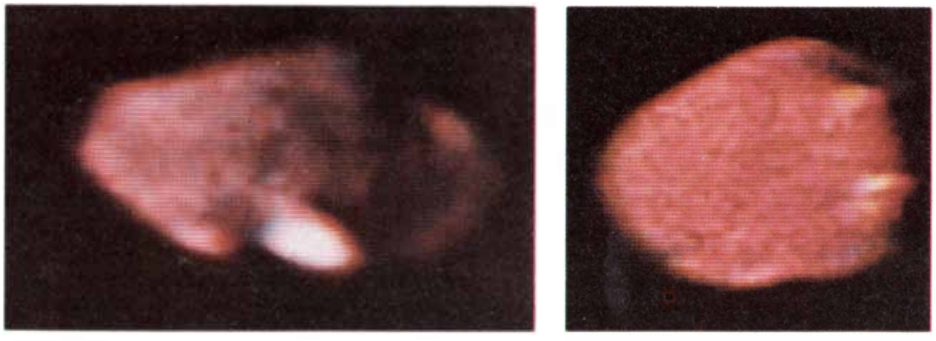
Io's surface shows a host of bizarre land forms, including sinuous scarps and faults [*see bottom illustration on page 95*]. Pictures returned just before *Voyager 1*'s closest approach to Io finally offered a clue to the process that might be renewing the satellite's surface. One such picture shows a circular depression about 50 kilometers in diameter, rimmed by a cliff and surrounded by a radiating pattern of flow [*see top illustration on page 94*]. The feature resembles a volcanic form seen on other terrestrial bodies: a caldera, or large collapse crater, created when lava flows out of or is withdrawn from a magma chamber so that the surface above it collapses. More than 100 such features larger than about 25 kilometers have been identified in the hemisphere seen in high resolution. The flows radiating from the presumed volcanic centers are multicolored: black, yellow, red, orange and brown. To account for the observations members of the Voyager teams have suggested that the flows consist of everything from sulfur-colored basalt to pure sulfur; some have even suggested the existence of sulfur lakes and oceans.

If the surface of Io is so young, and if volcanic activity was as recent as the pictures indicate, that activity should be continuing even today. Those of us on the Voyager imaging team did not dream, however, of detecting active volcanism on Io. The probability of seeing an active volcano on the earth at a comparable distance with cameras like those on the Voyagers is extremely small.

A few days after the encounter of *Voyager 1* with Jupiter, Linda A. Morabito, an engineer at the Jet Propulsion Laboratory of the California Institute of Technology, was examining some images of Io taken at long range for the purposes of navigation. The pictures had been overexposed to bring out the stars in the field behind the satellite. As Morabito was examining one picture she noted a large bright form shaped like an umbrella off the limb, or edge, of Io's southern hemisphere. When efforts to explain away the form as an artifact failed, the conclusion was reached that it was real. Evidently an enormous cloud was arching 270 kilometers above the surface of Io [*see top illustration on page 95*].

After the reality of the cloud was accepted a search revealed that there were no fewer than eight active volcanoes on Io throwing up plumes from 70 to 300 kilometers high with velocities ranging up to a kilometer per second. Some of the plumes are highly symmetrical. In most instances the ejected material rises to a height of about 100 kilometers and forms the umbrella-shaped cloud as it spreads out and plunges back to the surface. Some of the plumes exhibit a large diffuse globe in the ultraviolet and an inner core in the visible wavelengths [*see cover and illustration at bottom left on page 94*]. The inner core, which is often asymmetrical, may consist of solid particles ejected on ballistic trajectories. The ultraviolet globe is possibly the condensate of a symmetrically expanding gas. Of the eight volcanic plumes observed by *Voyager 1* seven were still there when *Voyager 2* arrived four months later. Evidently the plumes erupt continually and last from a few months to a few years.

Another interesting phenomenon on Io is the presence of white or bluish white bright patches along scarps and faults observed in many areas but notably in the south-polar region. Such features are diffuse, are sometimes variable and apparently obscure the surface below, and so it was first suggested that

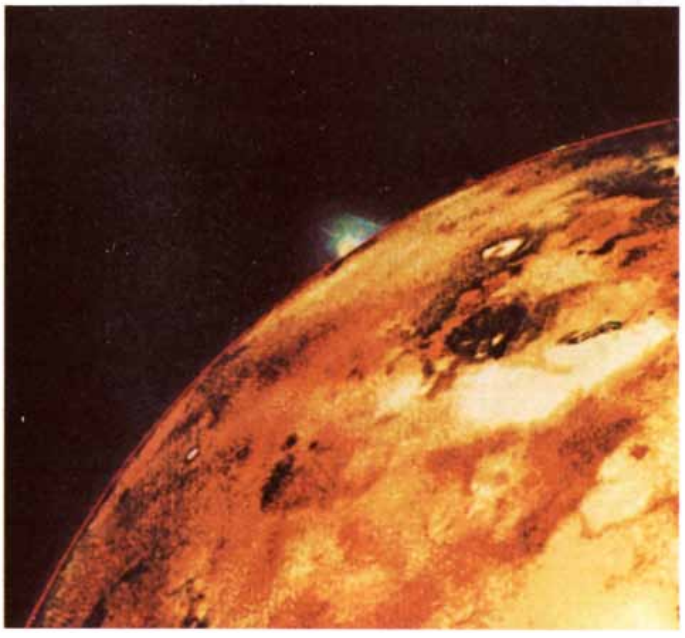


BEST VIEWS OF AMALTHEA were recorded by *Voyager 1*. This asteroidlike body is about 270 kilometers on its long axis and 155 kilometers on its short axis. The picture at the left, made from a range of 695,000 kilometers, has a resolution of 13 kilometers per line pair. Picture at right, made from a distance of 425,000 kilometers, has a resolution of eight kilometers.

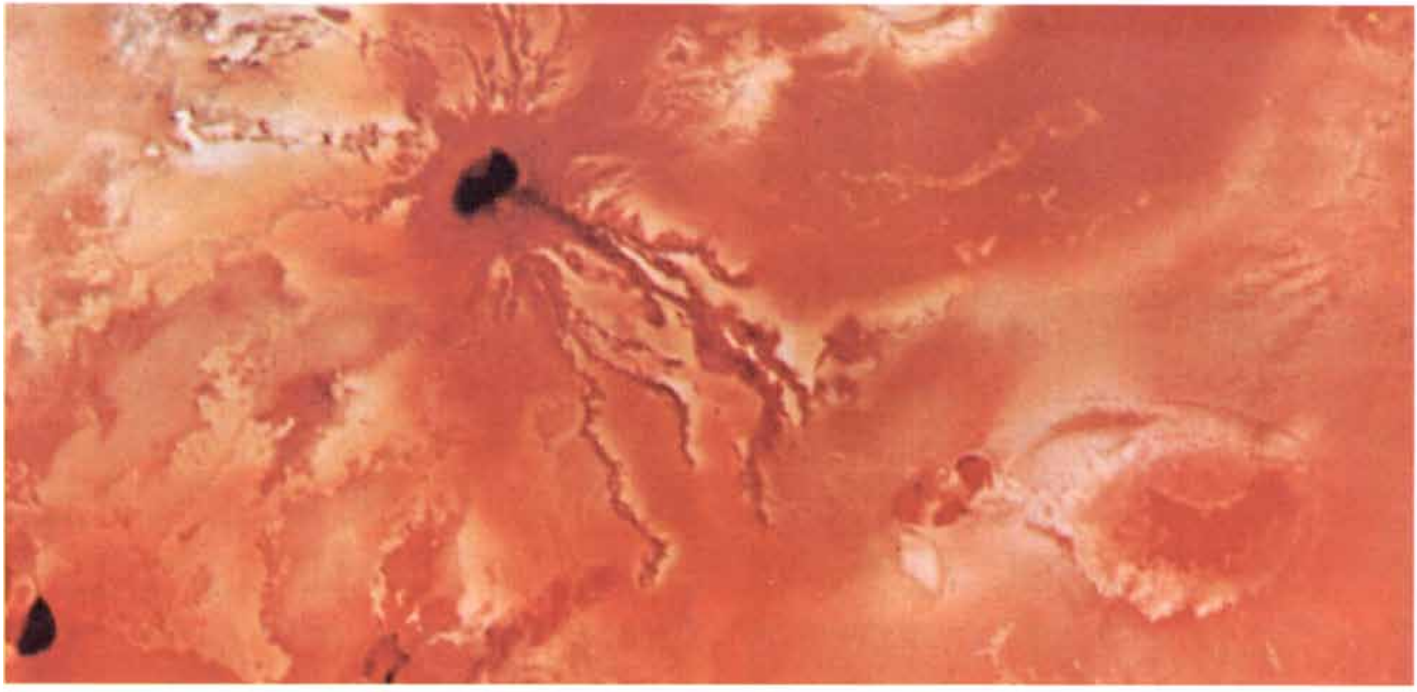
they might be clouds (or material deposited from clouds) produced by a gas leaking out of the interior of the satellite and condensing into some kind of snow. One of the key observations that has led to a preliminary explanation of both the plumes and the isolated white patches along scarps is the discovery by John Pearl of the Goddard Space Flight Center and members of the Voyager infrared experiment that Io has a tenuous atmosphere of sulfur dioxide. From measurements of infrared absorption it has been estimated that during the day this atmosphere has a pressure of about a tenth of a microbar (a ten-millionth of the atmospheric pressure at sea level on the earth).

Another infrared-absorption feature, discovered in the course of telescopic observations made from the earth by Dale P. Cruikshank of the University of Hawaii at Manoa, was interpreted at about the same time by a number of

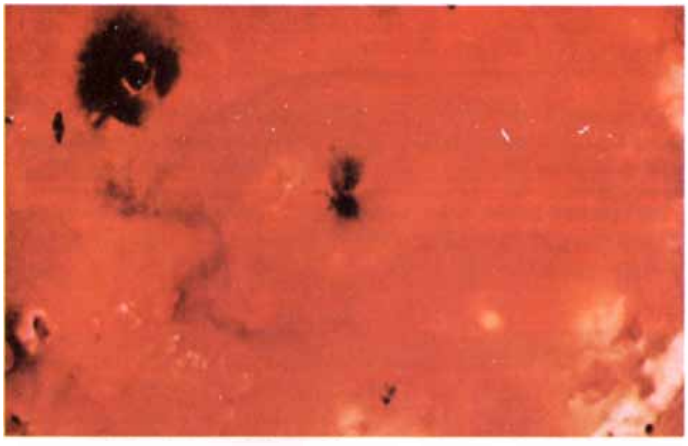
TWO FACES OF IO were recorded about 11 hours apart as *Voyager 1* approached and then overtook this innermost Galilean moon. The picture at the left, made from a distance of about 860,000 kilometers, shows the hemisphere of Io that always faces away from Jupiter as the satellite circles the planet every 1.77 days. The doughnut-shaped feature near the center is the site of an erupting volcano (Plume 3). The resolution is about 16 kilometers per line pair. The higher-resolution mosaic at the right was made on the day of the spacecraft's closest approach and shows Io rotated about 120 degrees to the right, or the east. The large heart-shaped region at the lower right is the site of Plume 1, the first actively erupting volcano to be discovered on the satellite. Four months later, when *Voyager 2* arrived, the eruption had stopped. Closeup views of the same region, both active and quiescent, appear in the illustration at bottom left on the next page.

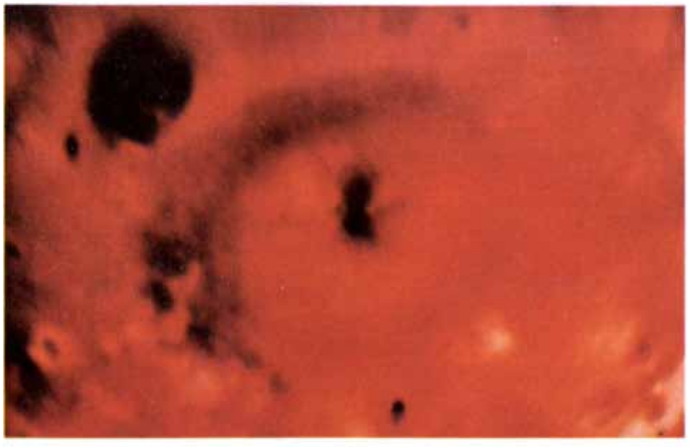


THREE MORE VOLCANIC PLUMES ON IO, of the total of eight sighted, rise to a height of about 100 kilometers in these two views. The picture at the left, made by *Voyager 1,* shows Plume 2. The one at the right, made by *Voyager 2* during a 10-hour "volcano watch" after the spacecraft's closest approach to Jupiter, shows Plume 5 and Plume 6. The image of Plume 2 is a companion to the image on the cover of this issue of SCIENTIFIC AMERICAN. The picture of Plume 2 here is constructed from images made in the visible part of the spectrum and shows only the small central core of the plume. The picture on the cover incorporates an image made through an ultraviolet filter (passing wavelengths centered on 3,500 angstrom units) and reveals that the visible core is surrounded by a larger envelope that scatters uitraviolet radiation. In picture at right Plume 5 is the symmetrical lower plume. Plume 6, the upper one, is smaller and asymmetrical.

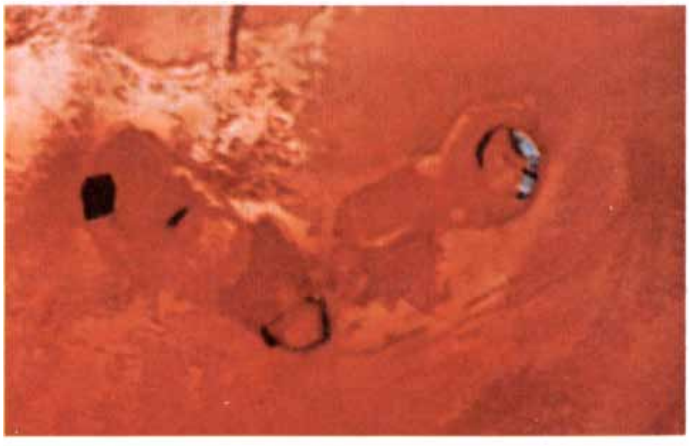


RADIATING FLOW PATTERNS ON IO around a dark caldera, the crater of a collapsed volcano, appear in this picture made by *Voyager 1* shortly before its closest approach. The resolution is about two kilometers. At that point the first of the volcanic plumes had not yet been recognized. The calderas gave Voyager investigators their first clue to the absence of the impact craters they had expected to see long before *Voyager 1*'s closest approach. Even in the pictures of the highest resolution, 600 meters, no impact craters were ever seen. Their absence indicates that the volcanism of the satellite is erasing features as large as one kilometer across in less than a million years.





LARGEST ACTIVE VOLCANO ON IO, the source of Plume 1, is shown in the view at the top made by *Voyager 1*. The heart-shaped pattern extends about 1,200 kilometers in its longest dimension. When the same region was photographed by *Voyager 2* (*bottom*), the heart-shaped pattern had been replaced by a more symmetrical one. Other *Voyager 2* images showed that the eruptions had stopped. The *Voyager 2* image, made at a greater distance, has a resolution of 24 kilometers compared with one of seven kilometers for *Voyager 1* image.



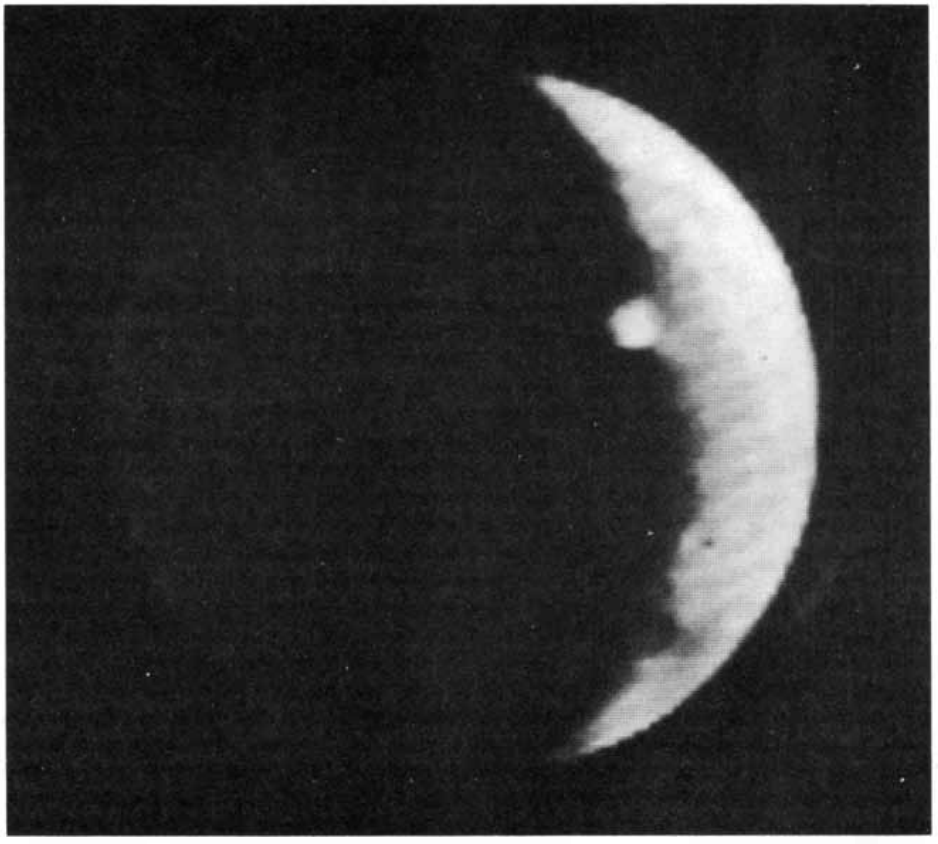
THREE CALDERAS ON IO are shown in images made six hours apart as *Voyager 1* closed in on the satellite from a distance of 374,000 kilometers to 130,000 kilometers. In the first view (*top*) the floors of all three calderas show up as black. In the second view (*bottom*) luminous bluish white patches have developed in the caldera that is farthest to the right. A possible explanation is that liquid sulfur dioxide has leaked out of the interior of the satellite and on reaching the surface has exploded into a large cloud of ice crystals and gas.

observers as being possibly due to sulfur dioxide frost on the surface of the satellite. Recent spectrophotometric analysis of the Voyager images suggests that the large white regions could be about half sulfur dioxide frost and half sulfur. These findings suggest that sulfur dioxide may be abundant in Io's surface layers. If that is the case, liquid sulfur dioxide will be stable in a zone the top of which will be a few hundred meters below the surface, and it will collect in that zone.

The presence of reservoirs of liquid sulfur dioxide, much like aquifers of liquid water on the earth, may provide a simple explanation of what appear to be ice clouds issuing from fractures in Io's crust. The fluid sulfur dioxide would have easiest access to the surface along a fault or at the base of a scarp. As it reaches the surface, perhaps forced out by artesian pressures, the confining pressure drops below a critical point and the liquid explodes into an ice fog that spreads out and falls back to the surface. John F. McCauley of the U.S. Geological Survey and his colleagues have proposed that this process actually erodes material along the scarps.

The sulfur dioxide "aquifer" has also been called on to explain the much larger volcanic plumes. Bradford A. Smith of the University of Arizona and his colleagues on the Voyager imaging team have suggested that the liquid sulfur dioxide causes violent volcanic eruptions in much the same way that water does in the earth's crust. In the model proposed for Io heat, tidally generated in the silicate lithosphere, is carried upward by molten sulfur that has been in contact with the hot rock. When the molten sulfur comes in contact with the liquid sulfur dioxide, the mixture begins to rise toward the volcanic vent and rapidly expands as the pressure drops and the sulfur dioxide vaporizes. When it reaches the surface, its velocity will be between 500 and 1,000 meters per second, which is consistent with the velocities deduced for plumes rising from 100 to 300 kilometers above the surface. This model requires large segregated pools of molten sulfur, possibly large enough to justify being called oceans, below Io's surface. Some of the volcanic flows may therefore consist of almost pure sulfur, as has also been proposed by Carl Sagan of Cornell University.

At the other extreme are models proposed by Michael H. Carr and Harold Masursky of the U.S. Geological Survey in which sulfur and sulfur compounds are simply coloring agents for the lava flows of ordinary silicate volcanism, such as that on the earth. Actually diverse mechanisms may be at work in different places on the satellite. It is generally agreed that the vivid colors seen on the surface are consistent with the wide range of molecular forms of sulfur



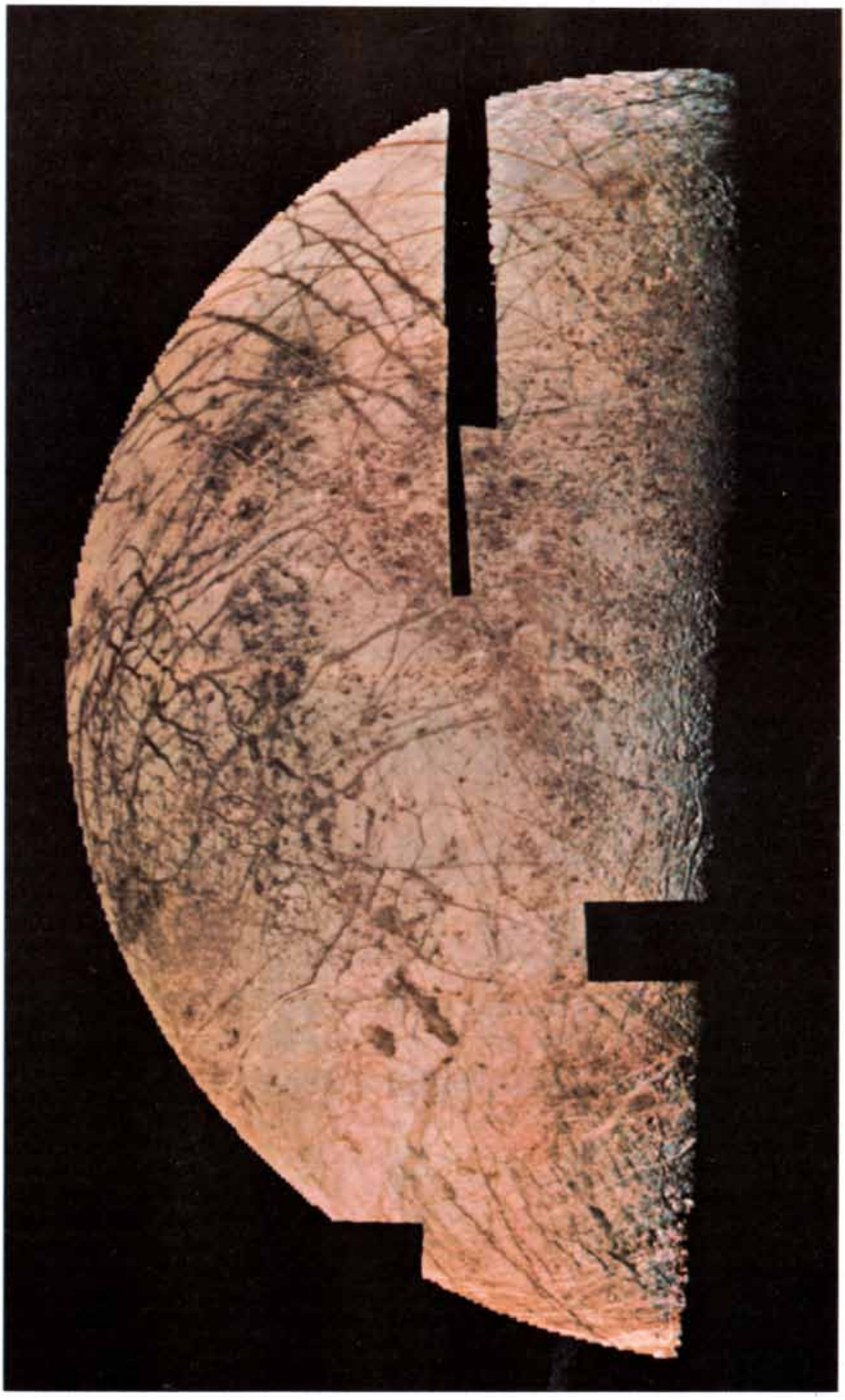
FIRST ACTIVE VOLCANO ON IO was discovered in this image made by *Voyager 1* on March 8, looking back 4.5 million kilometers at the satellite it had flown past three days earlier. Linda A. Morabito of the Jet Propulsion Laboratory detected the faint umbrella-shaped cloud (Plume 1) on Io's limb as she was inspecting a specially processed image. It was then recognized that the picture includes a second volcanic plume that has caught rays of the rising sun, creating a glow just inside the line of the terminator: the edge of the illuminated hemisphere.

CLOUDS OR SURFACE DEPOSITS ON IO can be seen in the vicinity of scarps in this *Voyager 1* picture of a region near the satellite's south pole. The resolution is about 16 kilometers. The white patches may be ice crystals, either aloft or on the surface, produced when liquid sulfur dioxide just under the surface breaks through fractures in the crust and explosively freezes.

known to be stable at Io's surface temperature, which ranges between 60 and 120 degrees Kelvin.

Europa

Europa was farthest away of all the Galilean satellites at the time of *Voyager 1*'s closest approach to it (732,270 kilometers) and therefore was photographed with the lowest resolution, about 33 kilometers per line pair in the best images. This resolution is comparable to a picture of the earth's moon occupying a third of the height of an ordinary television screen. *Voyager 2,* which approached to within 204,000 kilometers, provided an eightfold improvement in resolution. The *Voyager 1* images of Europa show a body that is nearly white, with global markings and patterns that are bland and low in contrast. Its equatorial region exhibits two basic types of terrain: mottled darker regions and brighter regions, both of which are transected by a series of narrow dark stripes tens of kilometers in width extending in some instances for thousands of kilometers.



EUROPA, the next moon outward from Io, was photographed by *Voyager 2* with a maximum resolution about eight times higher than that achieved by *Voyager 1*. This mosaic of *Voyager 2* images has a maximum resolution of about 4.3 kilometers. It shows Europa crisscrossed by stripes and bands that may represent filled fractures in the satellite's icy crust. Water, in solid and liquid form, may constitute about 20 percent of Europa's mass. Satellite's albedo, or surface reflectivity, is nearly 70 percent, or 10 times the albedo of the earth's moon (and twice that of the earth). Europa circles Jupiter every 3.55 days at a distance of 600,000 kilometers.

In the higher-resolution *Voyager 2* images the stripes emerge as a vast tangle of intersecting lines [*see illustration at left*]. The mottled dark terrain is resolved into a series of interlocking depressions and mesas whose typical dimension is a few kilometers. Many of the depressions may be impact craters, but their identification is uncertain.

Only three probable impact craters have been identified. They are between 18 and 25 kilometers in diameter and display widely different morphologies. One crater is fresh and bowl-shaped. Another is shallow and surrounded by a system of dark rays. The third seems to be raised on a pedestal as if the surrounding area had been etched away. In some pictures there is a hint of numerous smaller craters a few kilometers in diameter, particularly along the terminator zone (the zone at the edge of the satellite's illuminated side). The surface of Europa is at least hundreds of millions of years old and is possibly billions of years old.

The absence of substantial relief along the terminator, where it would be most visible, suggests that water, the most likely principal volatile component of the satellite, has risen from the interior to the surface and formed a thick mantle of ice that obscures topographic relief. A mantle 100 kilometers deep would certainly be thick enough to mask whatever relief could possibly exist on the silicate lithosphere.

Europa, along with all the other planetary bodies in the solar system, was presumably created by the accretion of material circling the sun four to five billion years ago. Thereafter, again like the other bodies, it continued to be subjected to an intense bombardment by meteorites, large and small, for perhaps another half-billion years. Its present appearance is probably a sensitive reflection of the relation between its early intense bombardment and its thermal history. If its icy crust had formed, frozen and become rigid in the period of early intense bombardment, the scars of the impacts would be clearly visible today. Evidently its crust remained warm, soft and mobile until late enough in its history to obscure the evidence of the bombardment.

Two mechanisms have been suggested to explain how the surface temperature of the satellite could have been

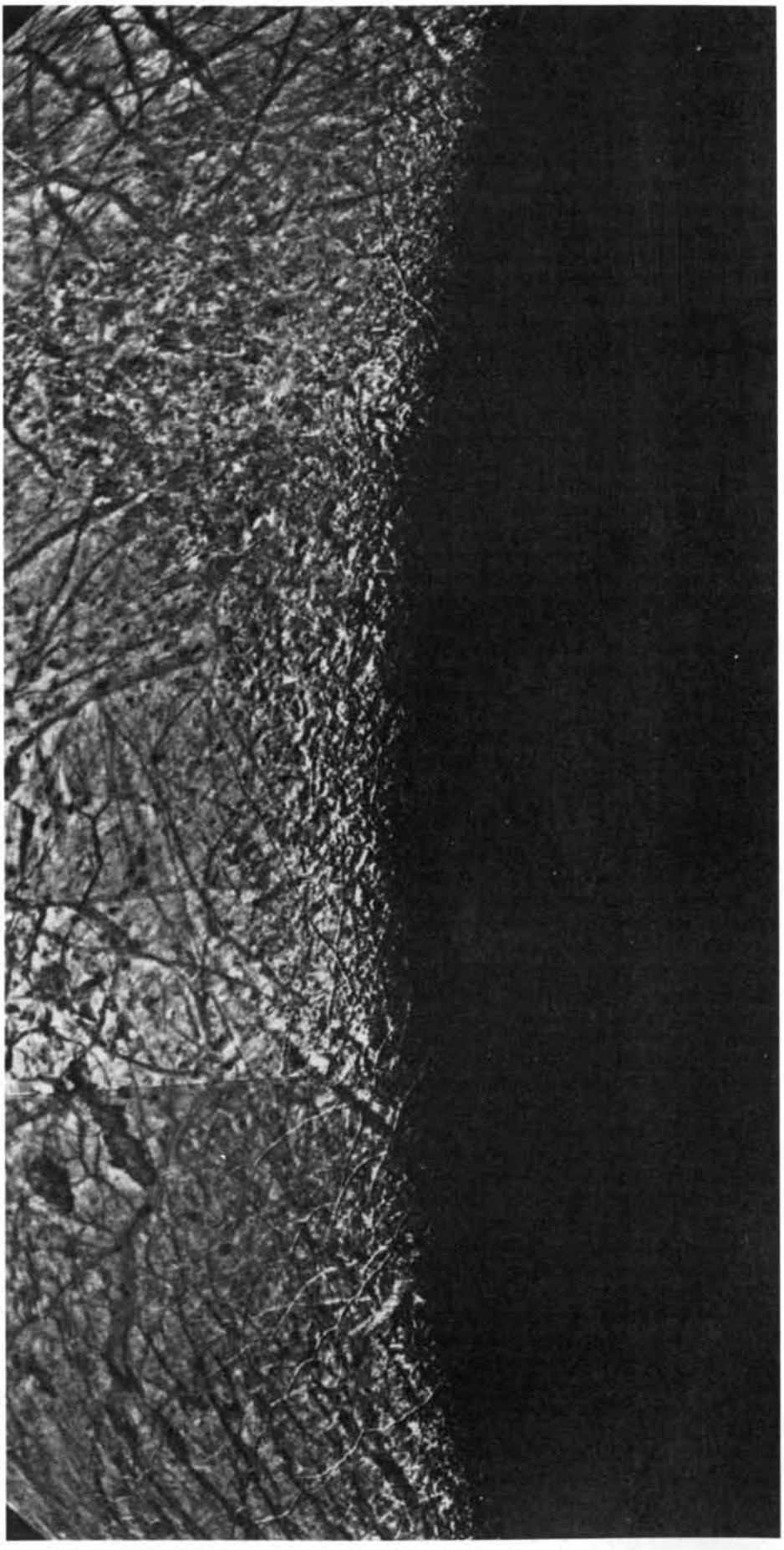
maintained during this early period. Fraser P. Fanale and his co-workers at the Jet Propulsion Laboratory proposed a few years ago that the heat released in the decay of radioactive elements would in itself have been sufficient to keep Europa's icy mantle soft, and perhaps liquid, at depths below some tens of kilometers. Recently Cassen and his colleagues have suggested as an alternative that the same kind of tidal friction heating Io would heat Europa, although to a lesser degree. They also suggest that the dark stripes on the surface of the satellite could be filled cracks resulting from simple expansion when the oceans froze. The hypothesis may not, however, be adequate to explain the increase in area of from 10 to 15 percent implied by the width of the large dark markings. Many other models are possible.

Ganymede and Callisto

Ganymede and Callisto, the two outermost Galilean satellites, will be treated here as a pair because their bulk properties are similar and because the Voyager evidence indicates that their surfaces evolved to some extent in parallel. Although Callisto is the more distant from Jupiter, I shall take it up first because it is easier to understand and because stages in its development seem to be evident in some of Ganymede's oldest terrains.

The Jupiter-facing hemisphere of Callisto was imaged at high resolution (between 2.3 and seven kilometers per line pair) as *Voyager 1* flew past the satellite at a distance of 124,000 kilometers [*see illustration on page 89*]. The surface is almost saturated with craters, yet Callisto looks quite different from the earth's moon. For example, there is a complete absence of visible relief at Callisto's bright limb. The left half of the hemisphere is dominated by a system of concentric rings centered on a bright circular region 10 degrees north of the equator and about 600 kilometers in diameter. The rings, from 50 to 200 kilometers apart, extend out to a radius of some 1,500 kilometers.

The population of craters within the inner part of the ring system is lower by a factor of three than it is elsewhere on Callisto. The population in the outer part of the system has a density similar to that in other regions, but judging by the transection and intersection of the features some two-thirds of the craters were formed before the ring system was. Evidently preexisting craters within about 350 kilometers of the central zone were erased and craters at greater distances were preserved. Younger craters are superposed on both the rings and the central zone. The most likely explanation is that the large impact that created the rings took place early in the history of the satellite, when the crust was not rigid enough to support and retain the



DETAIL ALONG THE TERMINATOR ON EUROPA is seen at high resolution (about 3.8 kilometers per line pair) in this mosaic of *Voyager 2* images. The low inclination of the illuminating sunlight makes it possible to estimate that the long, narrow ridges rise only about 100 meters above the average surface. The ridges are most abundant in the satellite's southern hemisphere. Much of Europa's terrain consists of interlocked depressions and mesas. Although many of the depressions could be impact craters, they are too small for positive identification.

GANYMEDE is the largest of Jupiter's moons. This picture, which was made by *Voyager 2* at a distance of 1,217,000 kilometers, shows the hemisphere of the satellite that always faces away from the planet. The hemisphere is dominated by an immense dark area, the largest remnant of an ancient crust riddled by the impact of meteorites. Early in its history Ganymede may have been as densely cratered as Callisto is today. Ganymede is only two-thirds as dense as Europa and therefore may have a higher fraction of water, perhaps as much as 50 percent.

GANYMEDE'S JUPITER-FACING HEMISPHERE was photographed at high resolution by *Voyager 1*. The area shown is about 1,600 kilometers across and has a maximum resolution of about four kilometers per line pair. Ganymede's crust presumably consists mostly of ice, which is darkest in oldest regions. Around rayed craters ice may be white because it is cleaner.

topographic forms normally associated with large impact basins.

The rings themselves could have been dynamically formed immediately by the impact of a large body or could have been a delayed response to the impact as the central region rebounded and the surrounding surface readjusted. Later the icy crust became rigid as it froze to substantial depths. As we shall see, a similar feature of comparable size is partially preserved on the most ancient regions of Ganymede's crust.

Pictures of Ganymede made by *Voyager 1* during its approach to the satellite show a body that remarkably resembles the earth's moon. One pattern visible at this scale that does not resemble anything seen on the moon, however, is a complex intersecting network of irregular, linear, broken and branching bright bands that crisscross the disk of the satellite. At this scale the patterns bear little resemblance to the rays of impact craters scattered elsewhere over the surface. The preliminary suggestion was that the patterns might be tectonic.

As *Voyager 1* flew closer to Ganymede two basic types of terrain were recognized: heavily cratered polygon-shaped regions up to a few tens of kilometers across surrounded by younger regions of grooved terrain. The grooved terrain was found to occupy the bright linear patterns seen at lower resolution. The grooved terrain consists of closely spaced parallel ridges and troughs, each from five to 15 kilometers wide and up to several hundred kilometers long. In some regions as many as 20 parallel grooves and ridges can be counted.

The density of the craters on the grooved terrain of Ganymede is extremely variable, ranging from a density equivalent to that found in the cratered terrain to a density lower by a factor of 10. The implication is that the formation of the grooved terrain began early in the satellite's history and continued for a long time through the period of intense bombardment. Most of the grooved terrain is a mosaic of discrete systems, with the grooves of one system ending abruptly at the boundary of an adjacent system. In some instances, however, several systems of grooves transect one another. In such instances the older system is not offset by the younger, implying that the process that forms the grooves does not involve lateral motion. It has been suggested that the grooves are formed by a process similar to the one that creates intrusive dikes in the earth's crust: fluid material might be injected into faults and then freeze.

The encounters of the two Voyagers with Ganymede were highly complementary. *Voyager 1* viewed the hemisphere of the satellite facing Jupiter; *Voyager 2* encountered the satellite before the spacecraft's closest approach to

Jupiter and therefore got high-resolution coverage of the opposite hemisphere [*see top illustration on opposite page*]. The most striking feature in the *Voyager 2* view of Ganymede is a large circular region of ancient dark cratered terrain covering about a third of the entire hemisphere. It is the only feature that can easily be identified with telescopes on the earth. Traversing the dark region is a series of giant bright streaks that are slightly curved and parallel. In the higher-resolution images they bear a striking resemblance to the ring structures on Callisto. Although studies have shown that the streaks on Ganymede form concentric circles, there is no evidence at their center of symmetry of a bright impact-disrupted zone like the one on Callisto. Evidently the formation of younger terrains, primarily grooved ones, has erased the record of an ancient impact site.

Although the ring systems on Callisto and Ganymede exhibit strong similarities, where the rings on Callisto take the form of flat-topped ridges those on Ganymede are furrows. The furrows suggest what the ridges on Callisto looked like before they were filled. Such differences may reflect variations in the detailed history of the two bodies, including differences in the strength of the early crust, differences in the chronologies of cooling and freezing and differences in the availability of fluid materials capable of being extruded.

Another important observation is the wide variation in the appearance of impact craters on Ganymede. Some of the craters look fresh and young. At the other extreme are ancient circular forms that are hardly more than "ghosts" or stains on the ancient crust. Evidently the brighter ghost craters were formed when Ganymede's crust was too warm and soft to hold a record of the impact that made them. Close examination indicates that the younger or smaller the crater is, the better preserved its relief is. Eugene M. Shoemaker and his colleagues in the U.S. Geological Survey have concluded that the cratering record in ancient terrains on Ganymede and Callisto is also a record of the thermal gradient and the strength of the upper crust as a function of time as the two bodies began to cool.

The discovery by *Voyager 2* of a large fresh impact basin near the south pole of Ganymede is further evidence of major changes in the strength and rigidity of the satellite's crust. The basin resembles some of the major impact basins of the inner solar system. In high-resolution views one can see that the flat-floored basin is surrounded by blocky massifs and a blanket of ejecta, indicating that by the time the basin had formed the crust had hardened enough to support large-scale mountainous relief. The fact that numerous large craters are superposed on the basin suggests it was formed in the first billion years after the accretion of Ganymede, when the satellite was still being bombarded by bodies of substantial size. Evidently the crust was cooling rapidly and freezing during the first few hundred million years of the satellite's existence.

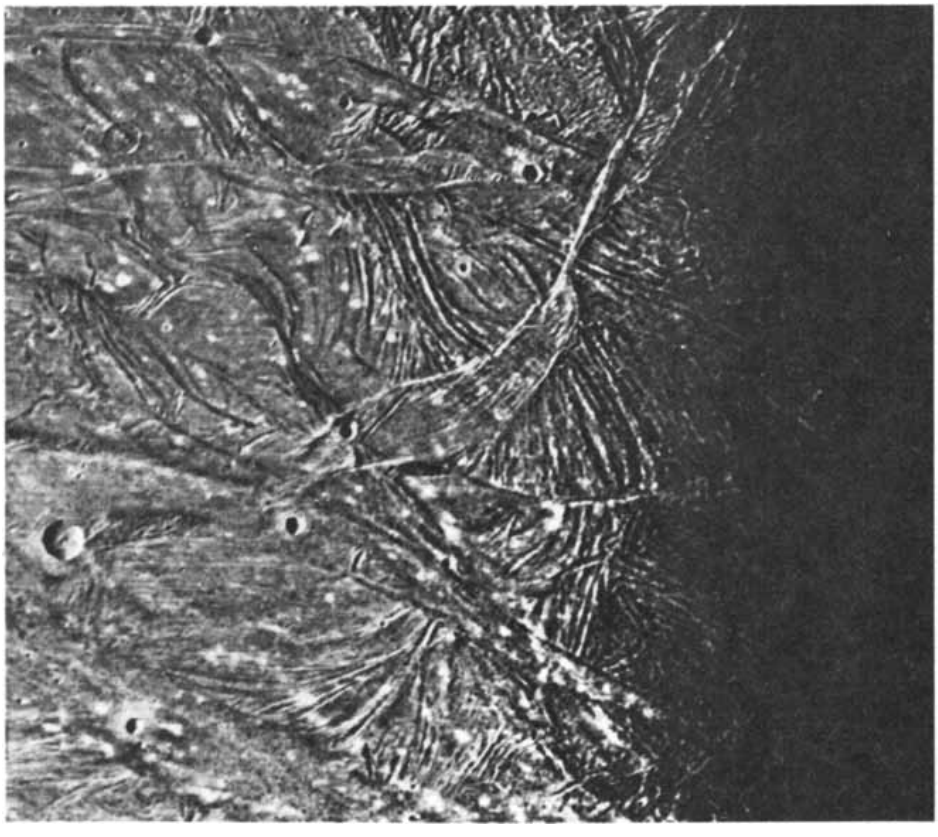
By comparing Ganymede and Callisto it is possible to reconstruct the evolutionary history of Ganymede in broad strokes. One can imagine that the ancient surface of the satellite was soft, icy and dark, resembling the present surface of Callisto. The huge array of slightly curved parallel streaks in the northern hemisphere was created when a large meteorite punctured the soft crust and set up oscillations and strains that fractured the crust in an enormous pattern of concentric rings. Water filled the fractures and froze. Later polar caps were formed, and grooved terrain slowly began to develop and grow. As sections of the crust slid past each other faults at an angle to the lines of slippage gave rise to offsets in the lines. Then the crust became rigid enough to preserve a lasting record of meteorite impacts. The young impact basin near the south pole and a large bright-rayed crater near the equator were the last large-scale features created in the hemisphere facing away from Jupiter that was viewed by *Voyager 2*. In the several billion years since these early events in the evolution of Ganymede the satellite has evidently remained quiescent.

History in the Craters

One of the major objectives of planetary science is to compare the geological histories of the terrestrial planets and satellites in the solar system as they are revealed by the detailed images returned by spacecraft. The principal means of establishing the relative time scales of these histories from images is to determine the record of impact cratering on the simple premise that the older the surface is, the more craters there are on it. Establishing absolute time scales from the cratering record is much more difficult because it requires a knowledge of the cratering rates for each of the bodies.

There is now ample evidence on how the rate of cratering differed from one part of the solar system to another. Models of the rates at which craters were made by the impact of asteroids (and to a lesser extent of comets), together with observations of the crater populations on different bodies, suggest that the average rates over the past few billion years were similar for Mercury, Mars, the earth and the moon within a factor of perhaps two.

Shoemaker has suggested, however,



GROOVED AND TWISTED TERRAIN is a distinguishing feature of Ganymede. This *Voyager 1* view along the terminator from a distance of 135,000 kilometers has a resolution of about 2.4 kilometers. The closely spaced parallel ridges and troughs are usually from five to 15 kilometers across. Younger systems of grooves can be seen transecting older systems. In the area shown here the grooved terrain has totally replaced heavily cratered ancient dark crust.

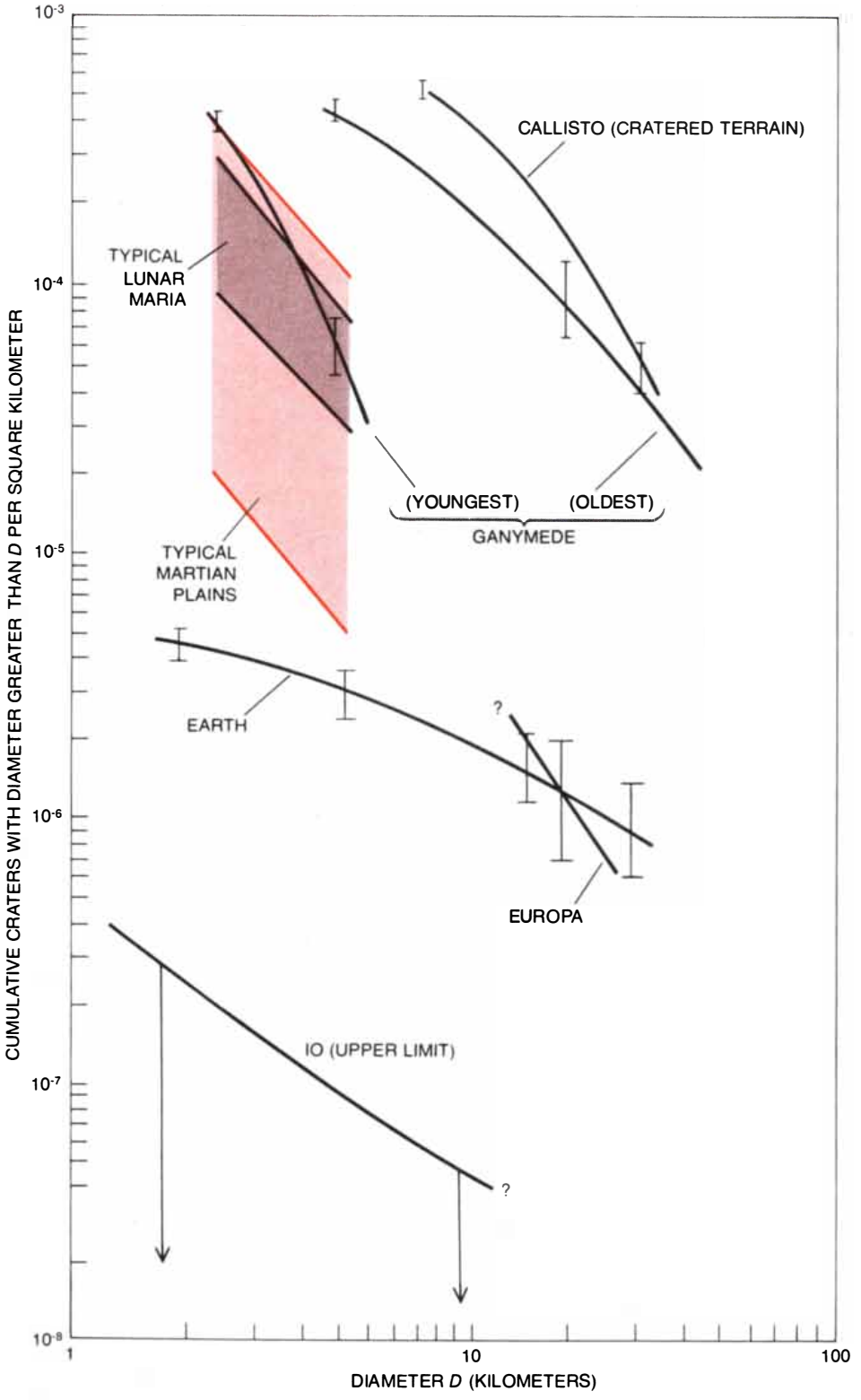
that over the past few billion years the cratering rates for Jupiter's satellites may have been substantially lower than those for the inner solar system, perhaps by a factor of between 10 and 100. His reasoning is as follows. Although asteroids are deflected outward from the inner solar system to the region of Jupiter's orbit, they are probably soon deflected again by Jupiter's powerful gravitational field to other parts of the solar system. Therefore they have only a small probability of hitting Jupiter or its satellites. That being the case, the flux of smaller objects in the Jovian system is probably dominated by periodic comets, which are less abundant than asteroids by a factor of at least 10 and perhaps as much as 100.

Another factor that influences the cratering rates of the satellites is Jupiter's great mass, which dramatically increases the velocity of incoming objects. Thus an impacting body of a given size would tend to produce a larger crater on a satellite near Jupiter than it would on one farther away. This enhancement of velocity is negligible for Callisto, but it is about a factor of two for Ganymede, three for Europa and five for Io.

When crater size is plotted against impact-frequency distribution for the Galilean satellites and for representative bodies of the inner solar system, one obtains curves for Callisto and Ganymede that are rather like those for the earth's moon and Mars [*see illustration at left*]. In contrast Europa (on the basis of only three probable craters) and the earth have fewer craters of a given diameter than the other bodies by a factor of between 10 and 50. Although Europa may have many small craters less than 15 kilometers in diameter, direct evidence for them is lacking. Io, as we have seen, has no craters at all down to the limit of Voyager image resolution. The low cratering value for the earth undoubtedly reflects the difficulty of identifying small craters and the obliterating effects of geological processes.

The following general conclusions can be reached. The heavily cratered terrains on Ganymede and Callisto, comparable to the heavily cratered highlands on the earth's moon, Mars and Mercury, must date back to the period of torrential bombardment some four billion years ago. Evidently the grooved terrain on Ganymede started to form before the early intense bombardment ended.

A lower limit on the recent cratering rates can be derived as follows. First, the surface of Europa certainly cannot be older than about four billion years or it would still bear the scars of large ancient craters formed during the early torrential bombardment. Three identified craters in four billion years sets a lower limit for the flux of crater-making objects in the vicinity of Europa at about a tenth of the flux in the vicinity of the earth's moon. This is consistent with Shoemaker's estimate of the flux in the vicinity of Europa and could imply that the surface of Europa is billions of years old. At the other end of the range of possibilities, the surface of Europa could not be younger than perhaps 100 million years or the flux in the vicinity of Europa would have to be much greater than that in the vicinity of the moon. As for Io, regardless of flux estimates the total absence of craters simply means that it has the youngest and most dynamic surface yet observed in the solar system.



SIZE AND FREQUENCY OF THE IMPACT CRATERS on Europa, Ganymede and Callisto fall within the range measured on the earth, the earth's moon and Mars. Io, however, has no impact craters down to the resolution of the Voyager images of the satellite: about one kilometer. The curve for Europa is highly uncertain since only three craters can be confidently identified in the Voyager pictures. Estimates for impact craters on the earth were made by Richard Grieve and Michael R. Dence of Canadian Department of Energy, Mines and Resources.

UN MOMENT DE MARTELL

80 PROOF IMPORTED BY THE JOS GARNEAU Co., NEW YORK, N.Y.

DINNER IS OVER, BUT THE NIGHT IS STILL YOUNG.

WHAT BETTER TIME TO EXPERIENCE THE WORLD'S FINEST COGNAC—MARTELL.

AFTER ALL, MARTELL COGNAC HAS HELPED MAKE EVENINGS MEMORABLE SINCE 1715.

MARTELL V.S.P., V.S.O.P., AND CORDON BLEU.

MARTELL COGNAC

Geometrical Illusions

In these classic figures of psychology lines appear different from the way they really are. The effects appear to be related to clues to the size of objects in the three-dimensional world

by Barbara Gillam

Geometrical illusions are line figures in which the length, orientation, curvature or direction of lines is wrongly perceived. For example, in certain figures two lines of equal length appear to differ in length. Why does this happen? In normal three-dimensional viewing two lines of unequal length at different distances from the observer can cast on the retina images of equal size, but the lines are not perceived as being equal because the visual system takes into account the fact that they lie at different distances. The lines appear to have the lengths they actually have in the three-dimensional world because the perceptual mechanism known as size constancy seems to compensate for the difference in distance by making the far line appear to be larger and the near line appear to be smaller. It has been suggested that size constancy is responsible for geometrical illusions. In other words, if the visual system processed the lines of a geometric figure as if they were at different distances, then the result would be an illusion.

Attractive as this explanation of illusions may be, it is not correct because in most illusory line figures there is no depth, either real or apparent. It has nonetheless seemed to several investigators, including me, that some process contributing to accurate perceptions of the three-dimensional world might give rise to illusions in two-dimensional figures. My own finding is that geometrical illusions depend not on apparent depth but on clues to the scale and size of objects in the visible world, clues such as linear perspective and foreshortening. I shall be returning to this distinction between apparent depth and perspective clues in somewhat greater detail.

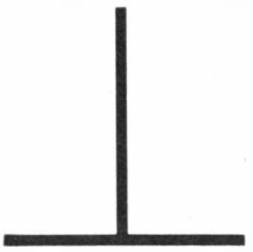
Most of the more than 200 geometric illusions that have been recorded by investigators were discovered in the second half of the 19th century. Some of the best-known illusions are shown in the illustration on the opposite page. In the Müller-Lyer figure the two lines, one with outward-pointing arrowheads at the ends and the other with inward-pointing arrowheads, are actually the same length. In the Ponzo figure the horizontal lines are also the same length. In both the Zöllner and Lipps figures the oblique lines are parallel. In the Titchener figure the two inner circles are the same size. In the Delboeuf figure the outer circle on the left is the same size as the inner circle on the right. In the Poggendorff figure the oblique lines are collinear. In the upside-down *T* the horizontal and vertical lines are the same length. And in the Judd figure the dot is at the middle of the horizontal line.

Each of these illusions has traditionally been thought to consist of two parts: an "inducing" component that does the distorting and a "test" component that is distorted. For example, in the Müller-Lyer figure the arrowheads are the inducing component and the horizontal lines are the test component, and in the Poggendorff figure the parallel lines are the inducing component and the oblique lines are the test component. This division can, however, oversimplify what is going on. The distortion of the test component is sometimes only the most obvious manifestation of a host of misperceptions that involve the inducing component as well.

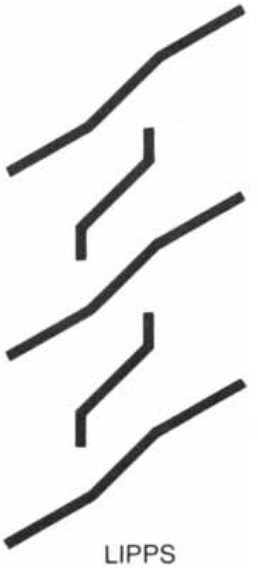
In the 100 years that geometrical illusions have been studied many different explanations for them have been advanced. The most compelling of these explanations agree on three fundamental points. First, the illusions are not conceptual but perceptual; knowing that a particular effect is illusory does not diminish the strength of the illusion, although most of the illusions are drastically reduced when the figure is viewed repeatedly over a short period of time. Second, the illusions do not originate in the retina; they emerge at almost full strength when the inducing component is presented to one eye and the test component is presented to the other, and so they must originate at a point in the visual system beyond the lateral geniculate nucleus of the brain, where the inputs from the two eyes first come together. Third, illusions do not result from the movements of the eye; experiments show that the illusions, usually of full magnitude, emerge when a figure is exposed too briefly for the eye to scan it or when the retinal image of the figure is artificially stabilized by a special apparatus that causes the image to remain still on the retina even as the eye moves back and forth.

Explanations of illusions fall into four main categories: classification theories, activity theories, physiological theories and functional theories. The theories are not necessarily mutually exclusive; in some cases they might emphasize different aspects of the same process. The first of these categories is the least ambitious. The classification theories point to common properties among a variety of illusory figures. What the figures have in common suggests the presence of an underlying perceptual process. Neither the mechanism nor the function of such a process is addressed by the classification theories. The illusions are attributed to contrast when they are characterized by a perceptual exaggeration of the degree to which the test component differs from the inducing component in a pre-

NINE GEOMETRICAL ILLUSIONS are presented on the opposite page. In the upside-down-*T* figure the vertical line and the horizontal line are the same length. In the Lipps figure the oblique lines in the middle are parallel. In both the Ponzo figure and the Müller-Lyer figure the horizontal lines are equal in length. In the Judd figure the dot is at the midpoint of the horizontal line. In the Poggendorff figure the oblique lines are collinear. In the Zöllner figure the oblique lines are parallel. In the Titchener figure the two inner circles are the same size. In the Delboeuf figure the outer circle on the left is the same size as the inner circle on the right.

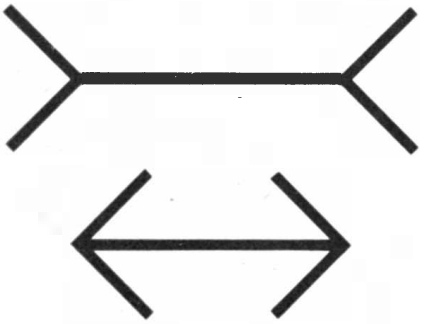


UPSIDE-DOWN *T*



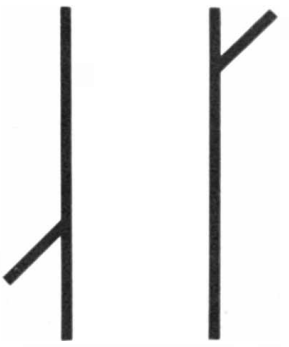
LIPPS

PONZO

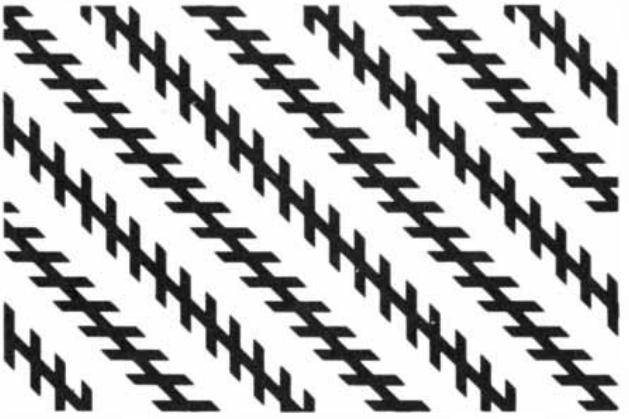


MÜLLER-LYER

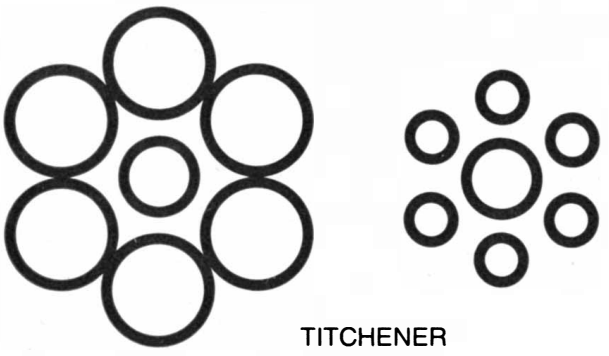
JUDD



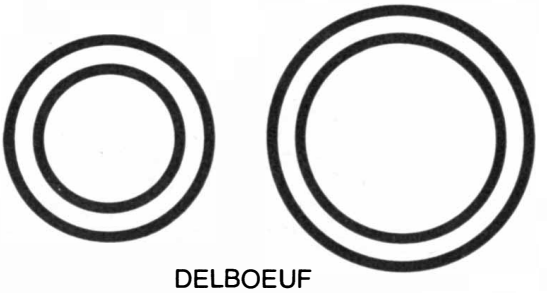
POGGENDORFF



ZÖLLNER



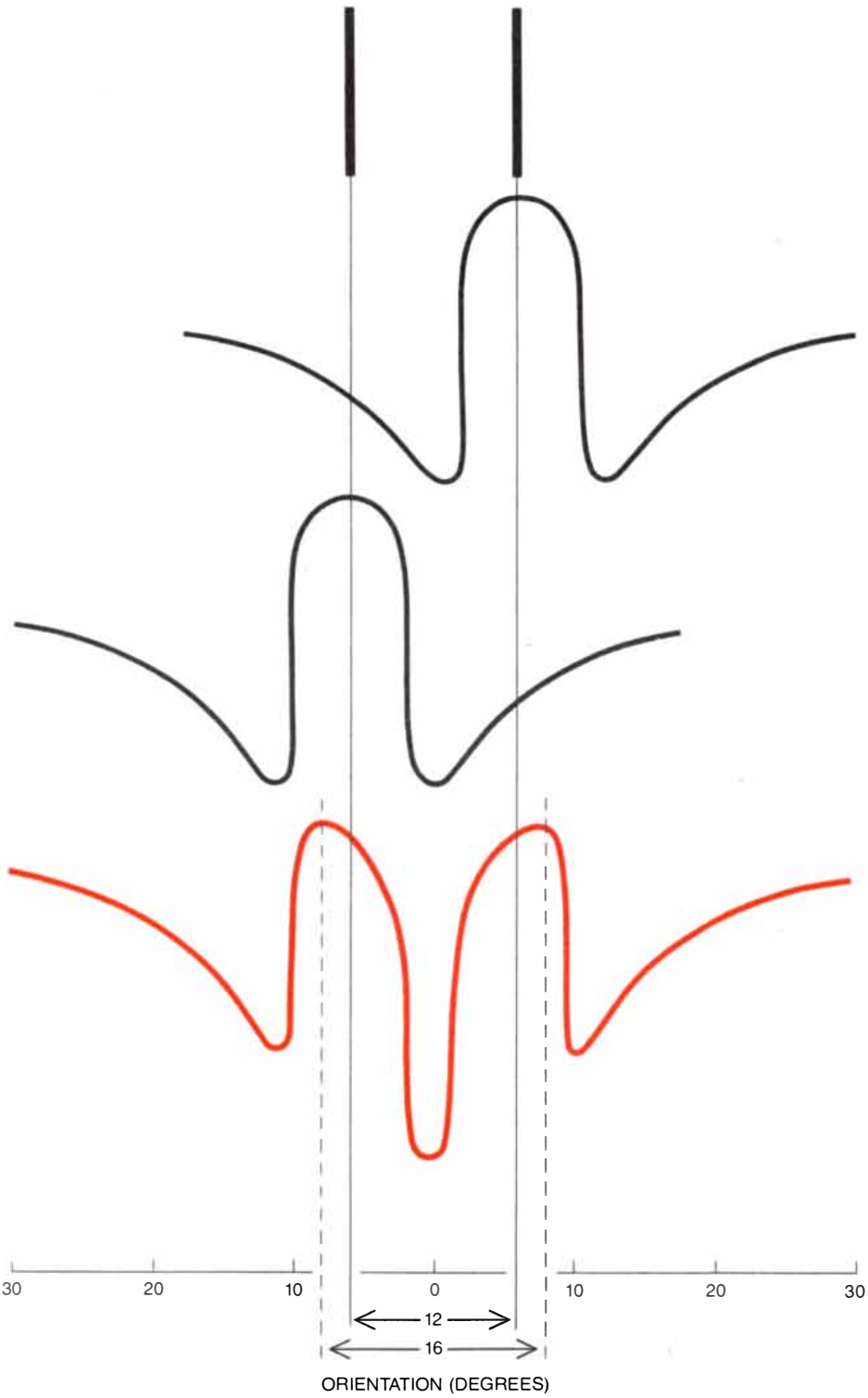
TITCHENER



DELBOEUF

vailing quality such as size or orientation. The Titchener illusion is an example of size contrast, whereas the Zöllner illusion is one of orientation contrast, or more specifically of angle contrast.

The illusions are attributed to assimilation or confusion when they are characterized by a perceptual underestimation of the degree to which the test component differs from the inducing component. In other words, the test element is distorted in the same direction as a prevailing quality of the inducing element. The Müller-Lyer illusion is an example of size or position assimilation because the horizontal lines are lengthened in the direction of the arrowheads. The Lipps illusion is an example of orientation assimilation because the orientation of each parallel line is perceptually shifted toward the orientation of its neighbors. A single illusion can incorporate both contrast and assimilation. In the Ponzo illusion the apparent length of the upper line is distorted toward the surround (assimilation), whereas the apparent length of the lower line is distorted away from the surround (contrast).



VISUAL RESPONSE TO LINE STIMULI is shown in a highly schematic way for two lines (*a*) that differ in orientation by 12 degrees. Below each line is the distribution of orientation-detecting cells of the visual cortex that the line activates. The colored distribution shows what might happen if the two lines were presented simultaneously to the visual system. The sum of the individual distributions gives rise to the colored distribution in which the peaks of activity are slightly displaced from each other. This illustration is based on the findings of Colin Blakemore, Roger H. S. Carpenter and their colleagues at the University of Cambridge.

This classification scheme has done little more than help to bring order to the large number of illusions. It does suggest, however, that the processes could have something in common with those of brightness perception, which exhibits both contrast and assimilation with respect to the brightness of neighboring areas in a scene. Apart from the classification scheme's lack of explanatory power, it is limited by the fact that it does not apply to some important illusions, for example the Poggendorff figure and the upside-down *T*. Moreover, the lines that are assimilated in the Lipps figure are well within the orientation range of the lines that show contrast in the Zöllner figure.

The activity and physiological theories go beyond mere classification by asserting that illusions are incidental side effects or errors of normal visual perception. Activity theories attribute illusions to the responses an individual has to certain stimuli or prepares to have to them. The efferent-readiness theory, proposed by Leon Festinger of the New School for Social Research, maintains that illusions develop because of the way the eye "gets ready" for saccadic movements. (As I have mentioned, experimental evidence rules out the eye movements themselves playing a role in the illusions, but that evidence does not rule out the possibility that the visual system's preparations for eye movements do play such a role.) According to this theory, people tend to look not at the entire figure but at the part of the figure that optimizes the number of details seen with high acuity. Festinger and his co-workers have found that when people try to fixate on the ends of a Müller-Lyer figure, they actually fixate within the arrowheads. According to Festinger, that has the effect of lengthening the line with the inward-pointing arrowheads and shortening the line with the outward-pointing ones.

It is known that the decrease in the strength of an illusion that comes with repeated exposure is accompanied by an increase in the accuracy of eye movements, but this correlation does not indicate whether or not the eye movements cause the illusions. Many other motor responses, such as directing a pointer at the illusion, might provide information

IBM'S GREATEST HITS.

These aren't the kind of "platters" they spin at discos. But they are recordings, and they've sold in the millions.

They're magnetic disks, and what they record is information for storage in computers.

They were introduced by IBM in the mid-50's and revolutionized data processing because they gave the user quicker access to his information.

Since then, disk storage has been universally adopted by the computer industry.

Without it, making airline and rental car reservations, credit card transactions, and 24-hour banking, to name just a few uses, would be all but impossible.

But the most important thing about disk storage is that at IBM we're still inventing it. Finding new ways to pack more information into less space, and retrieve it faster, all for less cost.

Today a single 8″disk (the small one at left above) can store as much information as 120 of the 24″ disks (the big one at the top) used in the first disk storage system we built. The 8″disk can, in fact, pack all the information on 22 newspaper pages into a space the size of a postage stamp!

Information that used to be stored at a cost of $150 can now be stored for about $1.

Technology has gone through the roof.

The cost of using a computer has gone through the basement.

And the person who's getting more for his money is you. IBM®

Chevrolet Monte Carlo for 1980. Even standing still it knows how to move you.

To help keep your new Monte Carlo looking good, look into the double-panel construction of its doors, hood and deck lid. Fenders inside fenders provide extra protection against corrosion.

If that impressive new front-end styling flowing into the unmistakable mark of Monte Carlo's sculptured fenderline moves you, read on.

For as commanding as the 1980 Monte Carlo is to the eye, its true beauty comes when you turn the key.

You'll be moved by what you see. And by what you don't.

Prepare to feel special behind the wheel.

Dials and gages with international markings cluster in front of you. A rich new look of vinyl wood-grain sweeps across the instrument panel, integrating with elegant new interiors for 1980.

As you drive, a highly tuned sport suspension imparts a solid feel of the road.

Up front, a standard 3.8 Liter V6 engine moves you with a smooth rush of power, a sensation you can amplify in 1980 with a new available turbocharged V6.

All in a Monte Carlo with these impressive 1980 EPA estimates.

CHEVROLET MONTE CARLO	EPA EST. MPG	HIGHWAY EST.
Standard 3.8 Liter V6	19	26
Available 3.8 Liter V6 Turbo	18	25

REMEMBER: Compare the "estimated MPG" to the "estimated MPG" of other cars. You may get different mileage depending on how fast you drive, weather conditions, and trip length. Actual highway mileage will probably be less than the estimated highway fuel economy. Monte Carlo is equipped with GM-built engines produced by various divisions. See your dealer for details.

Its Chevrolet price will move you, too.

If by now you think there's never been a Monte Carlo quite like our 1980, you're quite right.

A refined new Monte Carlo. With automatic transmission, power front disc brakes and power steering.

All standard. All at a Chevrolet price.

See your Chevy dealer about buying or leasing Monte Carlo today.

And be moved.

Chevrolet

TURBO

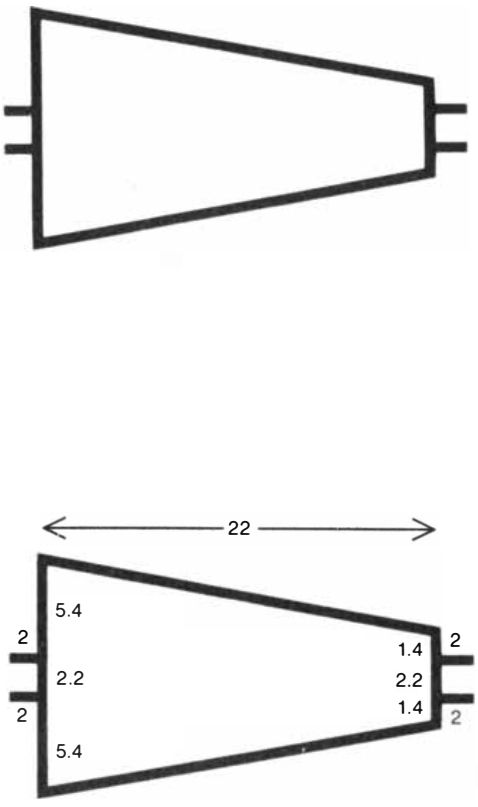
New available 3.8 Liter turbocharged V6 with up to 47% extra power on demand for passing or merging.

Remove the available glass roof panels and invite the sun into Monte Carlo's rich new interior.

GM

about the perceptual error and thereby help to reduce the illusion in the same way that eye movements do. Inasmuch as the illusion would not be attributed to the intention to direct a pointer, it would be premature in the absence of other evidence to attribute the illusion to preparations for saccadic eye movements. Moreover, it is difficult to understand why inaccurate eye movements in the presence of angle and line intersections would not have been eliminated in the normal environment where such intersections are frequently encountered.

The physiological theories attribute illusions to the "hardware" of the visual system. Many of these theories are based on lateral inhibition, a neural process that serves to restrict the stimuli that can fire a cell in the visual cortex of the brain. According to these theories, the perceived orientation of a line is determined by the peak activity of orientation-detecting cells in the visual cortex, each of which is activated by line stimuli of a particular range of orientations. The introduction of a second line of a different but similar orientation gives rise to a somewhat different distribution of activity among the orientation-detecting cells, activity that either facilitates or inhibits the activity generated by the first line, thereby shifting the apparent orientation. On this view facilitation, which would result from very similar lines, would cause assimilation in the illusory figure, and inhibition, which would result from less similar lines, would cause contrast. The same view is supported by the experiments of Colin Blakemore and his co-workers at the University of Cambridge. To explain size illusions a similar physiological theory has been proposed that is based on size-detecting cells in the visual cortex.



SUBJECT	UPPER LEFT LINE	LOWER LEFT LINE
1	+3.2	−12.8
2	+6.8	−9.6
3	−.1	−4.4
4	+1.6	−3.2
5	+2.6	−3.2
6	+1.1	−6.9
7	+1.3	−2.2
8	+5.9	−6.4
9	+2.3	+.5
10	+7.0	−5.3
11	+14.1	−11.8
12	−3.2	−1.6
MEAN	+3.6	−5.6

VARIATION OF THE PONZO ILLUSION at the top left consists of collinear line segments that seem to be displaced. The line at the upper right appears to be too high and the line at the lower right appears to be too low for them to be aligned with their counterparts at the left. The author asked 12 subjects to indicate how much each line at the left should be adjusted so that it would be collinear with the corresponding line at the right. For the figure at the bottom left with dimensions in centimeters the subjects' adjustments are tabulated in millimeters at the right. In the normal Ponzo illusion in the illustration on page 103 the apparent length of the upper line is distorted toward the surround (the distortion called assimilation), whereas the apparent length of the lower line is distorted away from the surround (the distortion called contrast). Yet there is more to the illusion. The figure at the top left shows that the Ponzo configuration affects not only length of lines placed in it but also alignment of lines attached to its ends.

Physiological theories based on lateral inhibition are best able to handle angle-contrast illusions with multiple inducing lines, for example the Zöllner illusion, which exhibits assimilation at extremely small angles and contrast at larger angles. George Wallace and his co-workers at the University of Reading have found that the Zöllner illusion is enhanced by changes that should promote lateral inhibition, such as increasing either the brightness contrast or the number of inducing lines.

Most angle-contrast illusions persist when the test line is replaced by a set of dots or by a moving dot. Physiological theories based on lateral inhibition cannot explain this phenomenon because there is no actual test line to be inhibited. That may not be a problem, however, because little work has been done on whether the orientation-detecting cells might in fact respond selectively not only to a line but also to a row of dots or a moving dot.

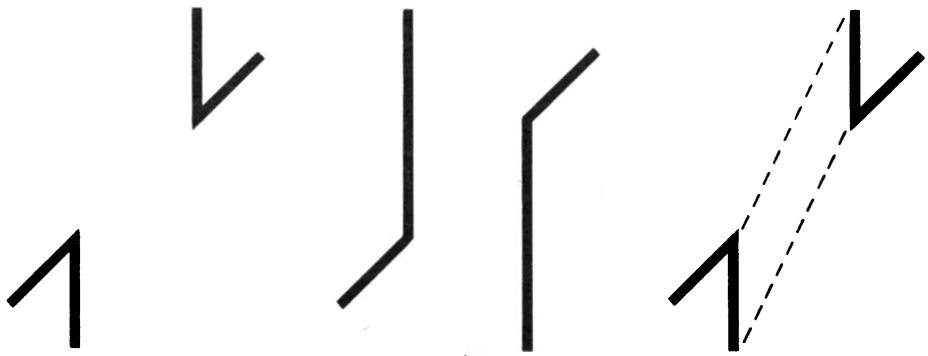
Physiological theories based on lateral inhibition fall short for the simplest illusions involving angles. For example, in the Poggendorff figure the acute angles, which should show the greatest effect of lateral inhibition, contribute nothing to the illusion [*see top illustration on next page*]. Another exception is the Müller-Lyer illusion, which no physiological theory has been able to explain. A comparison of the horizontal lines of the Müller-Lyer figure with a line lacking arrowheads reveals that the inward-pointing-arrowhead configuration (obtuse angles) contributes two or three times as much to the total illusion as the outward-pointing-arrowhead configuration (acute angles). No plausible arrangement of orientation-detecting cells could give rise to such lopsided contributions.

Another physiological theory, filter theory, is based on neural channels in the visual cortex that seem to be tuned not to line stimuli but to particular spatial frequencies of sinusoidal modulations of light, the spatial frequency being the number of modulations per unit of visual angle. Mathematicians have proved that any waveform, no matter how complicated, can be expressed as the sum of simple sinusoidal waveforms. In 1968 Fergus W. Campbell and John G. Robson of the University of Cambridge proposed that the visual system processes a waveform of light by breaking it down into its sinusoidal components. Since then much psychophysical and physiological evidence has accumulated in support of this proposal, and it is now a cornerstone of visual theory.

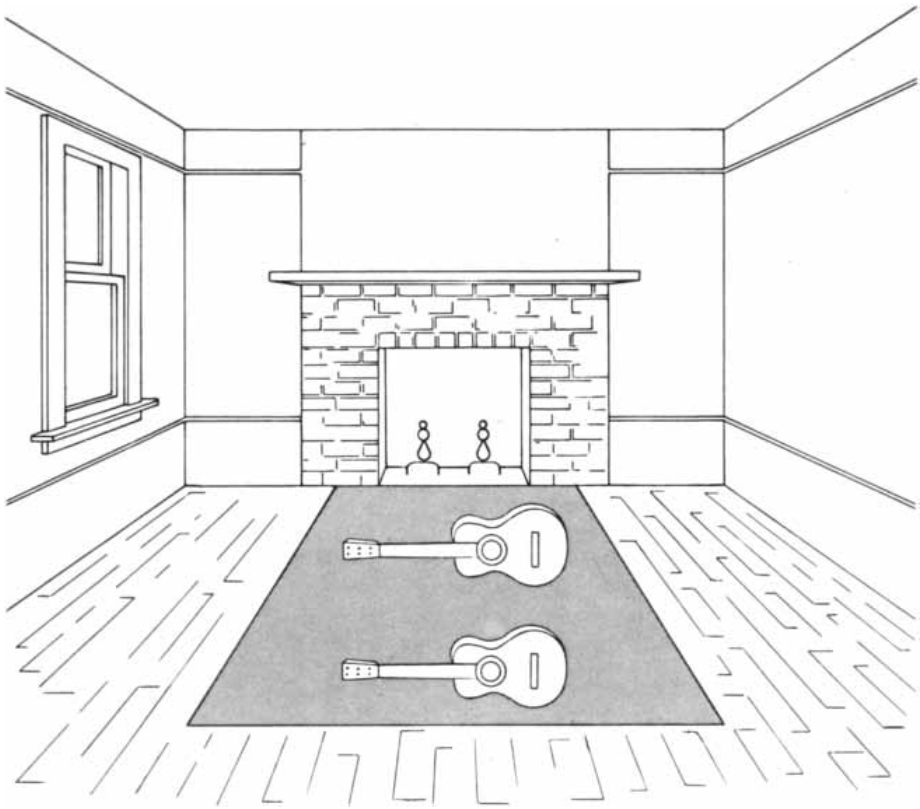
Arthur Ginsberg of the U.S. Air Force attributes illusions to the responses of the neural channels that are tuned to low spatial frequencies. He has broken down the images of certain illusory figures into sinusoidal waveforms of different frequencies and amplitudes, eliminated the high-frequency waveforms and added back together the low-frequency ones. With the high-frequency waveforms filtered out the reconstructed image turns out to have properties that correspond to perceived distortions of the original figure. For example, in the reconstructed image of the outward-

pointing-arrowhead configuration of the Müller-Lyer illusion the arrowheads and the shaft have merged to form a short blurred shaft, and in the reconstructed image of the inward-pointing-arrowhead configuration the arrowheads and the shaft have merged to form a long blurred shaft. The respective shortening and lengthening of the shaft is of course the substance of the illusion.

Ingenious as Ginsberg's work may be, there is little reason to believe that in arriving at size judgments the visual system filters out the information provided by high-frequency waveforms. It would be odd (although not impossible) if the visual system judged the apparent size of a distinct line by the responses of neural channels in which the line was indistinctly merged with its surround, particularly because such judgments would lead to gross perceptual errors.



POGGENDORFF FIGURE is decomposed so that the illusory effects of the acute-angle components (*a*) can be distinguished from those of the obtuse-angle components (*b*). In *a* the illusion is zero or even slightly negative, the line at the right appearing to some people to be slightly lower than the one at the left. In *b* the illusion is strongly present. An illusion in *a* may be due to the visual system's processing *a* as if it were part of *c*, in which parallel lines form a receding plane. In *c* the line at the right is in fact lower on the receding plane than the one at left.



PICTURE OF A LIVING ROOM includes many well-known illusory configurations. One guitar looks longer than the other (Ponzo illusion), the rear edge of the carpet looks shorter than its front-to-back dimension (upside-down-*T* illusion), the molding looks too high to be collinear with the baseboard (Poggendorff illusion) and the front edge of the carpet looks shorter than the bottom edge of the back wall (Müller-Lyer illusion with half of each arrowhead).

A major drawback of both activity theories and physiological theories is that they cannot account for the decreases in the strength of illusions. The fact that by repeated exposure illusions can diminish almost to zero for as long as a period of days does not support theories that posit immutable physiological mechanisms. Only functional theories, which treat illusions not as errors but as essential processes of the visual system, begin to deal adequately with decreases in the strength of illusions.

Over the past century several investigators, including the 19th-century psychologist Armand Thiéry, Richard L. Gregory of the University of Bristol and Reinhardt Tausch of the University of Marburg, have shown that most illusory figures can be found in two-dimensional representations of three-dimensional scenes. What are called distortions or illusions in sketchy drawings promote accurate perception in the normal three-dimensional world and increase the realism of a picture. Consider the bottom illustration at the left. One guitar looks longer than the other (the Ponzo illusion), the rear edge of the carpet looks shorter than its front-to-back dimension (the upside-down *T*), the molding looks too high to be collinear with the baseboard (the Poggendorff illusion) and the front edge of the carpet looks shorter than the bottom edge of the back wall (the Müller-Lyer illusion with half of each arrowhead).

Although all these percepts are distortions compared with the picture plane or the retinal image, they would not normally be called illusions because they are not surprising. They reflect characteristics of the three-dimensional world represented in the picture plane. In normal three-dimensional viewing it is quite irrelevant whether or not a baseboard is retinally collinear with a molding and whether or not the edge of a door has the same length in the retinal image as the corner of a room. Such facts about the picture plane, which depend mostly on the viewing position of the observer, are not registered.

The main function of perception is to decode the transient retinal image in order to achieve constancy: the perception of the external world in terms of its stable and intrinsic characteristics. In the illustration of the room it is clear that the "distortions" are just instances of perception doing its job. If the contextual details that are necessary for a three-dimensional impression are eliminated, however, the same perceptual responses to the same configurations seem unjustified and are called illusions. Does this mean that illusions are only constancy responses (the decoding of perspective) operating in a context that is too minimal, since it lacks apparent depth, to give the distortions meaning? This view

is appealingly parsimonious, but it goes against a strong tradition in psychology that constancy responses are triggered by apparent depth.

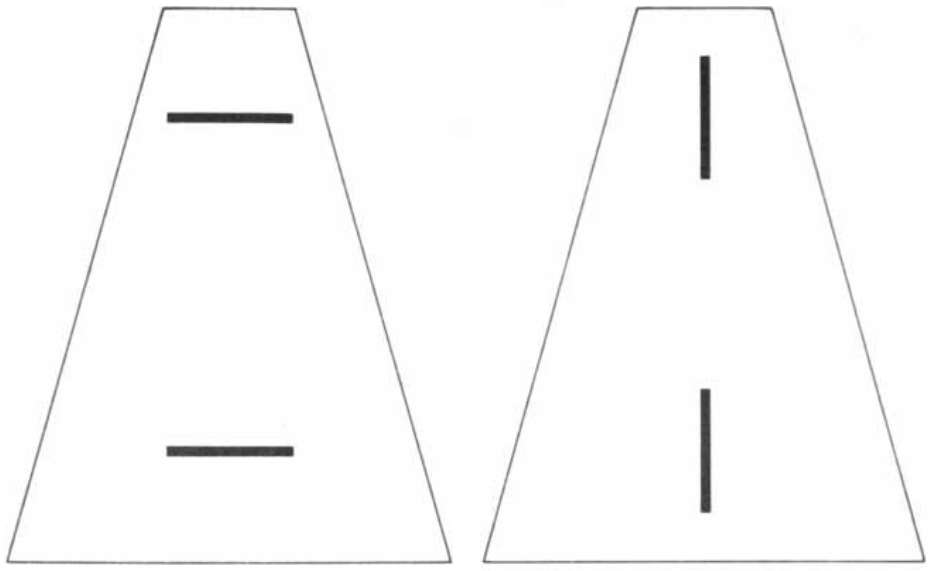
There are a number of reasons to doubt that a depth response always underlies constancy scaling. J. J. Gibson of Cornell University has pointed out that gradients in the apparent compression of objects and textures by perspective provide information about the size and shape of objects. The size of an object can often be judged by the ratio of its width to the width of the background. Consider railroad tracks receding into the distance. Since the viewer knows that the distant railroad ties are as large as the close ones, an object spanning the rails in the distance is perceived as being the same size as an object spanning the rails in the foreground, in spite of the fact that the sizes of the objects' retinal images are quite different.

Conclusive evidence for a size-constancy process that is not secondary to perceived depth comes from my own work with backgrounds in which equivalent depths are represented by different types of perspective [*see illustrations at right*]. In effect I was able to separate the response to depth from the response to perspective. I investigated two kinds of background drawing. In the first kind, based on the receding oblique lines of the Ponzo figure, the horizontal dimensions are increasingly compressed toward the "far" end, or top, of the drawing (the distortion known as linear perspective). In the second drawing, based on horizontal parallel lines that get closer to one another toward the top of the drawing, the vertical dimensions are increasingly compressed (the distortion known as foreshortening).

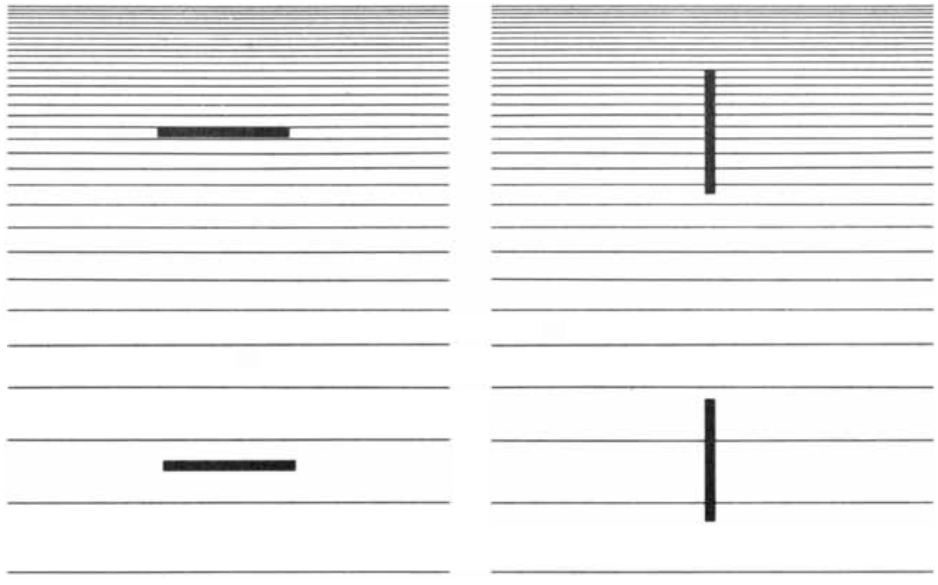
Next I put test lines at various orientations on the two kinds of background. I found that the perceptual lengthening of lines put at the top with respect to lines put at the "near" end, or bottom, occurred only when the test lines were oriented along the compressed dimensions. There was no illusion for test lines at orientations for which the background drawing did not provide a scale of size. In other words, a linear-perspective scale does not affect the apparent length of lines along the dimension subject to foreshortening. This result constitutes compelling evidence for a perceptual mechanism promoting size constancy that is based on the scale itself and not on a depth response to the scale, which would lengthen any apparently distant lines regardless of their orientation.

This is not to deny that apparent distance can influence size perception. Experiments have demonstrated that changes in perceived size and shape accompany depth reversal in ambiguous figures. The fact that perceived size and shape can be primary responses and yet can still be influenced by apparent depth is not surprising when one considers that depth perception itself can be influenced by apparent size and shape. An example is the well-known moon illusion. When the moon is near the horizon, it looks larger than it does high in the sky. As a result when it is near the horizon, it looks closer. Under normal circumstances primary and secondary ways of judging size and distance reinforce each other. My work shows, however, that the amount of information in a line drawing needed to trigger a primary process of size scaling is less than the amount needed to trigger a process of depth scaling. Gregory has suggested that this result may be due to a perceptual conflict between the drawing's representing objects in depth and its actually being flat.

How well can a functional theory

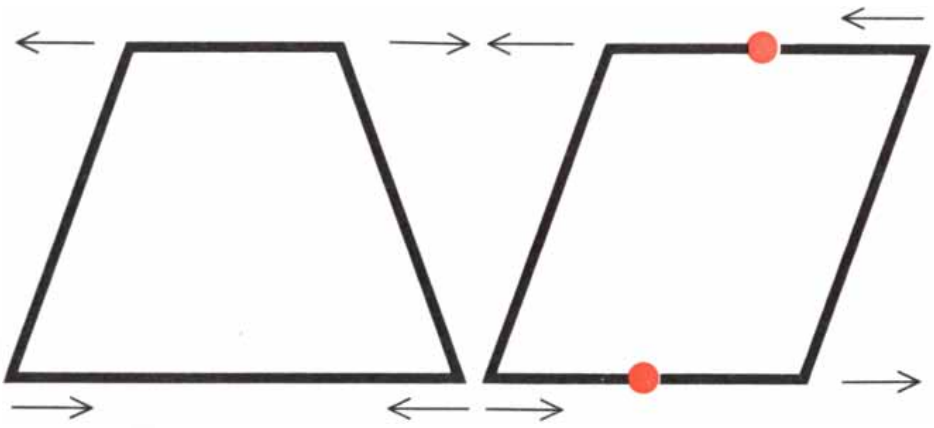


LINEAR PERSPECTIVE, the compression of horizontal dimensions toward the top of a drawing, is achieved by the receding oblique lines of the Ponzo figure. The horizontal test lines (*left*) are the same length, but the upper line appears to be longer because of linear perspective. There is no illusion for vertical test lines (*right*). This means that here the perception of size depends not on depth, which would increase the length of all the test lines regardless of their orientation, but on perspective, which affects only lines for which the background shows a scale of size.

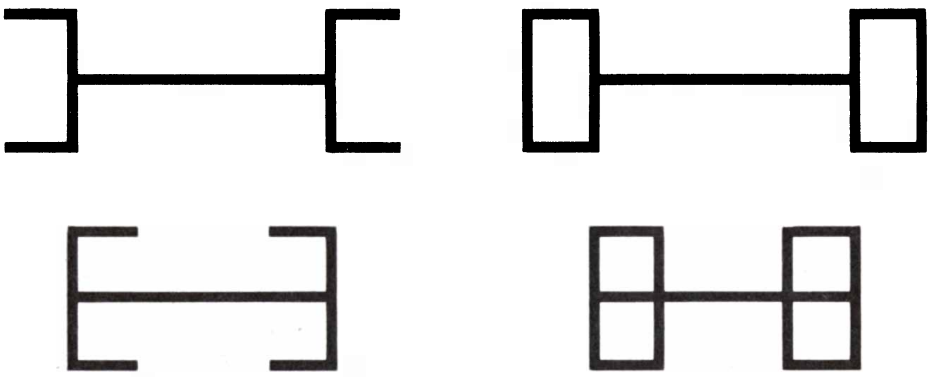


FORESHORTENING, the compression of vertical dimensions toward the top of a drawing, is achieved by parallel horizontal lines that get closer to one another toward the top of the drawing. The vertical test lines (*right*) are the same length, but the upper one seems longer because of foreshortening. There is no illusion for the horizontal test lines (*left*). This result confirms that the perception of size here does not depend on depth, which would lengthen horizontal lines as well as vertical ones. Perception of size seems to depend on the perspective scale of size.

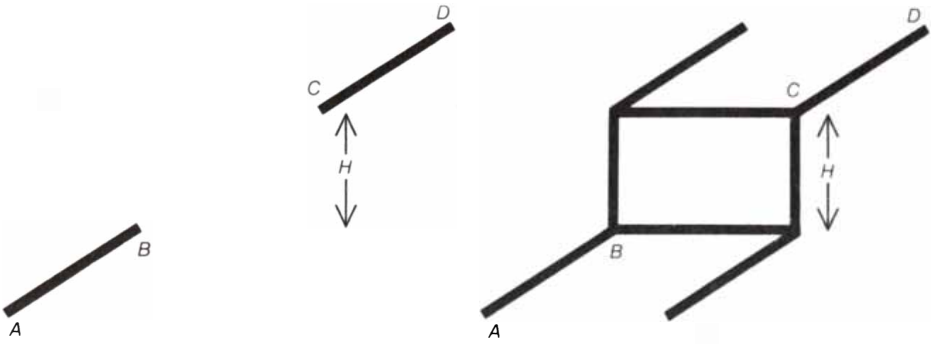
based not on apparent depth but on perspective explain illusions that do not have an interpretation as obvious in terms of perspective as the Ponzo illusion does? How does it explain the Müller-Lyer or the Poggendorff illusion? The resemblance between the Müller-Lyer figure and objects and scenes involving depth has been recognized since the 19th century. Most recently Gregory has pointed to the similarity between the two Müller-Lyer configurations and the corner of a building seen respectively from the inside and the outside. On the other hand, Tausch and other investigators, including me, have found that the Müller-Lyer illusion is close to being the sum of the separate effects of the four individual oblique lines forming the arrowheads on the length of the horizontal line, effects that seem to be divorced from depth. Two lines that form acute angles are perceptually shortened, whereas two that form obtuse angles are lengthened. It is the distortion of these fundamental angle relations that must be understood in order to comprehend the Müller-Lyer illusions and other illusions in which angles influence perceived line length.



QUADRANGLES ON THE RETINA are usually the projections of rectangles in three-dimensional space. A perceptual mechanism that lengthens horizontal lines bounded by obtuse angles and shortens horizontal lines bounded by acute angles would compensate for the difference in size caused by perspective projection. The arrows indicate the direction of the perceptual expansion or contraction. The colored dots mark the midpoints of the horizontal lines of the quadrangle at the right. The midpoints appear to be too close to perceptually shortened ends.



RECTANGULAR MÜLLER-LYER FIGURES (*left*) also create an illusion, although it is much weaker than the one created by the normal Müller-Lyer figures that are shown in the illustration on page 103. The illusion gets quite strong when the appendages are closed up to form rectangles (*right*). No theory of geometrical illusions can account for this phenomenon.



COLLINEAR LINES ON THE RETINA need not necessarily represent collinear lines in three-dimensional space. The points *B* and *C* at the left could represent an interruption in a continuous receding horizontal dimension, in which case the line *ABCD* would lie on a single horizontal plane in three-dimensional space. The other possibility (*right*) is that *B* and *C* represent points that are separated both horizontally and vertically, in which case *AB* and *CD* would lie on different horizontal planes and would therefore be noncollinear in three-dimensional space. The presence of parallel vertical lines in the Poggendorff figure favors the arrangement at the right, and so the visual system interprets the lines *AB* and *CD* as having different heights.

Do the lengthening of lines bounded by obtuse angles and the shortening of lines bounded by acute angles promote perceptual constancy? They do. An overwhelmingly large proportion of the obtuse and acute angles formed on the retina in the course of human visual experience are projections of right angles. A perceptual mechanism that lengthens horizontal lines bounded by obtuse angles and shortens horizontal lines bounded by acute angles would compensate for the differences in size created by perspective projection [*see top illustration at left*].

The illustration also shows that the process I have described would promote size and shape constancy regardless of whether the obtuse angles represent far corners and the acute angles near ones or the obtuse angles represent near corners and the acute angles far ones. It is therefore not surprising that the presence or absence of these distortions has nothing to do with depth. The parts of the perspective projection of a rectangle that form obtuse angles are always the most contracted and the parts that form acute angles are always the most expanded. The fact that the obtuse angles have a much larger effect in the Müller-Lyer illusion suggests that constancy scaling tends to equalize lengths by expanding the most contracted parts (bounded by obtuse angles) to match the rest of the projection rather than to diminish the more expanded parts (bounded by acute angles) to match the rest of the projection.

If contraction and expansion illusions serve to compensate for properties of the projection that are not properties of the scene that was projected, then the greater the distortions caused by perspective projection are, the greater is the illusion needed to compensate for the distortion. That is the case with the Müller-Lyer illusion, which gets stronger as the angles deviate more from right angles and as the arrowheads become larger. (There is a limit I cannot account for, however, on the effect the size of the arrowheads has on the magnitude of the illusion.)

I want to turn now to the Poggendorff

illusion and discuss what it means in terms of perspective. The oblique lines of the Poggendorff figure do not seem to be significantly misaligned when they are viewed alone. Why does insertion of the vertical parallel lines disrupt the apparent alignment of the oblique lines to such a great degree? Although collinear lines in the three-dimensional world always project as collinear lines on the retina, noncollinear lines in the three-dimensional world can also project as collinear lines. I contend that the parallel lines of the Poggendorff figure, particularly the components forming obtuse angles, provide a context suggesting that the oblique lines do not represent collinear lines in three-dimensional space. As a result the oblique lines do not appear to be collinear. This is shown in the bottom illustration on the opposite page.

From the point of view of perspective, oblique lines represent receding horizontal lines. When two oblique lines are lined up, as they are in the Poggendorff illusion, the conventions of perspective dictate that the space between them on the picture plane could represent one of two possible arrangements. The points *B* and *C* in the illustration could represent an interruption in a continuous receding horizontal dimension, in which case the line *ABCD* would lie on a single horizontal plane in three-dimensional space. The other possibility is that *B* and *C* represent points that are separated both horizontally and vertically, in which case *AB* and *BC* would lie on different horizontal planes and would therefore be noncollinear in three-dimensional space.

Visual processing favors the noncollinear arrangement when details placed within the gap *BC* are consistent with equidistance rather than with a depth difference for *B* and *C*. The parallel lines of the Poggendorff figure are particularly effective in this regard because they place *B* and *C* on a plane seen head on. What actually seems to happen is that the context changes the arrangement from one in which *AB, BC* and *CD* are each seen as having less slope than they do on the picture plane, because they all represent horizontal lines, to one in which the apparent slopes of *AB* and *CD* remain the same, because they continue to represent horizontal lines, whereas the apparent slope of *BC* increases, because it no longer represents a horizontal line. The result is a disruption of perceived collinearity. This explanation is supported experimentally by the work of Ross H. Day and R. G. Dickinson of Monash University in Australia, who asked experimental subjects to estimate the slopes of *AB, BC* and *CD* in the Poggendorff figure.

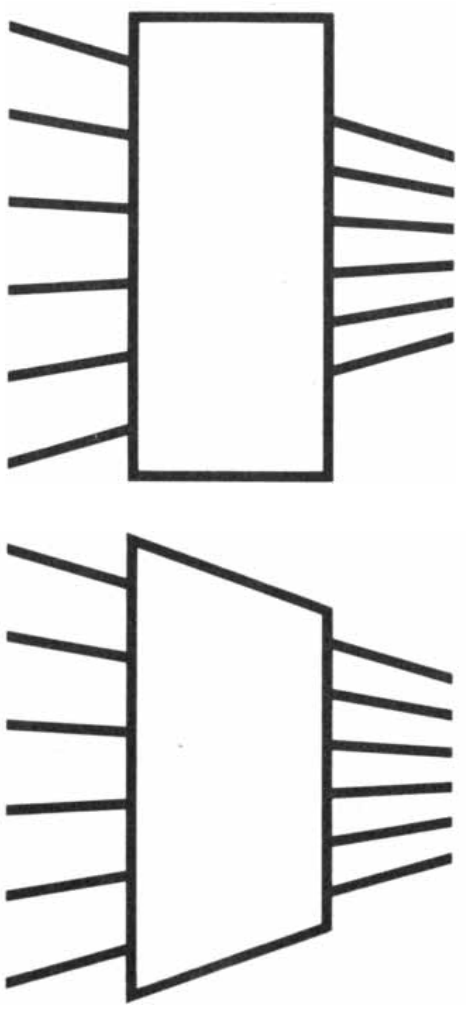
A consequence of my explanation of the Poggendorff figure is that the illusion should be greatly reduced if the length of the parallel lines is changed so that they outline not a plane seen head on but an appropriately receding plane, in which case *AB, BC* and *CD* are collinear in three-dimensional space. This change can be achieved by making the ends of the parallel lines and the oblique lines converge on a single vanishing point. When the vanishing point is clearly defined, the magnitude of the illusion is reduced by half. This provides strong evidence that the context relevant to the three-dimensional layout represented by two-dimensional lines strongly influences the perceived collinearity.

The evaluation of a functional theory based on perspective decoding is hampered by the fact that so little is known about perceptual responses to perspective, in spite of the fact that much of Western art has been based on perspective for 500 years. It is unlikely, however, that all illusions are the result of perspective decoding. For example, no obvious constancy function is served by illusions of orientation assimilation, such as the Lipps figure. Yet my functional theory, unlike the other theories, can handle in a general way the maintenance of illusory responses under normal three-dimensional viewing conditions, where the responses are reinforced because they promote accurate perception, and the diminution of illusory responses under two-dimensional viewing conditions, where the responses are inhibited because they serve no function.

Work remains to be done on how the diminution is achieved and on what role eye movements play in the process. Stanley Coren of the University of British Columbia and Joan S. Girgus of Princeton University have proposed that illusions that strongly diminish in the course of repeated exposure have a considerable judgmental or learned component, whereas illusions that only weakly diminish, for example assimilation illusions such as the Delboeuf one, are chiefly built in. Although constancy responses are traditionally regarded as being based on learned algorithms, there is no reason some of the responses could not be built-in ones that evolved because they promoted constancy. In principle theories emphasizing function are not incompatible with theories emphasizing mechanism.

Functional theories based on perspective are supported to some degree by responses to geometrical illusions among peoples who do not live in a world dominated by rectangular rooms, buildings and cities. For such people the illusions tend to be not as strong as they are for people in our own kind of culture. Their response to the upside-down-*T* illusion, on the other hand, which seems to depend not on the exposure to a rectangular environment but on the foreshortening of distant terrain, is not reduced. Striking as such results are, they should not be allowed to carry much theoretical weight until it is clear the people involved in the experiments have had the same understanding of what it is they are being asked to do as subjects in our own kind of culture. If such results hold up, however, they could significantly bolster the functional theories.

Perhaps the most serious challenge to functional theories, and for that matter to all theories of illusion relying on processes in the visual system, could come from investigations into the sense of touch. Experiments have revealed touch illusions that are analogous to optical ones when subjects ran their fingers over raised versions of well-known illusory figures. At this stage, however, the work is too tentative to establish any identity between misperceptions of touch and misperceptions of vision.



ORIENTATION OF THE PLANE strongly alters an illusion for a set of collinear oblique lines with a common vanishing point. The top lines, which interrupt a plane seen head on, appear to deviate more from being collinear than the bottom lines, which interrupt a receding plane that has the same vanishing point as the lines. This provides evidence that the context relevant to the three-dimensional layout represented by lines in two dimensions strongly influences the perceived collinearity.

The Ancestry of Corn

The progenitor of modern corn is probably the wild grass known as teosinte. The hypothesis has been disputed, but it is supported by new evidence from plant breeding, archaeology and folklore

by George W. Beadle

The origin and growth of agriculture over the past 10,000 years or so has made possible a thousandfold increase in the human population and has freed a substantial fraction of the population for participation in aspects of cultural evolution other than the procurement of food. It is relatively easy to trace the origins of wheat, barley and rice, on which the emergence of agriculture in eastern and western Asia was based. In much of Asia today the wild grains from which these crops were derived are still growing, enabling plant geneticists to determine the route by which they were cultivated and—at first unconsciously and later deliberately—selected and bred for characteristics that would make the harvest easier and more fruitful. No such continuum can be established, however, for the most important food crop of the Western Hemisphere: maize, or corn (*Zea mays*). Modern corn, with its many-kerneled, husk-wrapped ear, bears no readily apparent morphological resemblance to any New World plant that might have been its ancestor. Yet corn must have been derived from some wild grass through selective breeding by the ancient Indian agriculturists of Central America. Whether there is or ever was such a thing as wild corn continues to be a subject of vigorous scientific controversy.

Modern corn is so highly domesticated that it has in a sense become a biological monstrosity. Although it is superbly adapted for producing grain, it could not survive under natural conditions because it has no satisfactory mechanism for dispersing its seed. If an ear of corn is left on its own under conditions favorable for germination, it will give rise to a group of seedlings so densely clustered that they will compete among themselves for water and soil nutrients, with the result that all may fail to reach reproductive age. Without human intervention modern corn would become extinct in a few generations. Corn can survive only if man removes the kernels from the cob and plants them.

The origin of corn is a matter of more than academic interest. Corn is the most efficient of all cereals in converting solar energy, carbon dioxide, water and mineral nutrients into foodstuff. It is now the world's third most important food crop: more than 200 million metric tons is produced each year, only slightly less than the harvests of wheat and rice. Although relatively little corn is consumed directly as human food, when it is transformed into meat, milk, eggs and other animal products, it is the principal food plant of the Western world. Hybrid corns, which represent most of the corn grown today, tend to be susceptible to disease, and the strains of corn that are crossbred to produce the hybrids tend to become weakened by excessive inbreeding, so that correctly identifying the original genetic stock of corn might well supply valuable sources of genetic diversity and hybrid vigor.

There are currently two competing hypotheses about the origin of corn. One school contends that wild corn can in fact be found: it is and always was the wild grass teosinte, which grows in Mexico, Guatemala and Honduras. The hypothesis was popular among 18th- and 19th-century botanists but lost its prominence during the first half of this century. The other school contends that if there was a wild corn, it was swamped out of existence by later, man-bred strains. This school is led by Paul C. Mangelsdorf, emeritus professor at Harvard University and now at the University of North Carolina at Chapel Hill [see "The Mystery of Corn," by Paul C. Mangelsdorf; SCIENTIFIC AMERICAN, July, 1950]. Recent work by my colleagues and me, however, has provided new and impressive support for the teosinte hypothesis. It now seems quite likely that a teosinte of some 8,000 to 15,000 years ago was the direct ancestor of modern corn and was transformed into a primitive corn through human selection.

Corn was totally unknown in Europe and Asia until the arrival of Columbus in the New World in November of 1492. His men discovered on the island of Cuba great fields of the strange new plant, which was later found to be cultivated throughout the Western Hemisphere. The large populations and impressive cultural achievements of the Aztecs, the Mayas and the Incas could not have come about without corn, which not only was nutritious but also could be dried, transported and stored for long periods. The tribute list of Montezuma, the last Aztec emperor, shows that the 20 provinces of the empire were taxed an annual total of some 300,000 bushels of corn.

The first corn known to exist was identified from archaeological evidence; it dates back some 7,000 years. Over a period of a few millenniums pre-Columbian peoples bred most of the major varieties of corn that exist today, including red corn, blue corn, yellow corn, field corn, sweet corn, dent corn, flint corn, flour corn, pod corn and popcorn. By the time of Columbus between 200 and 300 varieties of corn were already in cultivation. The ancient Indian agriculturists had carried the plant from its place of origin, probably southern Mexico, and adapted it to a wide variety of temperate and tropical climates ranging from the mouth of the St. Lawrence River to what is now central Chile. Although modern plant breeders have greatly increased the yield of corn through the development of new varieties and hybrid lines adapted to various regions and resistant to diseases and insect pests, the development of corn by the Indians remains man's most remarkable plant-breeding achievement.

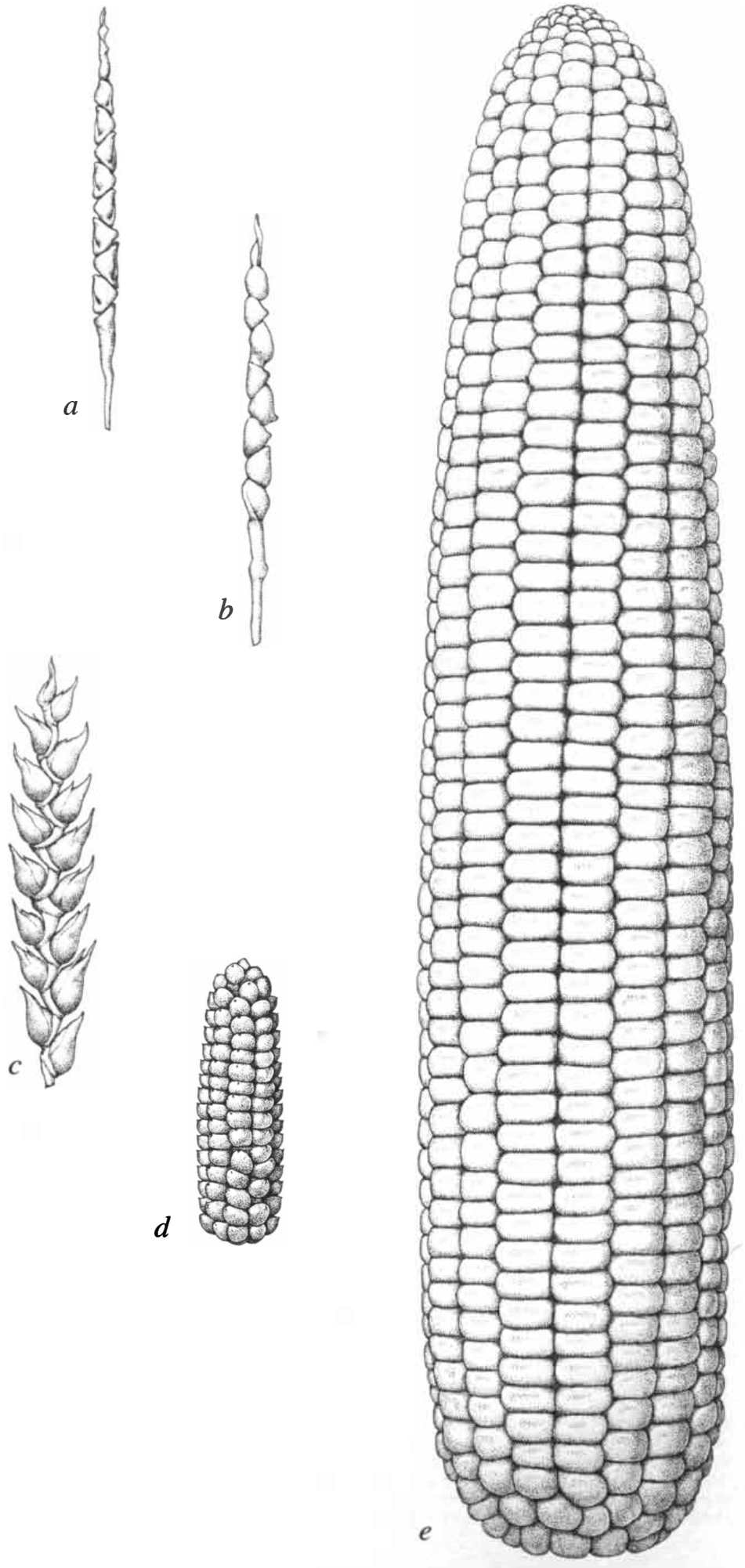
Long before Columbus the peoples of parts of Mexico and Guatemala were also familiar with teosinte. In 1780 the Spanish plant explorer Francisco Hernández described teosinte (in translation) as a plant that "looks like corn but has a triangular seed." Teosinte differs from corn in that it has several stalks growing from the base of the plant where modern corn usually has only one. Its seeds are borne in hard, individual fruit cases, which form a single row

of from six to 10. This row is called a female spike and corresponds to an ear of corn. Where the entire corn plant bears only one ear or at most a few, the teosinte plant typically produces several spikes. At maturity the individual fruits of the teosinte spike disarticulate, allowing them to be disseminated.

Although the seed spike of teosinte has only a seemingly remote resemblance to an ear of corn, teosinte was thought by early students of the question to be the most likely ancestor of corn because it grows wild in the same regions of the New World where corn was probably first domesticated. The botanist A. Vinson wrote in 1877, "Following the thinking of Darwin, teosinte is the ancestor of corn." Most of his contemporaries disagreed, because the teosinte spikes were so different from ears of corn that it seemed unlikely human selection could have transformed one into the other.

Later what seemed a more likely candidate was discovered: a plant intermediate between teosinte and corn. It bore a small ear with pointed kernels resembling a dog's teeth partially projecting from individual jackets and was therefore called dog corn or coyote corn. This type was soon found to be a naturally occurring first-generation cross between teosinte and corn that can easily be repeated experimentally and is quite fertile. The fertility of the cross indicated that the genetic relation between teosinte and corn was much closer than had been assumed.

Such was the state of the field in 1928, when I began graduate work at Cornell University with R. A. Emerson. Emerson had shortly before demonstrated that teosinte is a "short day" plant: it wants no more than 12 or 13 hours of

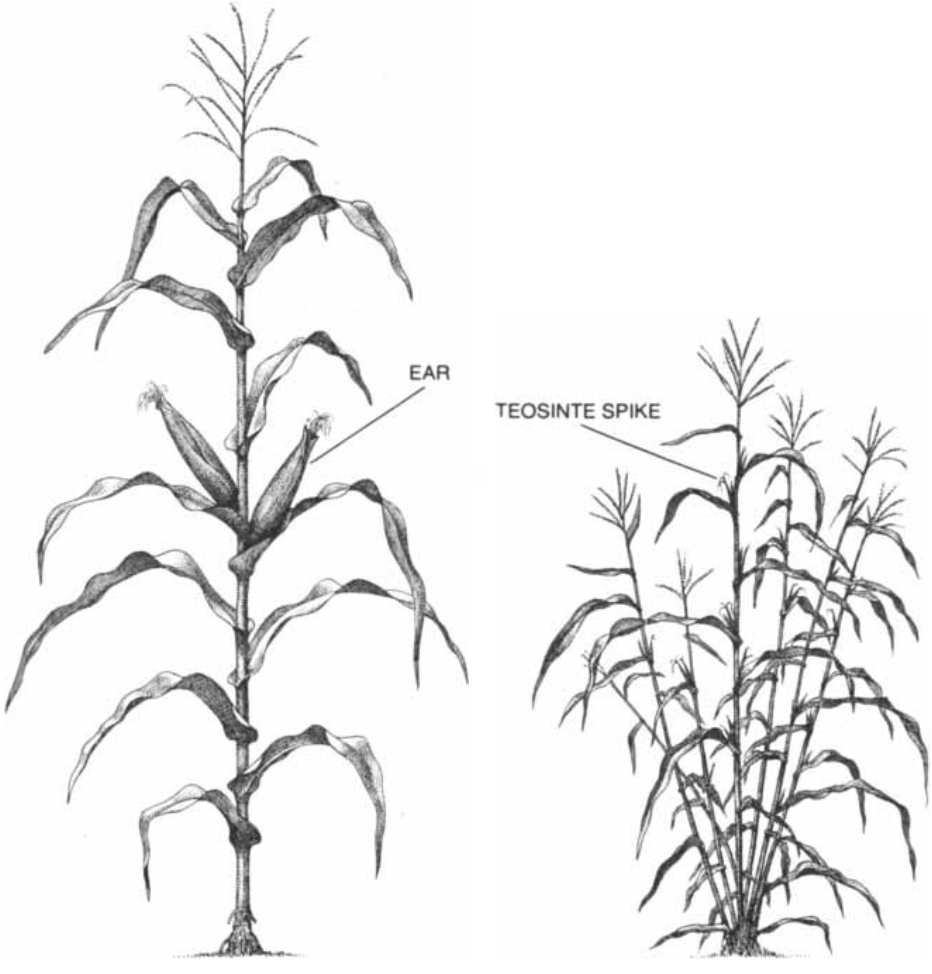


DOMESTICATION OF CORN from the wild grass teosinte by a few millenniums of human selection has been reconstructed by breeding experiments undertaken by the author. All the specimens are drawn approximately life-size. The "spike" of teosinte (*a*) is equivalent to the ear of corn and consists of a single row of kernels in hard, shell-like fruit cases. On ripening, the spike shatters, scattering the seeds. Crosses between teosinte and corn yield a modified teosinte (*b*), which may be similar to an early transitional form. A single mutation in teosinte can also give rise to a tunicate variety (*c*), in which the hard fruit cases have been converted into soft, husklike glumes from which kernels can be threshed with ease. This mutation may well have been a crucial step in the domestication of teosinte. Crosses between teosinte and modern corn give rise to small, primitive ears (*d*), which are similar to 7,000-year-old archaeological specimens found in the southwestern U.S. and Mexico. Modern corn (*e*) is a biological monstrosity created by prolonged domestication. It is well adapted for production of grain but is unable to survive under natural conditions.

sunlight a day and warm temperatures. It grows wild only where these conditions exist. Emerson was able to grow teosinte at the Cornell latitude by artificially shortening the summer day; he did so by putting dark boxes over the plants for the critical period in which flowering is initiated.

As a part-time graduate-student assistant I was assigned to study the cytology and genetics of corn-teosinte crosses. Emerson and I confirmed the fertility of the cross and showed that the 10 chromosomes in the cells of Mexican teosinte were highly compatible with the 10 chromosomes of corn. The chromosomes paired normally during the formation of sex cells in the crossed plants, and the nine chromosomes we could identify with markers formed pairs in the crosses and exchanged segments in essentially the same way they did in pure corn. Our conclusion was that cytologically and genetically corn and Mexican teosinte could even be considered one species. The hypothesis that wild teosinte was the direct ancestor of cultivated corn therefore seemed entirely plausible. We regarded the problem as being essentially solved, believing relatively few minor gene changes could convert the wild plant into the cultivated one and had probably done so.

I remember well discussions with Emerson in which he pointed out that just two mutations would be required to make teosinte a more useful food plant: one mutation to a nonshattering spike, so that the kernels would not be scattered and lost, and the second to a soft fruit case, so that the kernels could be threshed free. These postulated mutations are not far-fetched; they are much the same as the known changes in the evolution of cultivated wheat, rye, barley and oats from their wild ancestors. In spite of the strong arguments in favor of the teosinte hypothesis, however, it began losing support in the late 1930's because of studies suggesting so many genes of corn were different from those of teosinte that they could not have originated in a few thousand years of domestication.



SIZE AND SHAPE of a modern hybrid-corn plant and a teosinte plant are markedly different, a fact that accounted for early resistance to the idea that corn could have arisen from teosinte by human domestication. Corn has a single thick stalk, whereas teosinte has numerous thin stalks. Corn also bears one or two large ears, whereas teosinte bears numerous small spikes at each node. The tassels (male flowers) of the two plants are similar, however. Where corn could not survive without human intervention, teosinte is well adapted to survival in the wild.

Forty years after my work with Emerson on teosinte as a graduate student, after a career spent studying the genetics of fruit flies and bread mold, I decided to return to the study of the relation between teosinte and corn because of the disrepute into which the teosinte hypothesis had fallen. On my retirement from academic administration at the University of Chicago I set out to determine the approximate number of genetic differences between teosinte and corn by growing large-scale populations of corn-teosinte crosses, populations that were far larger than any that had been grown before.

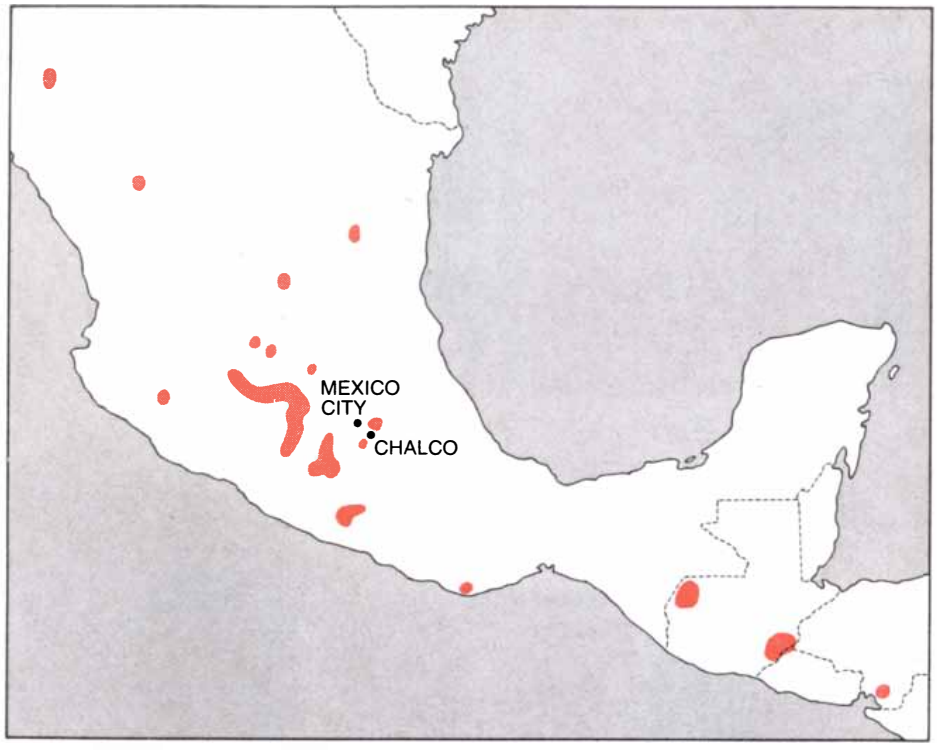
The need for a large number of segregating crosses is a consequence of Mendelian genetics. Mendel's laws state that if there is a difference of only one gene between the parents, each original parental type will reappear in the plants of the second generation with a statistical frequency of one in four. The progression is geometric: with two independent gene differences the frequency of reappearance is one in 16 and with 10 gene differences the frequency approaches one in a million. If teosinte and corn differed by even as few as 10 genes, I would not be able to grow and examine populations of crosses large enough to produce good segregant offspring equivalent to the original teosinte and corn parents. I therefore resolved to grow as many as 50,000 second-generation plants if necessary; that number would give a reasonable chance of finding types similar to the original parents if there were no more than six or seven major gene differences. Since teosinte is a short-day plant that will not mature at U.S. corn-belt latitudes, I made arrangements with Edwin J. Wellhausen to grow the hybrids at the El Batán station of the International Maize and Wheat Improvement Center near Texcoco in Mexico. The work there was done in collaboration with Mario Gutierrez of the El Batán station and Walton C. Galinat of the University of Massachusetts.

To minimize the probable number of gene differences between the parent stocks, we selected a particularly primitive Mexican variety of corn known as Chapalote and the most cornlike variety of Mexican teosinte, known as Chalco. After several seasons of crossing, the frequency of reappearance of good parent corn types and teosinte types in the second-generation crosses proved to be about one in 500. This frequency indicated that there could be approximately five major and independently inherited gene differences between the parental stocks of corn and teosinte, far fewer than had been suspected previously.

These experiments clearly demonstrated that the genetic differences between corn and teosinte are not so great

that they render untenable the hypothesis of an ancestral relation between teosinte and corn. It therefore seems reasonable that pre-Columbian man could have selected and preserved the relatively few mutants required to produce a useful crop plant from teosinte, namely a primitive corn.

Nevertheless, Mangelsdorf has continued to be a major critic of the teosinte hypothesis. In 1939 he and Robert G. Reeves of Texas A. and M. University proposed that teosinte was not ancestral to corn but rather was a relatively recent species derived from the crossing of corn and a grass of the genus *Tripsacum*. Their principal argument was that teosinte is intermediate between *Tripsacum* and corn in many morphological traits. I never believed this hypothesis and said so in 1939. My reasons were several. First, corn and *Tripsacum* have never been known to cross naturally, in spite of the fact that they grow in close proximity over millions of acres. Man-made crosses can be accomplished only with special techniques. Second, none of the 18 chromosomes of *Tripsacum* pairs normally with any of the 10 chromosomes of corn. Third, the man-made crosses of corn and *Tripsacum* are completely male-sterile.



TEOSINTE STILL GROWS WILD in parts of Mexico and Guatemala. It does not grow at temperate latitudes because it is a "short day" plant: it wants only 12 or 13 hours of sunlight per day. It also needs warm temperatures. According to the author's hypothesis, pre-Columbian Indians bred it to yield 200 or 300 varieties of corn, adapted to a wide range of conditions.

The *Tripsacum* hypothesis has been advocated by its proposers for more than a third of a century, with the result that it has thoroughly permeated genetic, botanical, plant-breeding and other scientific literature, often being transformed in the process from a hypothesis into an established fact. Two recent findings, however, have raised serious doubts about its validity. First, contrary to Mangelsdorf's assumption that teosinte is of relatively recent origin, two seeds much like those of Chalco teosinte were recovered recently from an undisturbed preceramic horizon dating back 7,000 years at a site about 20 miles southeast of Mexico City. Second, Galinat has demonstrated that the hard fruit case of teosinte, which Mangelsdorf postulated was contributed by *Tripsacum* as a block of closely linked genes, could not have arisen in that way. The reason is simply that the fruit case of *Tripsacum*, in contrast to that of teosinte, appears to be governed by genes that are not closely linked and hence could not have been transmitted to teosinte as a block.

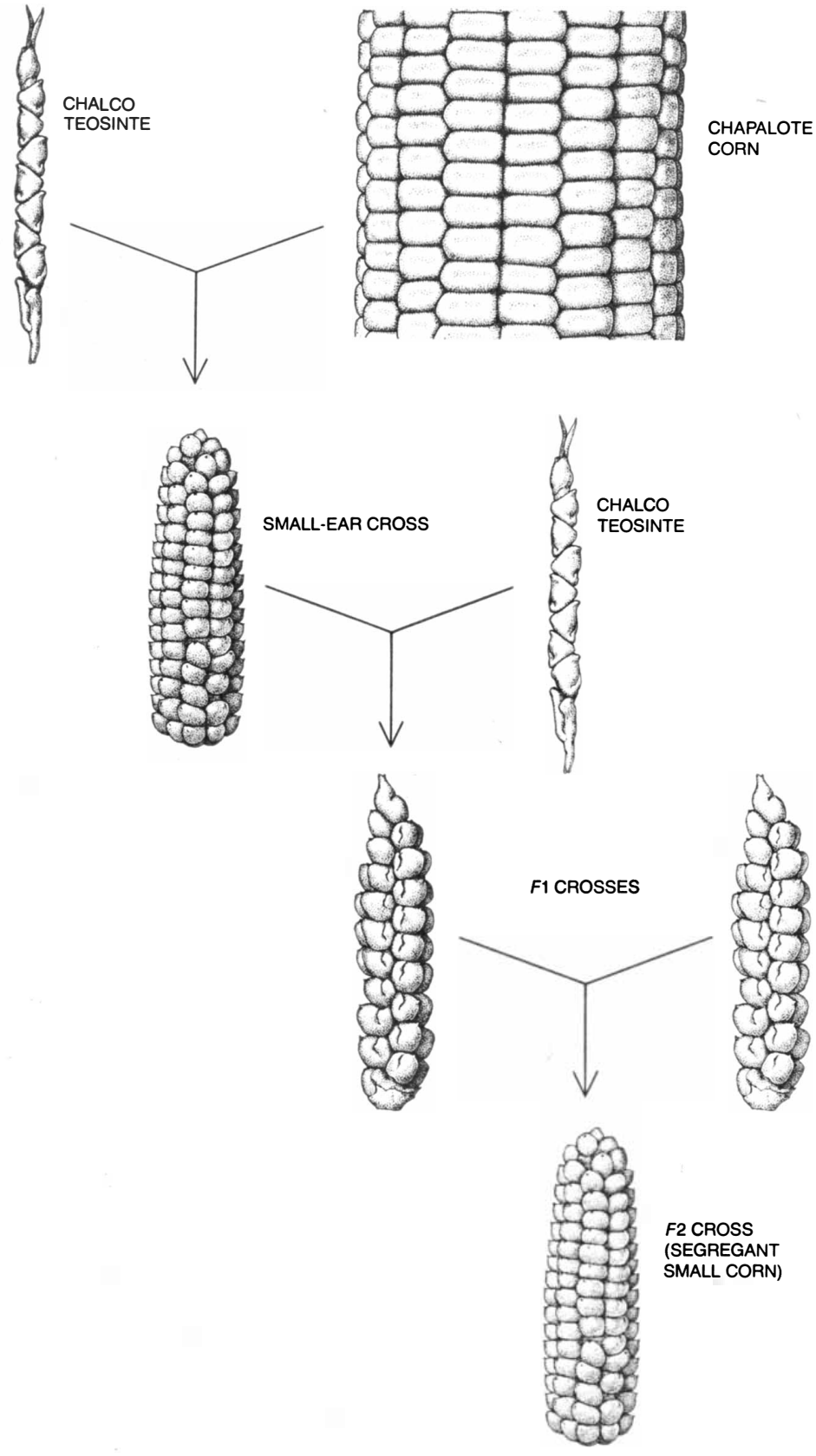
In response to the evidence against the *Tripsacum* hypothesis Mangelsdorf and his associates recently made an alternative proposal. They now contend that modern corn is derived from a wild type of popcorn known as pod corn and that teosinte arose from this wild corn by mutation. Their position is a dramatic reversal, since Mangelsdorf had formerly maintained that teosinte and corn were so different genetically that corn could not have descended from teosinte. If corn could have given rise to teosinte, surely the reverse is also possible. I would say it is much more probable, since teosinte is a highly successful wild plant and corn is not.

If there ever was a wild corn growing independently of teosinte, it is now extinct or is undiscovered in spite of extensive searches. Although it is logically impossible to prove that corn never existed as a wild plant, I see no compelling evidence to indicate that it did. Indeed, many ecological factors would have favored the survival of teosinte over the hypothetical wild corn. For one thing, the brittle spike of teosinte is a far more effective seed-dispersing mechanism than the cob of a wild corn would have been. Teosinte has many other characteristics that enhance its survival in the wild. Its seeds have a dormancy mechanism that ensures their germination under conditions favorable for survival, and the hard fruit case of teosinte provides significant protection to the enclosed kernels compared with the naked or poorly protected kernels of corn. Controlled experiments have shown that various rodents, birds, insects and other animals that feed on seeds much prefer corn kernels to teosinte seeds.

The colors and patterns of the teosinte fruit cases camouflage them remarkably well, thereby reducing animal feeding. The fruit cases vary from gray to brown to black, often with mottling, spotting or striping; in their color and their oblate shape they effectively mimic the pebbles and soil particles of the ground on which they fall. Moreover, unlike corn, teosinte responds adaptively to environmental conditions. In fertile soil with optimum moisture the plants are large and produce abundant seed, but under adverse conditions their growth is limited and they produce just as much seed as the available resources allow. Corn, on the other hand, gambles on producing a minimum of one ear, and if it does not do so, it produces no seed at all.

Therefore if a wild corn did exist, it seems highly improbable that it and teosinte could have shared the same habitat or overlapping habitats. Teosinte, with its many attributes of a successful wild plant, would surely have competitively replaced the wild corn. Conversely, if teosinte and wild corn occupied different habitats, it is difficult to believe they could have remained as completely interfertile as the evidence shows they have been for thousands of years.

What is the archaeological evidence for and against the two competing hypotheses of the origin of corn? In the 1940's many cobs were recovered from ancient refuse piles in a number of dry caves in the southwestern U.S. and the Tehuacán valley of southern Mexico. Carbon-14 dating has shown that the oldest Mexican cobs were grown approximately 7,000 years ago. Because successive generations of human occu-

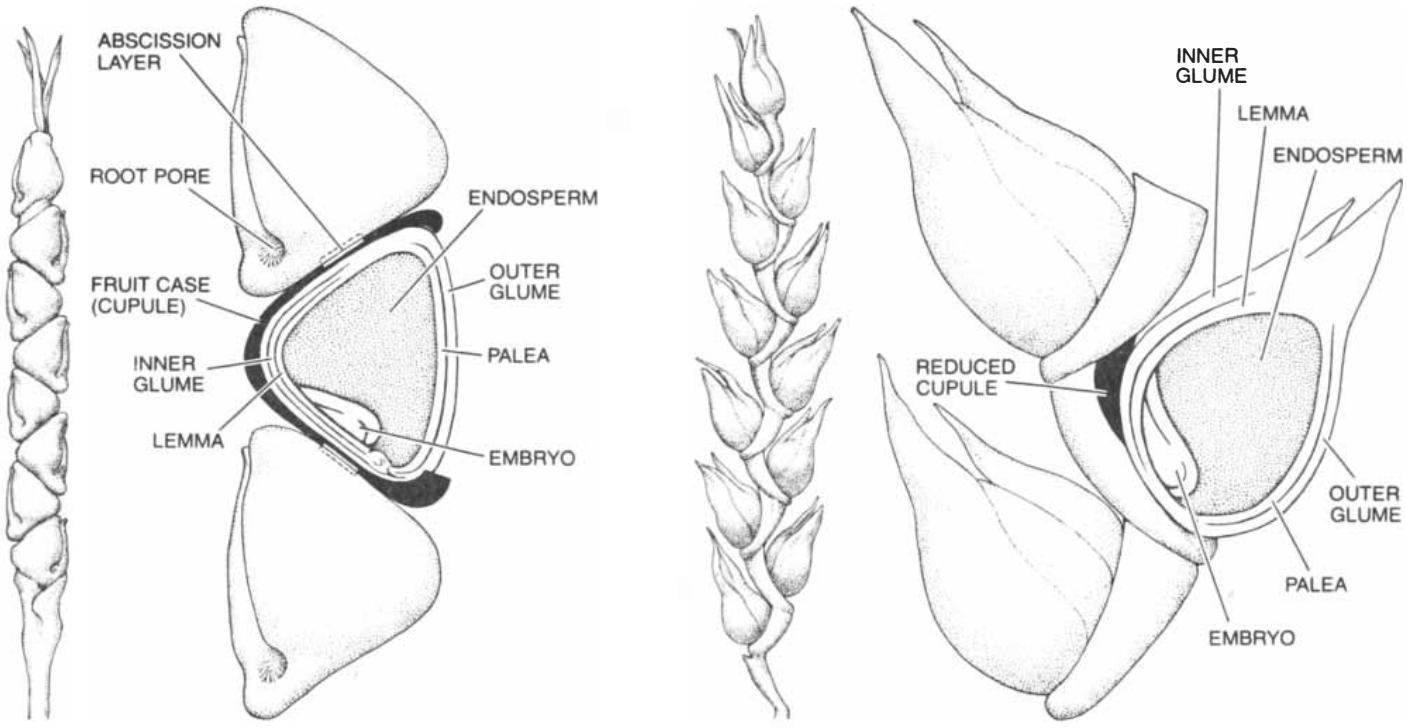


BREEDING EXPERIMENTS were done on a large scale by the author and his colleagues in order to determine the approximate number of gene differences between teosinte and corn. According to Mendel's laws, if two related varieties are crossed, progeny similar to the parent forms should segregate out in the second generation in direct proportion to the number of gene differences between the parents. In order to minimize the possible number of gene differences, a cornlike teosinte known as Chalco and a primitive corn known as Chapalote were selected for the experiment. Out of 50,000 crosses about one in every 500 closely resembled either parent, indicating that Chalco teosinte and Chapalote corn differ by only about five major genes. That corn arose from teosinte over a few millenniums of domestication is genetically plausible.

pants of the caves allowed refuse and trash to accumulate on the cave floor to a depth of about six feet, the cobs in different strata reveal a distinct sequence of evolutionary changes due to domestication. The oldest cobs are only about an inch long, and they bore about 50 to 60 kernels. Mangelsdorf has interpreted these cobs as being the remains of a genuine wild corn, mainly because of their reported uniformity and because they were uncovered in strata older than those in which the remains of other plants known to have been cultivated have been found.

The alternative explanation, and to me the more plausible one, is that the oldest archaeological cobs represent stages in the transition of teosinte to corn through a process of human selection. What is the evidence for this alternative? First, the ancient cobs are much closer morphologically to teosinte than the cobs of modern corn. For example, some of the earliest cobs are brittle and have two rows of kernels, both teosinte traits. Second, cobs closely matching the archaeological cobs are readily recovered in second-generation and later crosses of modern corn and teosinte. Third, if the earliest archaeological specimens were wild corn and dispersed their seeds by means of a brittle cob, it is difficult to explain how the cobs could have survived harvesting, transportation to the caves where they were found, shelling of the kernels and mixing with the other debris on the cave floor. If the cobs were not brittle, however, and the cobs in question clearly do not appear to have been, how would they have been able to disseminate their seeds without human help?

In spite of the impressive evidence in favor of the direct origin of corn from teosinte, Mangelsdorf and his associates persist in believing the direction of the change was from a primitive wild corn to teosinte. The major line of evidence to support this view rests on the analysis of fossil pollen. Corn and teosinte pollens are indistinguishable except by size: corn pollen is significantly larger. In 1954, during the excavation of the Belles Artes site in Mexico City, 14 analyzable pollen grains were recovered from construction drill cores at a depth of 70 meters, corresponding to an age of 80,000 years. Of the 14 fossil pollen grains five were judged by Elso S. Barghoorn of Harvard and his colleagues to be those of corn on the basis of size. Barghoorn concluded that because the stratum in which the pollen was found antedated the presence of man in the Western Hemisphere, the large pollen grains must have come from an ancient wild corn that had originated separately from teosinte. Mangelsdorf later wrote that the fossil pollen "shows beyond a reasonable doubt



TUNICATE MUTATION could have been a crucial step in the evolution of teosinte into corn. The teosinte spike, with each kernel completely encased in a hard fruit case, is shown in the drawing at the left. The tunicate mutant, with the fruit cases reduced to a series of shallow cupules and the kernels enclosed in individual husklike glumes, is shown in the drawing at the right. The tunicate form has less of a tendency to shatter on ripening and the kernels can be easily threshed free of the chaff, making it a much more useful food plant. A single mutation is therefore capable of transforming teosinte into a logical progenitor of modern corn. Further selection could then have reduced the size of the tunicate glumes, yielding naked kernels that could be removed from the cob of the plant without the need for threshing.

that the ancestor of cultivated corn was corn, not teosinte or any other of corn's relatives."

There is reason, however, to doubt this interpretation. For one thing, there is a close correlation between the size of pollen, the length of the ear and the distance a pollen tube must grow through a silk or style. Primitive corn and teosinte, with their tiny ears, have small pollen grains that cannot fertilize the kernels of large modern ears with their long silks. Assuming that the postulated wild corn was at least as small and primitive as the 7,000-year-old archaeological ears found in the Tehuacán valley, the correlation of pollen size and short ears indicates that the drill-core pollen is clearly too large for wild corn. One possible explanation is that the large drill-core pollen arose from a teosinte with four sets of chromosomes (a tetraploid) whose pollen had not yet been reduced by natural selection to the minimal effective size. Such tetraploids are known to occur.

A second and more likely possibility is that the pollen in question is not ancient at all but is the result of contamination of a drill core during or after the samples were taken. Indeed, extraordinary care would have been required to have completely avoided contamination of the cores by modern pollen, and it is doubtful that such precautions were observed with drill cores taken primarily for construction purposes. For these reasons it seems clear that the fossil-pollen evidence is at best inconclusive.

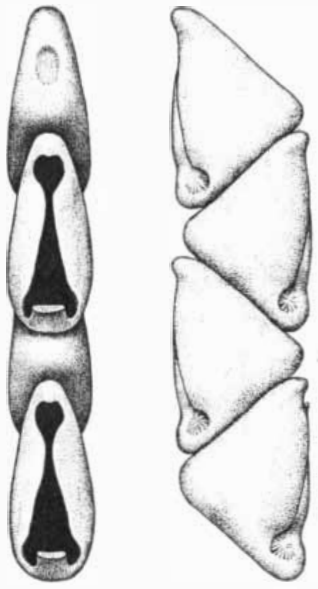
A second argument against teosinte as an ancestor of corn is that it is not clear how its nutlike seeds with their heavy shells could have been used as food by primitive man. This view has persisted in spite of early Spanish evidence that teosinte seeds were indeed used as food, since young teosinte spikes can be eaten before the fruit cases harden. In 1939, without knowledge of any documented use of teosinte as food, it occurred to me as a boyhood lover of popcorn that perhaps teosinte seeds would pop. A simple experiment demonstrated that they do; the popped kernels, exploded out of their fruit cases, are indistinguishable from popped corn. That teosinte could have been prepared in this way by preceramic man seems likely. Kernels placed in a fire, on glowing embers, on hot rocks or on heated sand pop very well and either pop free of the fire or can be retrieved with sticks or tongs.

I have also had a try at cracking dried teosinte seeds with the primitive grinding stones employed by the pre-Columbian Indians some 8,000 years ago: the mano and metate. The Field Museum of Natural History in Chicago kindly lent me 8,000-year-old archaeological specimens for the experiment. I concluded that in a day an energetic person motivated by sufficient hunger could separate enough partially shell-free teosinte meal from the ground shell-kernel mixture by a simple water-flotation method to feed a small family for a day or more. Whole mature teosinte seeds can be eaten directly by first soaking them in water until they have been sufficiently softened to chew. The shells can be either swallowed or selectively spit out.

I know of a simple and plausible mechanism by which teosinte could have been converted by domestication into a more useful food plant and a logical precursor of corn. The modern mutant corn called pod corn or tunicate corn has soft, leaflike flaps known as glumes covering each kernel. Mangelsdorf and his associates have suggested that the early archaeological corn could have had the genetic equivalent of the tunicate trait, since archaeological corn has noticeable tunicate glumes whereas the glumes of modern corn are reduced to scales on the cob that cannot be seen until the kernels are removed. In fact, Mangelsdorf has employed tunicate types of modern corn to breed reconstructions of the earliest archaeological specimens.

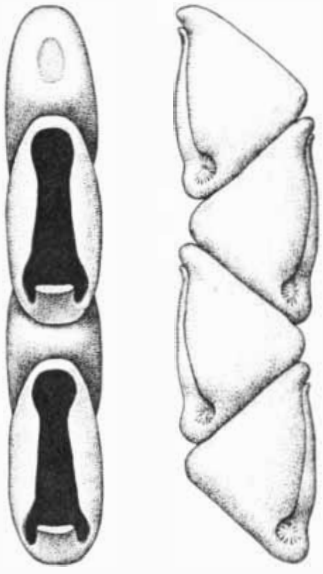
I now believe the tunicate trait had a far more important bearing on the origin of corn than has been recognized. Galinat generously gave me several strains of a teosinte cross carrying tunicate genes from corn, and with these crosses I was able to produce many additional strains by backcrossing them with teosinte. The resulting modified teosintes had many of the characteristics Emerson had postulated as being the

1



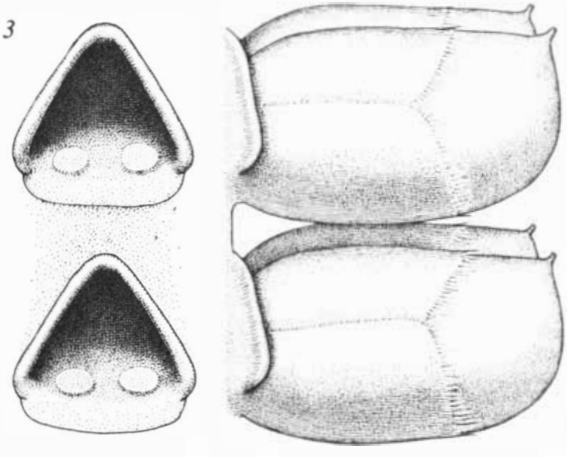
NOBOGAME TEOSINTE

2



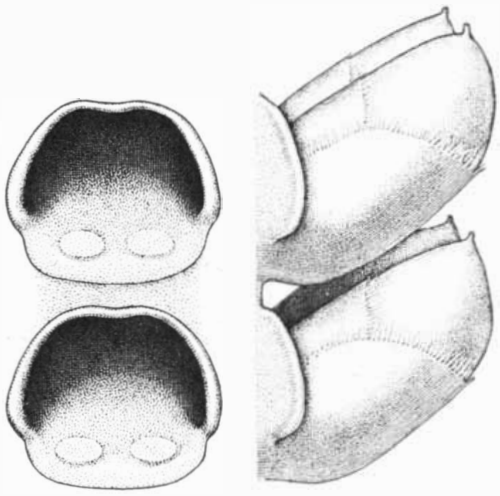
CHALCO TEOSINTE

3



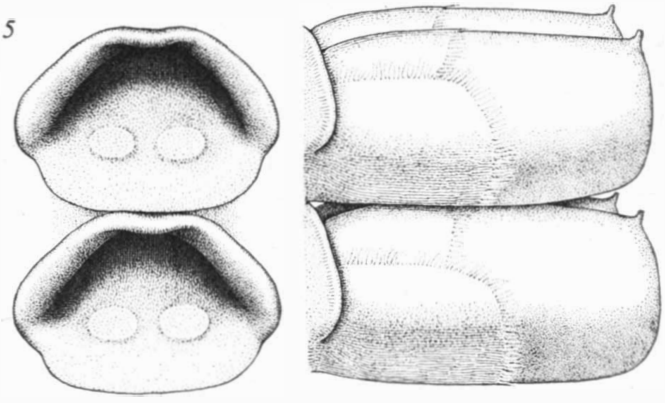
ARCHAEOLOGICAL CORN

4



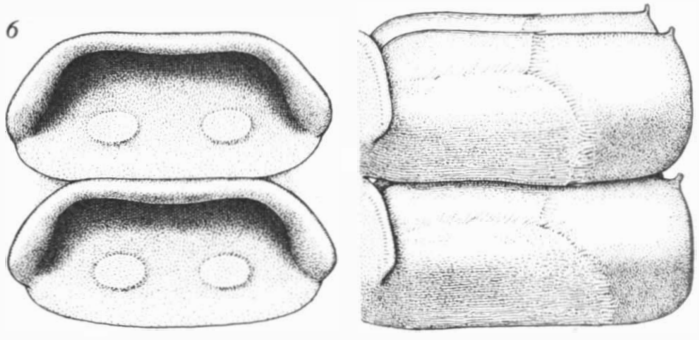
CORN-TEOSINTE CROSS

5



CHAPALOTE CORN

6



MODERN CORN

EVOLUTION OF THE CUPULE from the fruit case of teosinte to the compacted cob of modern corn provides strong support for the origin of corn from teosinte. Two progressive changes are apparent in this evolutionary sequence. First, the dimensions of the cupule gradually shift from a vertical orientation in teosinte to a horizontal orientation in modern corn. Second, the space between the adjacent cupules, corresponding to the back of the fruit case in teosinte, is retained in primitive corn but then is gradually reduced. In modern corn cupules are closely packed to yield a rigid, dense cob. Drawings are based on work of Walton C. Galinat of University of Massachusetts.

first steps in the transformation of teosinte into corn. The fruit cases of the tunicate-teosinte crosses were converted into shallow, softer, cuplike structures called cupules. Moreover, the glumes that protect the kernels and seal them within the fruit case were enlarged and softened, so that they could be scraped off with a fingernail to reveal the kernel underneath. There was also much less tendency for the modified teosinte spikes to shatter on ripening.

A single mutation is therefore capable of transforming teosinte, with its hard, shell-like fruit case, into a plant from which the naked kernels can be threshed with ease. If the tunicate-mutant change was indeed an early step in the transformation of teosinte into corn, the presence of the tunicate trait in the archaeological specimens would be accounted for. The plausibility of this hypothesis seems further enhanced by the fact that the gene responsible for the tunicate trait is dominant rather than recessive. If in the early stages of human cultivation a tunicate mutant of teosinte had arisen and had been selected for planting, it would not have seemed to disappear in the next generation through outcrossing, as a recessive mutant would have. The reduced tunicate glumes of modern corn can be easily explained. After the development of the tunicate form of teosinte, in subsequent stages of domestication the ancient Indian plant breeders would clearly have found it desirable to reduce the tunicate trait by selecting for cobs with naked kernels, which could be removed without threshing.

A new and important piece of evidence supporting the transformation of teosinte into corn has been provided by Galinat, who compared the cupules of the oldest archaeological corn with those of both modern teosinte and modern corn. In teosinte and sometimes in primitive corn the cupules are higher than they are wide. In contrast, in northern flint corn and its modern descendants the cupules are wider than they are high. Galinat has examined a series of archaeological cobs from New Mexico dating from 1400 B.C. to A.D. 1400 and has found that the cobs can be placed in an evolutionary continuum from teosinte to modern corn on the basis of progressive modifications of the cupule. In teosinte each cupule in a vertical row alternates with an interspace representing the back of the adjacent fruit case. This space also undergoes a progressive evolutionary change; it is retained in archaeological cobs but is progressively reduced, so that in modern corn-belt corn the cob is stiffly clad in vertical rows of compacted cupules, with virtually no space between them.

The fruit case–cupule relation seems overridingly persuasive as evidence that the direction of change was from teosinte to corn. How else can one account for the cupule of the corncob, clearly present in the earliest archaeological specimens? As Galinat has pointed out, the cupule indicates that corn must have evolved from an ancestor with a cupulate fruit case. With *Tripsacum* excluded, teosinte remains the sole candidate. There is no other living or known extinct species that so logically satisfies all the criteria.

Another line of evidence that supports the teosinte hypothesis comes from the study of regional language and folklore. The word teosinte comes from the Aztec *teocentli,* meaning "God's ear of corn." How did the Aztecs know that teosinte was related to corn? Is it a case of cultural memory? Moreover, in many parts of Mexico teosinte is known as *madre de maiz* ("mother of maize"), presumably from an earlier native designation. With many other existing native names for teosinte in the various parts of its range, this would appear to be a rewarding area for linguistic analysis. H. Garrison Wilkes of Harvard has also found evidence in regional folklore for the role of teosinte in prehistoric times. According to Wilkes, the explorer C. Lumholz, who traveled extensively in Mexico at the turn of the century, noted that in the Nobogame area teosinte growing in or near fields of corn was said to be "good for the corn."

Actually there is good evidence that teosinte growing adjacent to fields of corn improves the vigor of the crop. The reason is that under primitive conditions of culture corn is grown in small areas, often relatively isolated from other corn populations, and seed is saved from each crop for successive plantings. Inbreeding is thereby intensified, and after several generations the vigor and yield of the population is reduced. If teosinte grows adjacent to such isolated plots or even in them, crosses between it and corn occur. The frequency of the crosses depends on a number of factors, such as spatial relations and relative times of pollination. When corn is pollinated by teosinte, the resulting kernels are usually indistinguishable from pure corn and so may be unwittingly chosen for planting. The resulting first-generation crosses show marked hybrid vigor: we grew corn-teosinte crosses in Mexico that were 12 to 15 feet high and had as many as 4,000 kernels on one plant.

Although the ears of such crosses are small and undesirable and hence are not chosen as seed for the succeeding crop, the crosses will have shed abundant pollen, which is disseminated widely because of the plants' tallness and vigor. The cross will therefore be the male parent in backcrosses to corn in the next generation. Some of the backcross progeny will be good corn showing marked hybrid vigor and increased yield, and second- and later-generation backcrosses could result in an even higher frequency of high-yielding corn plants. In this way the corn will be rejuvenated according to the same principles employed in developing modern high-yielding hybrid corns.

The rejuvenation of corn populations by crossing with teosinte has been demonstrated in a well-controlled experiment by Robert J. Lambert and Earl R. Leng of the University of Illinois at Urbana-Champaign. They crossed a well-adapted, high-yielding inbred line of corn with four varieties of teosinte and made successive backcrosses to the inbred corn parent. Yields in the second and third backcrosses were increased by as much as 100 percent. Such crossing is not infrequent in parts of Mexico where corn and teosinte grow together. In these populations corn and teosinte genes are exchanged regularly at a low frequency, thereby maintaining hybrid vigor.

With modern practices of corn-seed production hybrid vigor is controlled and maximized without repeated introduction of genes from teosinte. In the long run, however, there may be a need for a wider range of genetic diversity, for example to increase resistance to disease, and teosinte might well supply it. In 1973, 40 corn breeders and corn geneticists were polled on the advisability of banking seeds from a large array of teosinte varieties and from various corn-teosinte crosses of the second generation and later ones. Of those polled 35 regarded the seeds as a valuable source of genetic diversity likely to prove useful in the next decade or so, five considered the proposal interesting but of dubious scientific or plant-breeding value and none thought the seeds unworthy of preservation. A similar poll on the desirability of preserving wild teosinte populations would certainly have elicited an equal or more positive response, since a number of wild populations of teosinte in Mexico and Guatemala have already been destroyed. Others are in danger because teosinte is a highly favored food of both wild and domestic grazing animals.

Except for the rather haphazard teosinte-seed-bank program relatively little has been done. In part the reason is that teosintes are short-day plants not readily cultivated at temperate corn-belt latitudes. There are now a number of genetically adapted day-neutral teosintes, however, that may prove to be of significant value in more extensive breeding programs. It would therefore be wise to establish teosinte preserves at appropriate latitudes and altitudes in such protected areas as parks, archaeological sites, botanical gardens, agricultural colleges and commercial and government corn-breeding stations. The cost in money and effort would be insignificant and the benefits could be enormous.

Rabies

This ancient scourge of man is perpetuated by infected wild animals, notably skunks, bats and foxes. The classic Pasteur vaccine treatment can now be replaced by a shorter and less painful series of inoculations

by Martin M. Kaplan and Hilary Koprowski

On May 21, 1908, the Philadelphia *North American* carried the headline, "Drugged at His Own Request, Victim of Hydrophobia Dies Peacefully." The accompanying article was the dramatic story of a successful manufacturer who had taken medical courses and recognized the inevitable course of hydrophobia, or rabies, after the appearance of the laryngeal spasms that are the early symptoms of the disease. He described to the attending physician the probable entry of the virus (he had scratched an eczematous patch on his thigh after treating bite wounds on his bulldog), arranged his business affairs and asked for opiates for the convulsions he knew would come before his death. Aulus Cornelius Celsus in the first century A.D. would have prescribed treatment by ducking in a pool, and if Celsus' preventive measures had been followed, including immediate excision of the bitten tissue and cauterization of the wound with a hot iron, the victim's life might have been spared.

There are references to rabies earlier than those of Celsus. Aristotle in the fourth century B.C. erroneously stated that man, but not any other animal, was exempt from contracting rabies from the bite of a mad dog. Achilles called Hector a "rabid dog" in the Iliad (before 700 B.C.), and the following reference appears in the Mesopotamian "Laws of Eshnunna" (before 1800 B.C.): "If a dog is mad and the authorities have brought the fact to the knowledge of its owner; if he does not keep it in and it bites a man and causes his death, then the owner shall pay two-thirds of a mina (40 shekels) of silver. If it bites a slave and causes his death, he shall pay 15 shekels of silver."

In 1546 the Italian physician Girolamo Fracastoro portrayed vividly what it is like for a human being to be afflicted with rabies. "Its incubation [following a bite by a rabid animal] is so stealthy, slow and gradual that the infection is very rarely manifest before the 20th day, in most cases after the 30th, and in many cases not until four or six months have elapsed. There are cases recorded in which it became manifest a year after the bite." Once the disease takes hold "the patient can neither stand nor lie down; like a madman he flings himself hither and thither, tears his flesh with his hands, and feels intolerable thirst. This is the most distressing symptom, for he so shrinks from water and all liquids that he would rather die than drink or be brought near to water; it is then that they bite other persons, foam at the mouth, their eyes look twisted, and finally they are exhausted and painfully breathe their last."

The standard modern treatment entails a painful series of from 14 to 21 inoculations of vaccine following an exposure to an animal that is known to be or is suspected of being rabid. Every year more than a million people, including more than 30,000 in the U.S., go through this disagreeable procedure. Fortunately it is rapidly being supplanted in several countries (but not yet in the U.S., where the Food and Drug Administration is still considering it) by a different type of vaccine that Tadeusz J. Wiktor, we and other workers developed at the Wistar Institute of Anatomy and Biology. It entails only four to six inoculations and so far has been without side effects.

The human fear of rabies is well founded; few people are known to have recovered from the disease after the clinical symptoms have appeared. The great Austrian actor and playwright Ferdinand Raimund shot himself on August 25, 1836, a day after he was bitten on the hand by a dog that had also bitten other people and was later killed by a policeman. Raimund had had a morbid fear of rabies since his childhood and once even thought he had contracted the disease because a dog had licked a piece of bread he was eating.

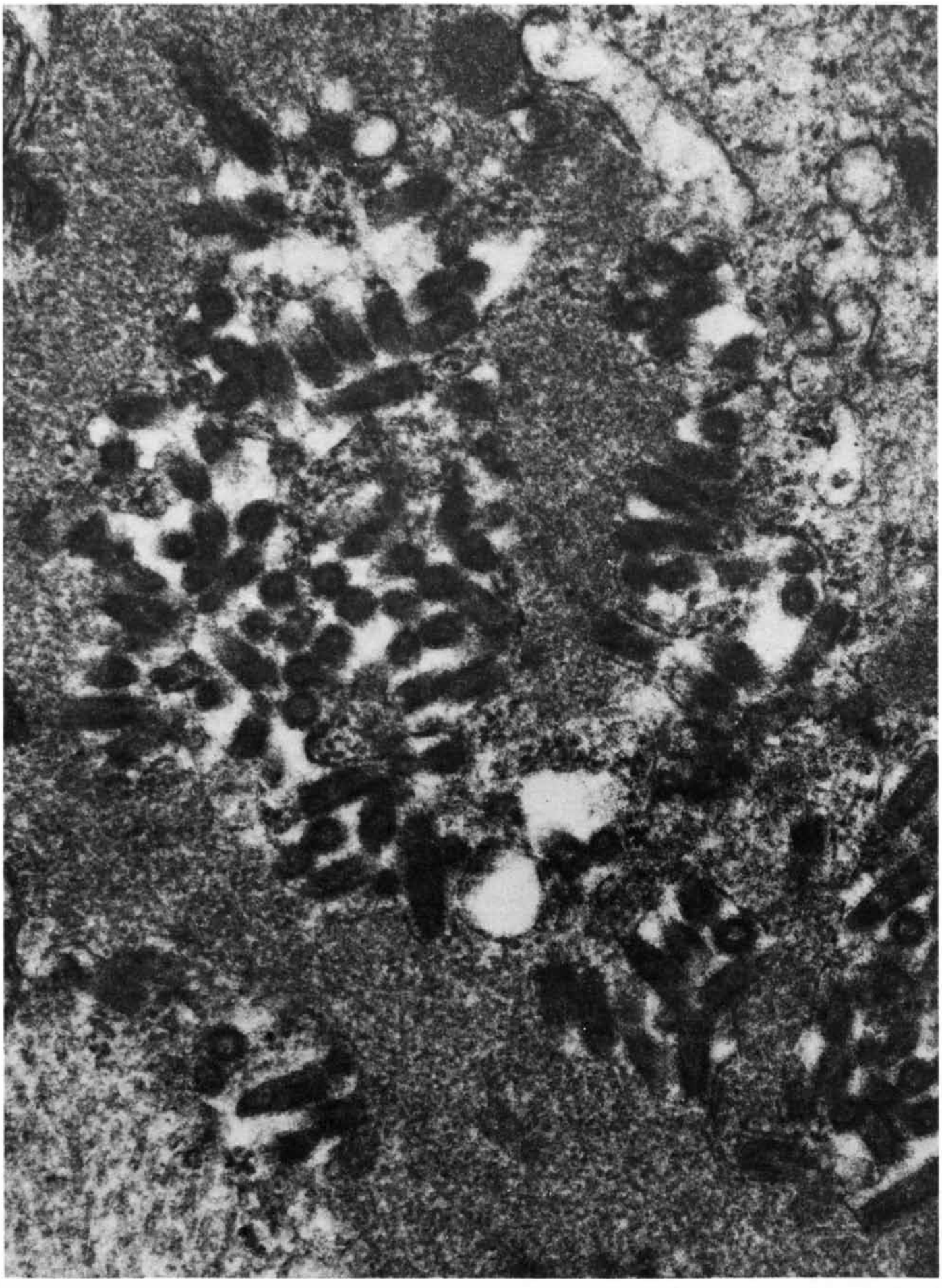
A successful defense against rabies infection requires that treatment start within a short time after a patient has been bitten. Notwithstanding the successes with vaccines and more than a century of laboratory work on rabies, such matters as when and how the rabies virus produces disease and how the vaccines achieve protection are still poorly understood.

Although rabies is found in a wide variety of mammals, in the past the main reservoir and transmitter of rabies in most countries has been dogs. The Greeks called the disease *lyssa* (frenzy) and the Romans used the word *rabere* (to rage), from which the modern term is derived. In antiquity the disease was associated with the appearance of Sirius (the Dog Star) in the "dog days" of summer, when dogs were considered to be particularly subject to madness.

For centuries it has been recognized that the causative agent of rabies is associated with the saliva of mad dogs, but it was not until early in the 19th century that proof of the relation was obtained experimentally (a considerable feat in the history of infectious disease). In 1804 the German worker G. Zinke demonstrated that the disease is infectious by inoculating normal dogs with saliva from rabid dogs, and in 1821 François Magendie and Gilbert Breschet infected dogs with human saliva, proving that the same agent was involved. Not until modern techniques utilizing the electron microscope and tissue culture began to be employed about 20 years ago was the viral agent itself characterized.

In the electron microscope the rabies virus resembles a bullet. Chemically the rabies virus consists of lipids, proteins and one species of ribonucleic acid (RNA). Taxonomically it belongs to the rhabdoviruses, a large group of viral agents that affect plants, insects and both warm-blooded and cold-blooded animals.

Until recently, because of a lack of reagents that could identify the specific antigenic determinants of the various components of rhabdoviruses, it was impossible to study the problems of the interrelation and evolution of these viruses. Now, however, the availability of monoclonal antibodies, particularly those that detect antigenic determinants on the nucleocapsids of viruses belonging to the rabies group, may allow the study of interrelations among various



RABIES VIRUSES appear at an enlargement of about 116,000 diameters in this electron micrograph of a slice of tissue from the brain of a rat infected with rabies. The characteristic bulletlike shape of the virus is evident in many of the dark forms. The dark circular forms are viruses in cross section. The micrograph was made by Carey Calaway of the Center for Disease Control, U.S. Public Health Service.

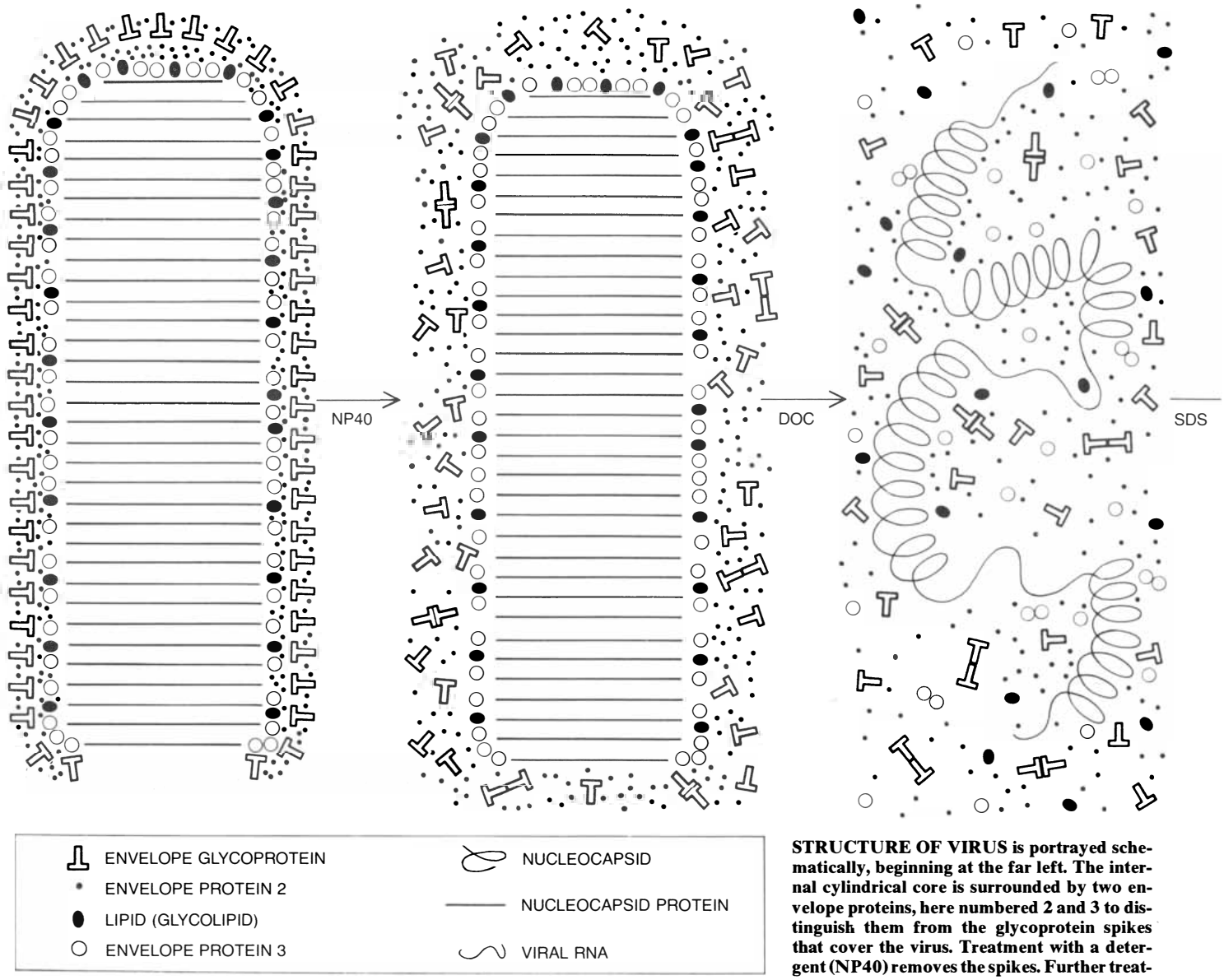
rhabdoviruses and provide a better insight into their evolution from a common ancestor.

The antigens represented by glycoproteins and nucleocapsid proteins until recently were thought to be similar, if not identical, regardless of the strain of rabies virus from which they could be isolated in relatively pure form. This assumption was based on a similar immunoreactivity of the antigens with antirabies serums obtained from animals immunized with rabies virus.

When rabies-virus preparations were studied in immunoassays with rabies-specific monoclonal antibodies produced by hybridomas in tissue culture, however, it became apparent that strains of rabies virus formerly considered as being identical or very closely related could be quite different antigenically. For example, two fixed laboratory-adapted virus strains derived from the rabies virus originally isolated by Louis Pasteur differed antigenically both from the Pasteur virus and from each other, and the much-attenuated Kelev strain reacted with only one hybridoma out of six hybridomas tested, all of which were producing monoclonal antibodies. In contrast to the marked antigenic differences observed among fixed laboratory-adapted strains of rabies, different "street" viruses (isolated from naturally infected animals in many parts of the world) show similar antigenic specificities when they are tested against a panel of monoclonal antibodies.

The results with monoclonal antibodies also reflect differences in the antigenic composition of the glycoproteins of rabies-virus strains. A. Flamand, Wiktor and one of us (Koprowski) have been studying the reactions of monoclonal antibodies against the nucleocapsid isolated from various strains of rabies virus. Contrary to expectations, we have discovered that nucleocapsid proteins can be classified into different antigenic groups. These results have also enabled us to establish the existence of antigenic determinants common to nucleocapsids of the rabies virus and what are known as rabies-associated viruses. Two such viruses, Makola and Lagos Bat, were isolated from human beings and bats in Nigeria. In contrast to the rabies virus, the Makola virus has been found to cause a nonlethal disease in human beings. Another African strain of rabies-associated virus, the Duvenhage strain, was isolated from a dying man in South Africa. Glycoproteins of this strain have been found to react only with the two hybridomas that also react with rabies-virus nucleocapsids.

Monoclonal antibodies can potentially be produced relatively cheaply and in large quantities. Since they react more specifically than antibodies in the animal serums employed for the diagnosis of rabies in infected cells, monoclonal antibodies should replace animal serums for such purposes. They may also eventually replace horse or mule antirabies serums, which cause unfavorable side reactions in from 15 to 20 percent of inoculated people. In addition they should replace antirabies gamma globulin of human origin, which is expensive and difficult to obtain, in the postexposure treatment of people bitten by rabid animals.

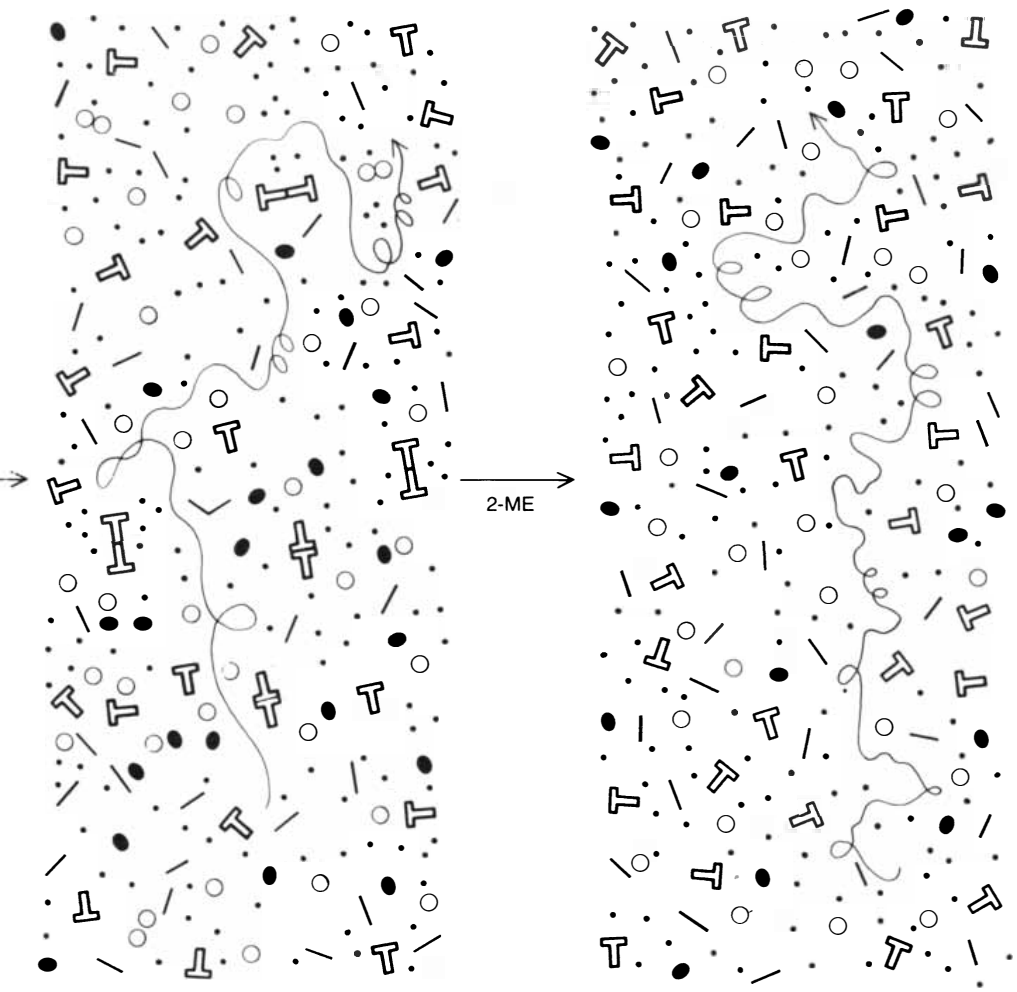


STRUCTURE OF VIRUS is portrayed schematically, beginning at the far left. The internal cylindrical core is surrounded by two envelope proteins, here numbered 2 and 3 to distinguish them from the glycoprotein spikes that cover the virus. Treatment with a detergent (NP40) removes the spikes. Further treat-

Although the events of viral replication in an infected cell are fairly well understood, it remains unclear what happens just before that, after the rabies virus has been deposited in the body tissue (usually in a wound caused by a bite). Recent work by Frederick A. Murphy of the Center for Disease Control in Atlanta found in newborn hamsters the first evidence of viral replication in muscle cells near the entry wound. This replication may provide a sufficient amplification of viral particles to create an infectious dose. On the other hand, the original inoculum of viral particles in saliva may be sufficient for the purpose.

The viral particles penetrate the nerve endings and progress toward the spinal cord in the axoplasm (the protoplasmic fluid that circulates in the axons of the neurons, or nerve cells, in the nerve trunks). The movement of the virus in the axons is passive, at a speed of about three millimeters per hour. The axons lead to the cell bodies of the neurons, which are situated in ganglia in the dorsal branch of the spinal nerves just before it enters the spinal cord. The infection spreads progressively up the spinal cord through the synapses of neurons with one another, also affecting the chain of other ganglia alongside the cord. The course of the virus through the spinal cord leads rapidly to the brain, where the virus multiplies. The infection then follows the nerve trunks outward to all parts of the body, including the salivary glands, which are the crucial site for the transmission of the virus to another animal. The virus multiplies further in the salivary glands.

The cycle may take weeks or months. It is not known how or where the virus remains in a latent and undetectable form during prolonged incubation periods. Nor is it known what balance of factors determines why rabies develops in certain individuals and not in others, but from the symptoms and the cause of death (respiratory failure) it is evident that the function of the neurons must be seriously impaired. Also not known is what arrests the disease in most naturally infected people, including those who have been vaccinated. Neutralizing antibodies, certain "killer" lymphocytes attacking rabies-infected cells and interferon may all be involved in halting infection. (Interferon is the protein manufactured by cells stimulated by a variety of infectious agents, including the rabies virus.) What is not clear is the interrelation of these factors.



ment with another detergent (DOC) disrupts the particle, making it possible to isolate a nucleocapsid consisting of from 30 to 35 coils of ribonucleoprotein. Further treatment with SDS and 2-ME digests the components of the virus so that single-strand ribonucleic acid (RNA) can be separated from the nucleocapsid protein. On analysis the glycoprotein is found to account for about 48 percent of the viral protein, the nucleocapsid protein for about 32 percent and the envelope proteins for 20 percent. A minor nonstructural protein is also present in the nucleocapsid. The analysis reveals that lipids such as cholesterol make up about 1 percent of the virus.

Rabies is found in most parts of the world. The major areas free of it are Australia and Antarctica, certain islands (including New Zealand, Hawaii, Britain and Ireland) and the Scandinavian countries except for the border region of Denmark. For centuries the principal vector of the disease was the dog, but for a variety of reasons (including ecological shifts and the vaccination of dogs) the dog has been superseded by other animals in many parts of the world in recent years, particularly since World War II.

All warm-blooded animals are susceptible to both natural and artificial infection with rabies virus, but their susceptibility varies widely from species to species. For example, skunks, opossums and fowl are relatively resistant, whereas foxes, cats and cattle are highly susceptible. Human beings and dogs occupy an intermediate position.

In the dog the incubation period is usually from three to six weeks, but it can be as short as 10 days and in rare instances as long as a year and perhaps even longer. The early signs are an alert, troubled air and a change of disposition followed by restlessness, snapping at imaginary flies and licking or gnawing the site of the bite. The animal is often fearful and hides.

Between the third and the fifth day after the appearance of these early signs the paralytic syndrome emerges in either the "furious" or the "dumb" form. In the furious form the dog is restless and snaps at moving objects, and toward the end the lower jaw hangs and saliva drools. Paralysis and convulsions follow. In the dumb form the owner may think the dog has something lodged in its throat. The dog is not irritable and rarely bites. It is lethargic and likely to hide.

In cats rabies follows a similar course. The signs of the disease in horses and cattle are irritation, restlessness and unusual aggressiveness. Rabies in such large animals is usually not transmitted, since they rarely bite people or other animals.

In wild animals the chief characteristic of rabies is the loss of fear of man and larger animals. The trait is particularly evident in foxes, jackals, skunks and bats. The signs usually resemble those in a rabid dog. An unprovoked biting attack by any wild animal, with the possible exception of a rat in a crowded urban setting, should give rise to a strong suspicion of rabies.

In many parts of the world the major problem with rabies is its transmis-

You can wait years for new developments to appear in standard texts, spend hundreds of hours extracting them from journals or you can subscribe to Scientific American Medicine.

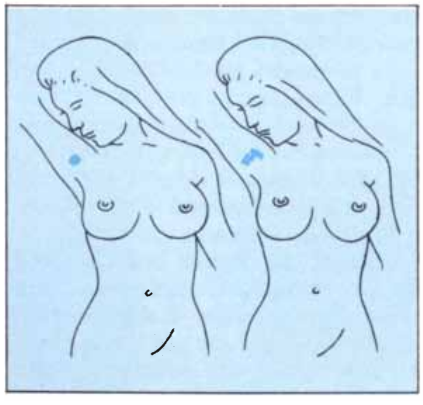
SCIENTIFIC AMERICAN *Medicine* is the clinician's answer to the diffuseness of the periodicals and the prepublication obsolescence of the textbooks.

Because its authors update SCIENTIFIC AMERICAN *Medicine* every month, it is always current. Because the new information that affects patient management appears in a single source, there can be no loss through diffusion.

This 2,000-page, innovative union of publishing and electronic technology is the work of leading scholar-practitioners from Harvard and Stanford. The editors are Edward Rubenstein, M.D., F.A.C.P., and Daniel D. Federman, M.D., F.A.C.P.

Each month as authors update their contributions, revisions are entered on the magnetic tape on which the text and index are stored. The tape drives high-speed phototypesetting equipment so that subscribers receive about eight new chapters and a

—Adjuvant chemotherapy for breast cancer utilizing CMF enhances survival among premenopausal women with axillary lymph node involvement.



—Posttransfusion hepatitis is usually caused by "non-A, non-B" viruses.

—Lidocaine fails to control ventricular arrhythmias in as many as 20 percent of acute M.I. patients. Procainamide, given by a bolus and continuous infusion method, is effective in most of these cases.

—Poorly ventilated recreational vehicles are a new, significant carbon monoxide poisoning hazard.

—Cerebral embolism, producing both transient ischemic attacks and permanent strokes, is a complication of mitral valve prolapse.

—Spectinomycin is effective for disseminated gonococcal disease caused by penicillinase-producing organisms.

—Adenine arabinoside has been licensed for treatment of herpes simplex encephalitis.

new index every four weeks; a bulletin highlights the new developments.

New material replaces old material in the living text, so that the information is there — up-to-date, at your fingertips.

A CME program of eight patient management problems offered over a 12-month period is available at no extra cost. Authorities from first-line institutions create the problems under the direction of Stanford's Office of Postgraduate Medical Education and the Educational Testing Service.

As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association provided it has been completed according to instructions.

This program is approved by the American Academy of Family Physicians for 32 Elective Hours of CME credit.

Trial Offer

Please send us the card and you will receive over a period of 60 days at no cost, the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, we will bill you $218, the price of the first year of service. This sum covers the cost of the initial two-volume text; renewal is at $165. Should you wish not to continue the subscription, just return the volumes to us.

The risk is all ours, so please mail the card today and let us take the hassle out of keeping up.

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

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine*. On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the price of the first year of service is $218, the current renewal price is $165, and the cost of the subscription is tax-deductible. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a complete refund.

☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*
☐ I shall also enroll in the CME Program
☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $218
☐ Please bill me

Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York

Name _____

MD Specialty _____

Address _____

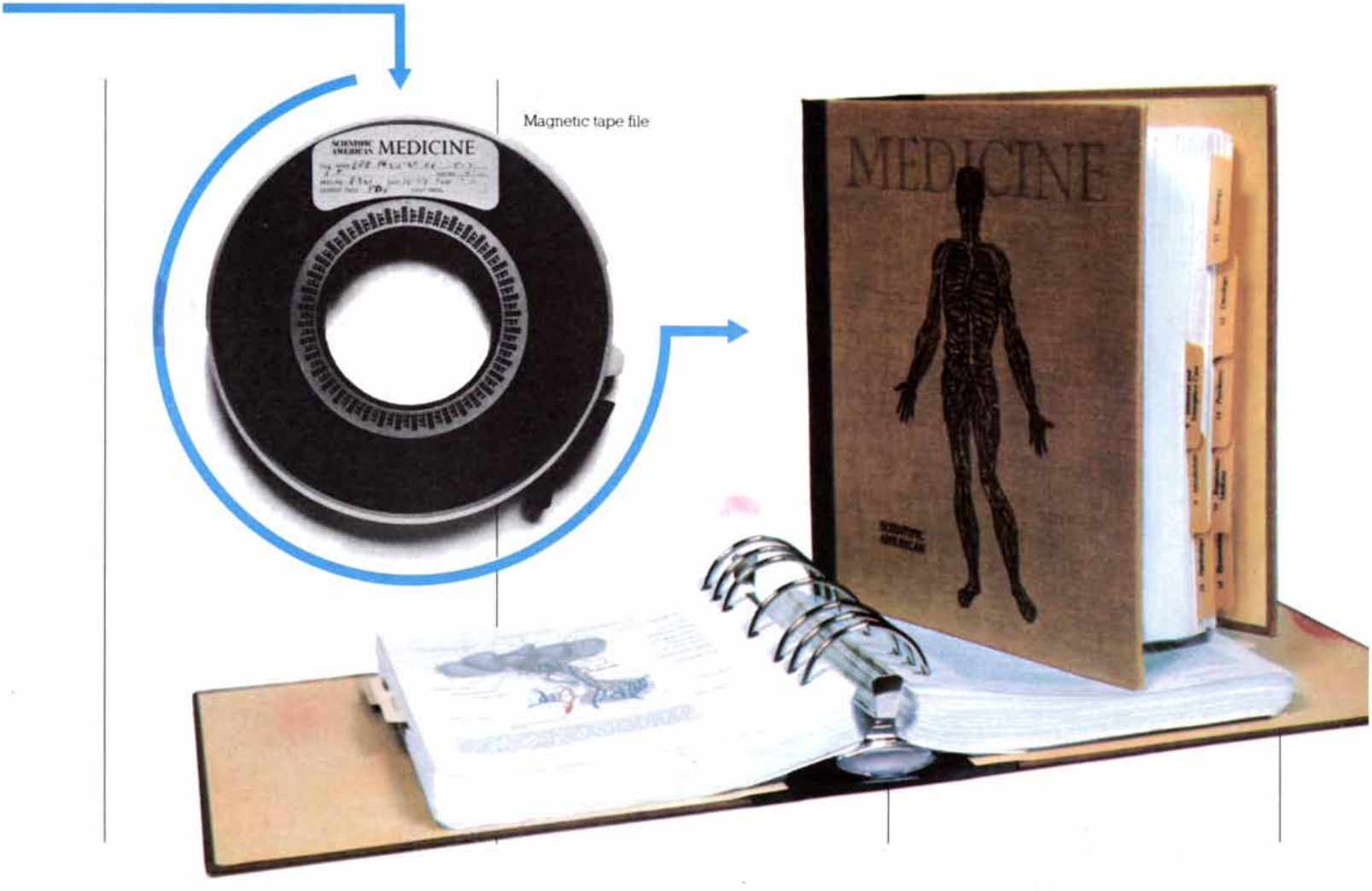
City _____

State _____ Zip _____

Signature _____

Subscribers outside of the U.S. and the possessions will be charged extra for shipping and handling; updates will be sent by surface routes unless air mail delivery is requested.

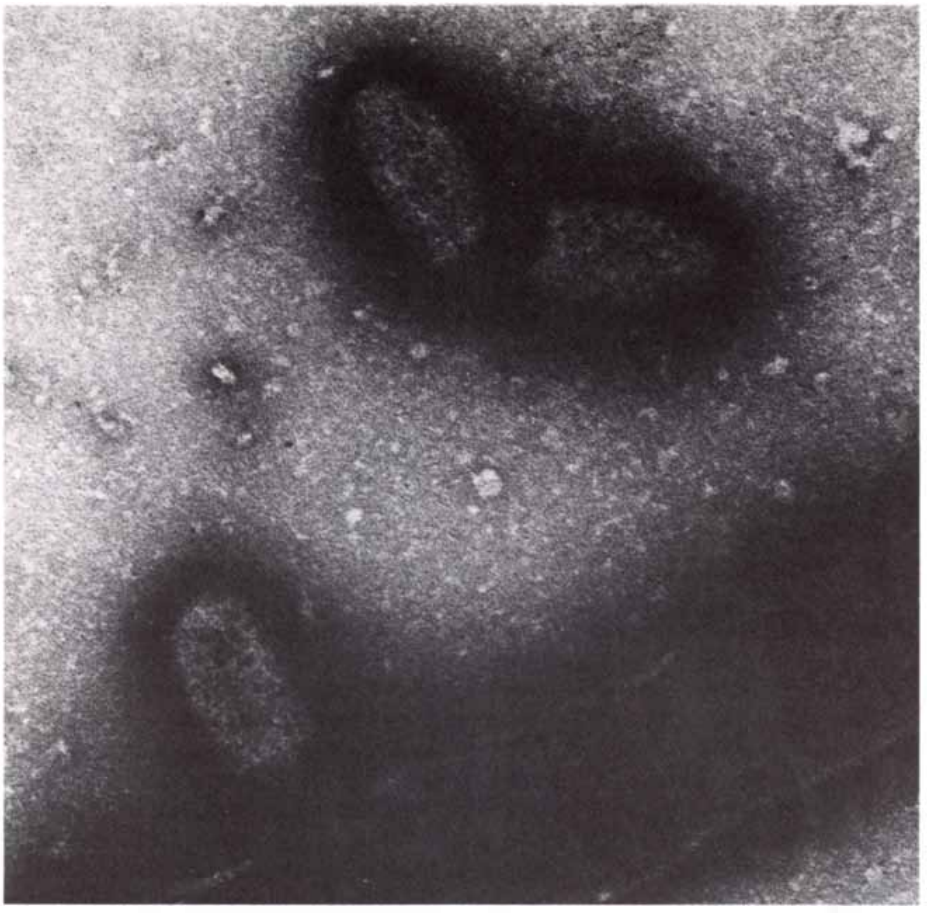
1H



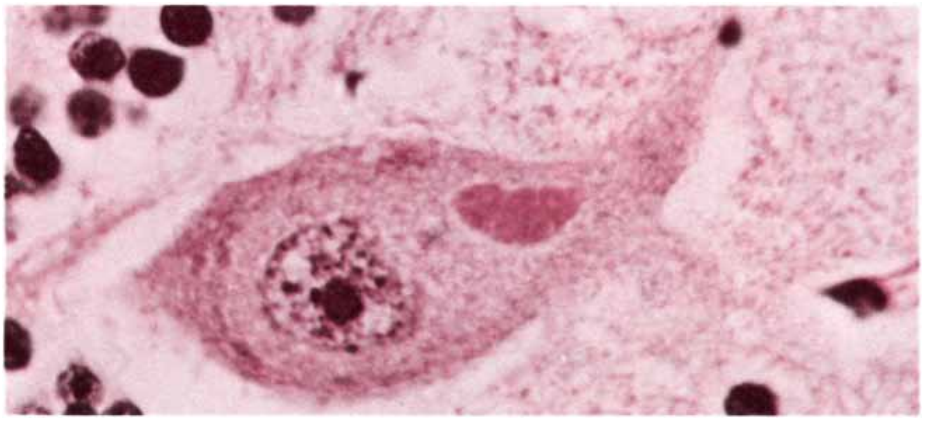
sion by wild animals. Spotted skunks (*Spilogale putorius*) were associated with human rabies in Lower California during the first quarter of the 19th century. In recent years skunks (mostly the striped skunk, *Mephitis mephitis*) have come to account for more than half of all rabies cases reported in the U.S. A rabid skunk may excrete the virus in its saliva for as long as 18 days before dying; it is also likely to be quite aggressive during that period. These characteristics make the skunk an effective perpetuator and disseminator of rabies both to other skunks and to other wild animals.

Fox rabies was first noted in Europe early in the 19th century, but it apparently did not attain major importance until the 20th century. The reasons for the increased incidence are not known. Today the red and the gray fox are the primary reservoir of rabies in continental Europe, the entire northern circumpolar region including Greenland and parts of Asia and North America. Fox rabies is particularly serious in Europe, where the disease has spread steadily at a rate of about 30 kilometers per year from east to west since World War II. Coming by way of Poland and Germany, it has almost reached the English Channel in northern France and has penetrated most of the eastern and southern parts of France and the rest of continental Europe.

About 50 percent of the foxes in a heavily infected area die from rabies. In West Germany fox rabies has been observed to recur in cycles of three years, so that even a death rate of 50 percent is insufficient to make the disease disappear spontaneously. A population density of one fox per square kilometer is enough to maintain an infection among the resident foxes. Experience in Denmark has shown that the disease disappears only when the fox population is reduced to one per four or five square kilometers.



REPRODUCTION OF VIRUS by budding from an infected cell is evident in this electron micrograph. The cell membrane is the dark body running diagonally across the bottom of the micrograph; one rabies-virus particle is just budding from the membrane and two are above it.



NEGRI BODY provides a means of diagnosing rabies. The fish-shaped structure in the center of this micrograph is a brain cell from a bovine victim of rabies. The dark elliptical object in the upper right section of the cell is the Negri body, named for the Italian physician Adelchi Negri, who discovered this characteristic structure in 1903. The structure appears in the cytoplasm of an infected cell late in the course of the disease. The Negri body is not always present and can be missed or misdiagnosed, so that it is now only one of several diagnostic signs.

Bram Stoker's classical novel *Dracula* deals with ancient folktales of vampires and vampirism. The story is set in Transylvania (Romania), and it centers on human vampirism, the sucking of human blood by an "undead" person who leaves his grave at night. This superstition spread widely in Europe in the 18th century, particularly in Hungary and Serbia. In Meduenga, a village near Belgrade, in 1732 and again in 1737 a medical commission including three regimental surgeons decreed the disinterment of many bodies, which received the prescribed treatment for vampires. Driving a stake of ash wood through the chests of supposed vampires, of whom suicides were the chief suspects, became such a pastime in Europe that in 1823 a law was enacted in England specifically prohibiting the mutilation of the body of a suicide by driving a stake through it.

In European witchcraft tales bats and vampires were often interchangeable. (Although Stoker never names the animal Dracula resembled in his nightly wanderings, the description of the specter suggests a bat.) Thus Pietro Martire d'Anghiera, the first bishop in the New World, perhaps did not consider it strange to record early in the 16th century: "In several places bats not much smaller than turtle doves used to fly at [men] in the early evening with brutal fury and with their venomous bites brought those injured to madness." Similar observations of bat bites and deaths in man and animals were recorded by early Spanish colonists in the Mexican peninsula of Yucatán and in what is now Panama.

Notwithstanding the long folk tradition of bat bites as a source of rabies, the

"I wanted to design a camera that takes photographs no other camera can."

Such was the challenge that Maitani, the celebrated Olympus camera designer, set for himself.

If he had listened to conventional wisdom, he would have folded up his drawing board and gone home promptly at five.

Instead, Maitani persisted. And created the Olympus OM-2.

The fully automatic Olympus OM-2 is the only camera that measures the light *during* exposure.

Because two sensors read the light that reflects off the film surface the instant the shutter opens.

So the shutter may stay open for up to two minutes as the camera adjusts for changes in light and automatically closes at the precise moment perfect exposure has been achieved. No other camera can do it.

Maitani

The OM-2 also embodies all that Olympus SLRs are justly famous for. It's a light, compact, rugged camera with a bright viewfinder and well-proportioned controls.

And that's just the beginning. You can snap on a five-frame-per-second motor drive, or a winder, or a 1,000mm lens, or any of 300 plus components from the stellar Olympus OM system.

The Olympus OM-2 is what happens when a brilliant designer sets his sights on the impossible.

To create what's possibly the world's ultimate camera.

For information write Olympus, Woodbury, N.Y. 11797.

OLYMPUS® OM-2

link was not proved until some 50 years ago. In Trinidad Joseph L. Pawan observed that natives sleeping outside their dwellings because of the heat were often attacked and bitten during the night by dark, fast-moving creatures Pawan recognized as bats. The human victims, who never felt the bites, became sick six weeks later and died of rabies. About 10 years earlier the relation of bites by rabid bats to paralytic rabies in cattle had been established in southern Brazil.

Vampire-bat rabies kills hundreds of thousands of cattle annually in Latin America, afflicting all countries except Chile. Although most vampire bats that have been infected experimentally become paralyzed and die within a few weeks, some of them remain without visible symptoms and shed virus in their saliva for months and even years. The bats live in colonies and have been observed to fight not only among themselves but also with fruit-eating and insectivorous bats, to which they can transmit rabies. The first bat-transmitted rabies case north of the Mexican border may have been an episode in the 1940's involving a giraffe that died of rabies in the San Diego Zoo. The episode was associated with a large incursion of insectivorous bats. The presence of rabies-infected bats in the U.S. was not believed, however, until 1953, when rabies virus was isolated from the brain tissue of an insectivorous bat that in broad daylight had bitten a boy in Tampa, Fla.

Many attacks on human beings by rabid insectivorous bats have since been recorded. They involve more than 50 species of bats in almost all the states of the U.S. and in many Latin-American countries. It is not clear whether the thousands of naturally infected insectivorous bats found in the U.S. over the past 25 years represent a spread of infection or merely reflect the discovery of an already widespread incidence of the disease among bats.

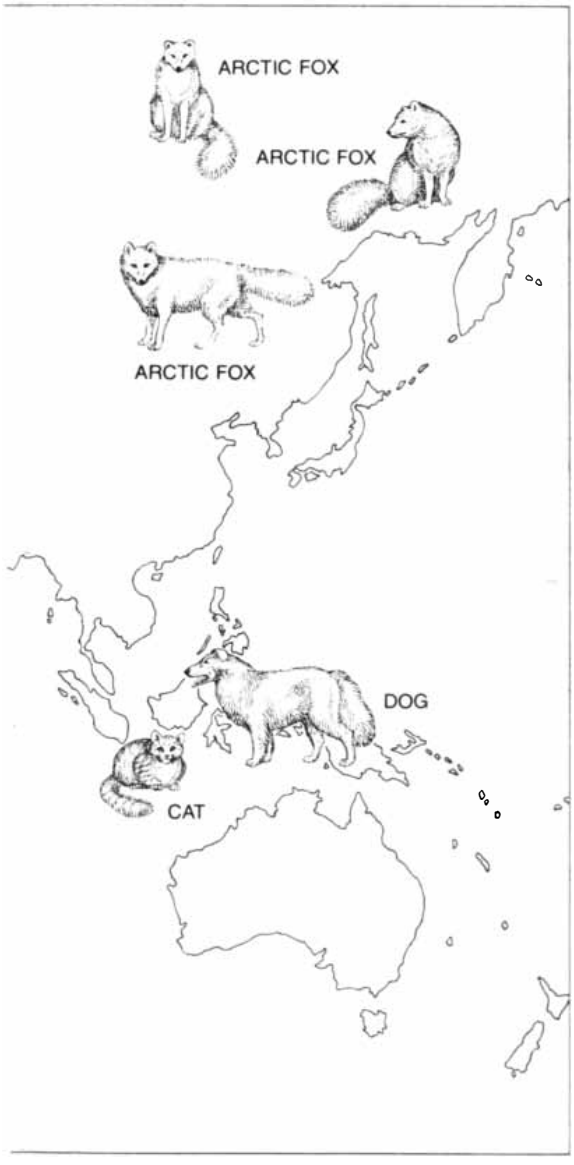
There is some evidence that bat rabies can be transmitted without a bite. In Texas a state entomologist who spent a great deal of time in the Frio Cave in Uvalde County contracted rabies in 1957 without having been bitten; the disease was probably transmitted from infected bats by inhalation. Coyotes, foxes and opossums in screened animalproof cages exposed to the air in the cave all died from rabies. Eventually rabies virus was isolated from samples of the air in the cave.

Before the advent of modern medical concepts the treatment of rabies in human beings included more than the usual quota of bizarre practices. Ancient Chinese physicians regarded musk and cinnabar as infallible. Ducking stools had their day, and so did treatments with wormwood or the ashes of seahorses. In 1806 the New York legislature awarded $1,000 to John M. Crous for discovering a "cure" consisting of a

PRINCIPAL ANIMAL VECTORS of rabies are indicated for the major regions in which the disease appears. In the U.S. skunks, mainly the striped skunk (*Mephitis mephitis*), account for more than half of all the rabies cases diagnosed in animals; some 50 species of insectivorous bats account for almost 15 percent. In places where dogs are the principal source of exposure for human beings, cats are usually

tablet made from the pulverized jawbone of a dog, the dried tongue of a newly foaled colt and corroded copper from an English penny minted during the reign of George I. Among Western frontiersmen it was not unusual to treat animal bites with "madstones," which were the gallstones of rare white deer or white cattle.

Rabies was the first human disease for which a systematically developed vaccine was successful. Pasteur and his colleagues, starting in 1881, found that the "poison" of rabies (the virus) could be obtained in relatively pure form from the brain and spinal cord of an infected animal. A major advance was the discovery that by the repeated passage of the virus in rabbits street virus, which has an extended and variable incubation period, could be converted into a fixed form with a short and constant incubation period of from five to seven days. This kind of virus was eminently suitable for the preparation of a vaccine.



second. Australia, the British Isles and much of Scandinavia are free of rabies; animals being imported are put in a lengthy quarantine.

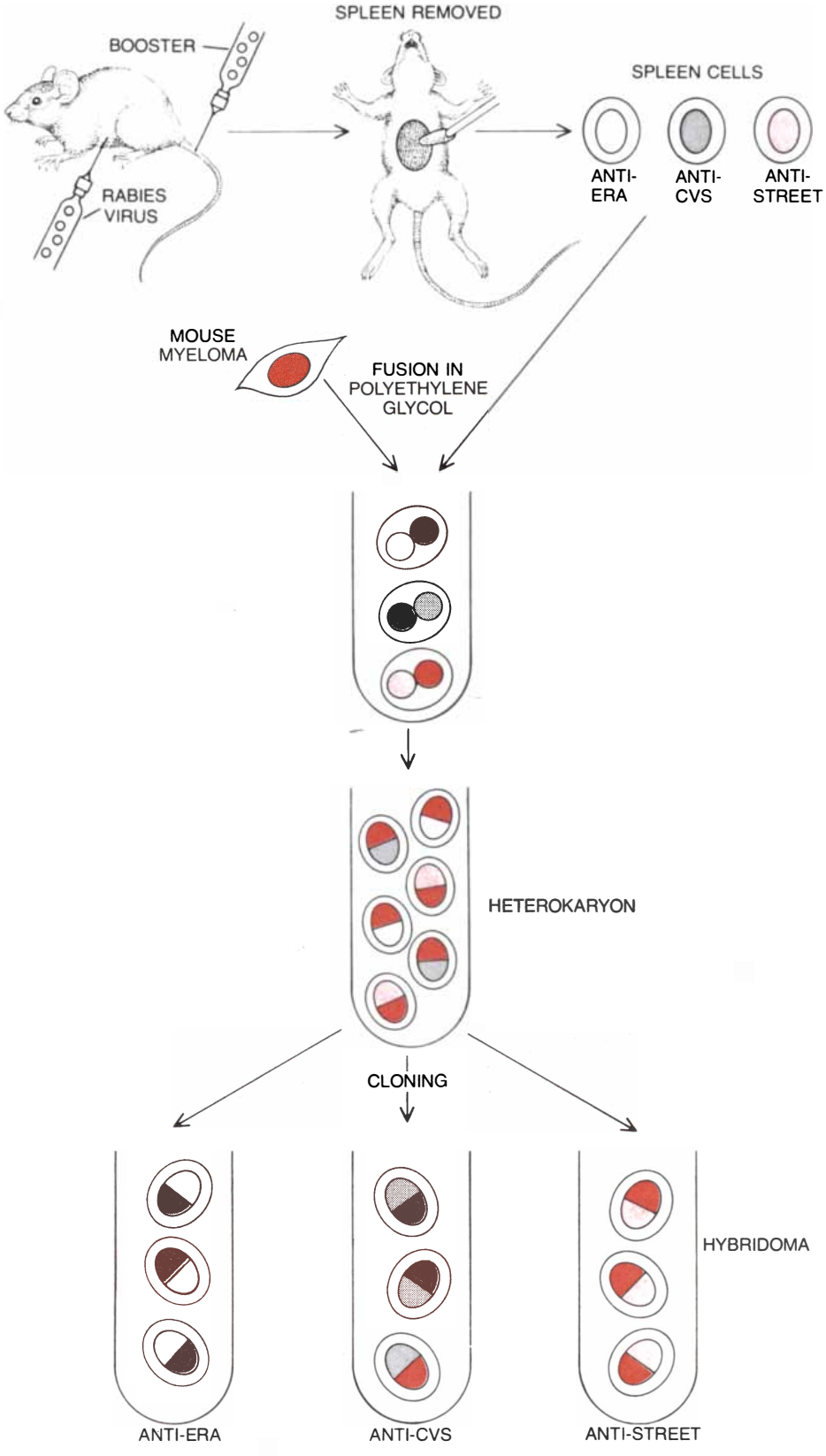
Pasteur's success with a vaccine in dogs encouraged him in 1885 to try it on one Joseph Meister, who later became a janitor at the Pasteur Institute. Although Meister had been severely bitten by a rabid animal, he survived. The test paved the way for wider trials in human subjects. On the whole they were successful.

Pasteur's success began with an experiment being done by his colleague Émile Roux, who had put in an incubator flasks containing the suspended spinal cords of rabbits that had died of experimentally induced rabies. Roux was trying to determine how long the virus would persist at a temperature of 37 degrees Celsius (98.6 degrees Fahrenheit, or body temperature). Pasteur modified the technique by adding potassium hydroxide as a drying agent and holding the flasks at room temperature.

His stroke of genius was to prepare a series of from 21 to 28 inoculations consisting of spinal-cord suspensions starting at the 14th day of drying, by which time the virus was innocuous, and progressing through suspensions dried for from 13 to three days. The last several suspensions were capable of causing rabies in rabbits, but by that time the earlier inoculations had conferred enough immunity to provide protection against the disease.

Since then a number of modifications of Pasteur's method have been introduced. Most of them sought to improve the keeping quality of the vaccine. Nevertheless, the principle of vaccine based on nerve-tissue suspensions has been retained and is still employed in most developing countries because it is easy to produce and inexpensive.

The Pasteur procedure was not without complications, the worst of which is that the fixed virus in the vaccine remains capable of causing rabies and sometimes does so when it is improperly prepared. Indeed, in 1886 Pasteur himself was accused of causing with his vaccine the death of a 10-year-old boy in Paris. Pasteur's opponent was Georges Clemenceau, who later twice became prime minister of France. In the 1860's Clemenceau briefly practiced medicine in New York and taught French in a girls' school in Stamford, Conn. He was a persistent advocate of spontaneous generation (opposed by Pasteur) and stuck to this viewpoint until he died in 1929. At the hearings Pasteur was supported by Paul Brouardel, the director of the laboratory, and by Roux (in spite of Roux's resentment over Pasteur's having copied his drying procedure). Brouardel and Roux recognized that a defeat for Pasteur could set back French medical science for decades and might expose the laboratory to lawsuits. Pasteur was vindicated, and his vaccinal procedure came into wide use through



MONOCLONAL-ANTIBODY TECHNIQUE provides a highly specific means of differentiating one strain of rabies virus from another. Mice are immunized with rabies virus or with isolated components of the virus, such as a glycoprotein or a nucleocapsid. One injection of such an antigen is given intraperitoneally and is followed three or four weeks later with a booster given intravenously. After three or four days the spleen is removed and a suspension of mouse lymphocytes is fused in the presence of polyethylene glycol with mouse myeloma cells grown in culture and deficient in the enzyme hypoxanthine guanine ribosyltransferase. The fused cells, called heterokaryons, are placed in a culture medium that prevents the growth of myeloma cells. Since mouse lymphocytes too cannot grow into a permanent line in culture, the colonies that appear in the culture dishes consist solely of hybrid cells combining traits of lymphocytes and myeloma cells. These cultures, called hybridomas, are cloned for the production of daughter cultures originating from one hybridoma cell. The clones are examined for the presence of specific antibodies against such rabies-virus strains as "ERA," "CVS" and "street." A hybridoma that makes a particular rabies antibody can be maintained indefinitely in culture.

Pasteur institutes established in other countries.

Notwithstanding a complete and intensive course of Pasteur treatment, some rabies patients died. Therefore in 1889 Victor Babes and M. Lepp developed a new approach in treating human beings who had been exposed to rabies. They immunized animals with rabies and obtained from them blood serum containing antibodies against the virus. They administered the serum in conjunction with the Pasteur vaccine. At the time the treatment was not widely accepted because of equivocal results.

In the 1940's Karl Habel of the National Institutes of Health reawakened interest in the serum technique, and as a result of additional serum studies by one of us (Koprowski) a trial in human beings of antirabies serum prepared in horses was recommended in 1950 by the committee on rabies of the World Health Organization (WHO). The opportunity came in 1954 when a rabid wolf bit 29 people and several domestic animals in a village in Iran. Under the plan of the experiment one group of human victims received the normal course of vaccine and the other received serum in addition to the vaccine. Of 13 who received the dual treatment only one died, whereas three of five people with similar wounds died after the standard treatment with potent vaccine alone. Eleven people more mildly exposed survived in spite of the treatment.

Today antirabies gamma globulin (the concentrated fraction of the blood serum containing antibodies) is widely employed in conjunction with vaccine. Yet because of the poor potency of many current vaccines the short-term immunity provided by serum can interfere with the lasting antibody-producing effect of the vaccine. It was therefore necessary to find a more potent vaccine to overcome this interference. Other reasons for trying to develop a more potent vaccine were to reduce the schedule of two or three weeks of daily inoculations required for the current vaccines and to eliminate their side effects, which were sometimes serious enough to cause paralysis and death.

In the 1960's Wiktor, Stanley A. Plotkin and we therefore undertook intensive laboratory studies to develop a vaccine prepared from virus grown in normal human embryonic cells. After almost 10 years of research a highly potent vaccine was produced that protected monkeys against challenge by virulent rabies virus, even when only one inoculation of the vaccine was given within a few hours after the challenge.

The new vaccine was subjected to a stringent test in human beings in 1975 and 1976 at the Pasteur Institute in Teheran under the direction of M. Bahmanyar, who had participated in the earlier wolf-bite tests. The recent tests

At last! Statistical Analysis (and a lot more) made easy by the low-cost TRS-80™ computer!

Store, retrieve and analyze volumes of data. Compute averages. Catalog collections. Use graphics to dramatize or prove your point. Balance your checkbook. It's all possible, and that's only the beginning!

TRS-80 Model I
Level II, 16K System
$849*

There must be a better way!
Statistical analysis, the organization and interpretation of numerical data, used to mean rolls and rolls of adding machine tape, hours of calculations, waiting for computer time, or an unchecked error at the start birthing a monstrous error in the end. But that's all over!

TRS-80 speeds and simplifies!
With Radio Shack's best-selling TRS-80 microcomputer, people having a minimum knowledge of data analysis techniques, and no computer programming experience, can now perform the sophisticated computations needed for accurate and detailed conclusions. Random samples, histograms, T-tests, linear regression and more are simplified. The secret? TRS-80's easy-to-learn BASIC computer language and Radio Shack's prewritten software (programs on cassette tapes) for statistical analysis.

If you can play a cassette tape, you can use TRS-80 computer.
Over 50 other software packages are available, including Real Estate, Word Processing, Payroll, Accounting, Personal Finance, even fast-paced computer games like Flying Saucers and Micro Chess. All software is highly interactive, meaning the computer asks you to "fill in the blanks" with data via the keyboard, then presents the answer on the screen — or prints it, if you add a TRS-80 printer.

The details are nearby!
We've covered only a small sample of what TRS-80 can do — and *what it's doing now* for well over 100,000 people from all walks of life. Visit one of our 7300 stores or 50 new Computer Centers in major cities for a "hands on" demonstration!

New TRS-80 Model II
Designed especially for business and professional applications. Built-in 8" disk — expandable to 4 disks, 2 megabytes memory. Priced far below competition. Complete Model II systems from $3450.*

FREE TRS-80 CATALOG!
24 fact-filled pages. Write: Dept. CMA-416, 1300 One Tandy Center, Ft. Worth, Texas 76102

*Retail prices may vary at individual stores and dealers.

Designed and built in the USA by

Radio Shack®
The biggest name in little computers™
A DIVISION OF TANDY CORPORATION • FORT WORTH, TEXAS 76102

Imagination and Research on the Frontiers of Science

Books of Special Interest to Readers of *Scientific American*

GRAVITATIONAL CURVATURE

An Introduction to Einstein's Theory

Theodore Frankel, University of California, San Diego

This fresh introduction to gravitational curvature captures the geometric essence of the subject in a modern mathematical setting. Replacing sometimes tedious analytical computations with geometric arguments, the book presents new derivations for Einstein's equations of gravitation along with other results. The derivations presented are at once clearly motivated and mathematically clean and simple. *Gravitational Curvature* includes new and attractive derivations for the deflection of light, Schwarzschild's exterior and interior solutions, and the Oppenheimer-Volkoff equations important in the study of stellar structure.

The book stresses the global aspects of cosmology and is suitable for independent study. The reader should have a background in Riemannian geometry and basic physics, or some familiarity with relativity theory. Background chapters, with derivations, are included on special relativity, continuum mechanics, and electromagnetism.

1979, 172 pages, 29 illustrations, cloth **$20.50**, paper **$9.95**

CANCER

Science and Society

John Cairns, Mill Hill Laboratory, London

"The best book that I have read explaining cancer to the intelligent layman." —Bruce N. Ames, University of California, Berkeley

A common cause of death in the industrialized Western world, cancer is one of the most feared of all diseases. Dispelling the misinformation surrounding cancer, John Cairns reviews its nature, its causes, its place among other diseases, the current state of research, and the prospects for its final conquest. He convincingly argues that cancer is largely due to controllable variables in the human environment and should therefore be preventable.

Written clearly but never simplistically, *Cancer: Science and Society* is for the general public as well as for physicians, medical students, and students of biology. References to the technical literature on the subject are given throughout the book for those who wish to read in more depth.

1978, 199 pages, 52 illustrations, paper **$8.25**

FRACTALS

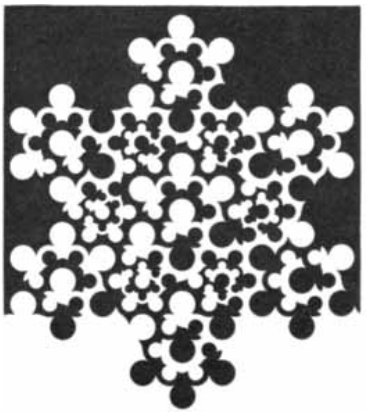
Form, Chance, and Dimension

Benoit B. Mandelbrot, IBM Fellow, International Business Machines Corporation, Thomas J. Watson Research Center

"The new and fascinating world of patterns presented here is all the justification this book needs. Read it, or better yet, look at the pictures. It is inconceivable that it would fail to give some food for thought." —*American Scientist*

Fractals: Form, Chance, and Dimension is an intriguing investigation that claims the geometry of fractals is necessary to describe geometric shapes in nature. Mandelbrot examines fractal applications for describing the surface of the earth, the clusterings of stars, and the shapes of clouds, snowflakes, and other forms appearing in nature.

1977, 365 pages, 68 illustrations, cloth **$20.00**



W. H. Freeman and Company
660 Market Street, San Francisco, CA 94104

Will there ever be another TV series as worthwhile as Sesame Street?

Yes, thank goodness. It starts Monday, January 14.

Joan Ganz Cooney and her Children's Television Workshop saw a way 10 years ago to teach kids, using the medium kids love best. The happy result was *Sesame Street*.

Does anyone have to be told *Sesame Street* went on to become one of the most successful TV shows ever?

Now Joan Cooney and the Workshop are ready with a new series every bit as significant as *Sesame Street*. It's called *3-2-1 Contact*. It will be on PBS stations Mondays through Fridays.

3-2-1 Contact is about science. If you just said *ugh*, maybe it's because you grew up without a show like this, which is a pity. And a national problem.

Thousands of bright and inquisitive youngsters start their school years with a natural interest in the sciences. But they get turned off before they ever reach junior high school. Thousands of scientists and engineers are lost to America in the process. *3-2-1 Contact* will keep curiosity alive in 8-to 12-year olds by bringing the same excitement to science and technology that *Sesame Street* brought to the child's first recognition of letters and numbers.

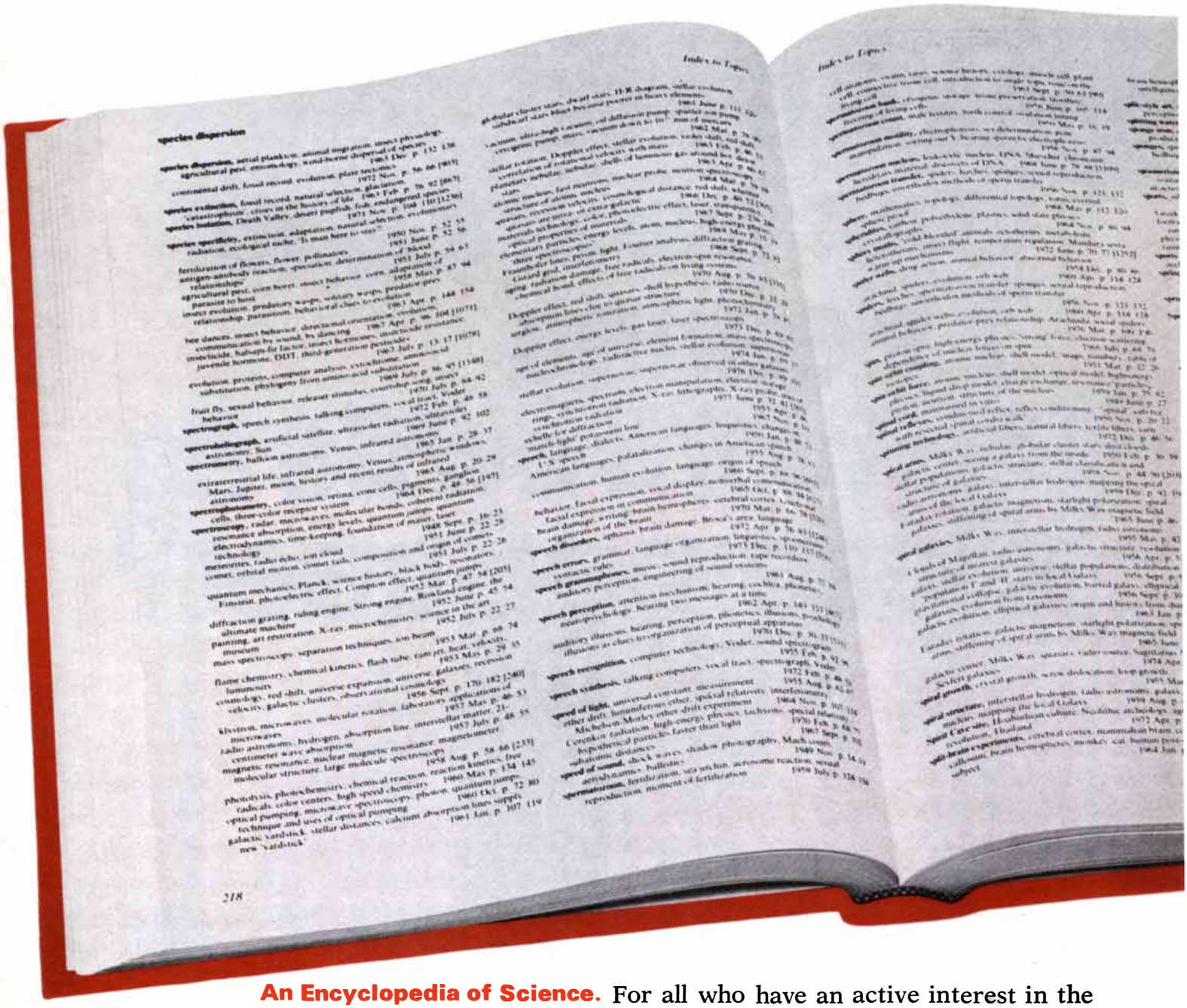
The National Science Foundation, The U.S. Department of Education, and the Corporation for Public Broadcasting all contributed sizable sums to help put *3-2-1 Contact* on the air.

So did we.

Our motives are simple. We invest more than a million-and-a-half dollars a day in scientific research and development to keep our Otis elevators, Pratt & Whitney Aircraft jet engines, Carrier air conditioners, Sikorsky helicopters and other high-technology products the best in their fields—and to create new and still better products. The United Technologies $2 million grant for this show is part of our investment in the future. We're pleased to make it.

And we're happy to invite you and your 8-to 12-year old friends to watch *3-2-1 Contact*. See your local TV listings for the broadcast schedule.

UNITED TECHNOLOGIES
Hartford, CT 06101



An Encyclopedia of Science. For all who have an active interest in the work of science, the new Index transforms the collected issues of SCIENTIFIC AMERICAN into an encyclopedia of the substance and history of modern science. It is actually eight different indexes in one:

Index to Topics—a rotated key-word index of the subject matter of all articles and Science and the Citizen features (see above).

Index to Proper Names—the names of all persons mentioned in the articles and Science and the Citizen and of places and institutions featured in a primary role.

Listing of Tables of Contents—a chronological listing of the complete contents of all 362 issues that provides ready identification, by titles and authors, of the articles cited in the two foregoing indexes.

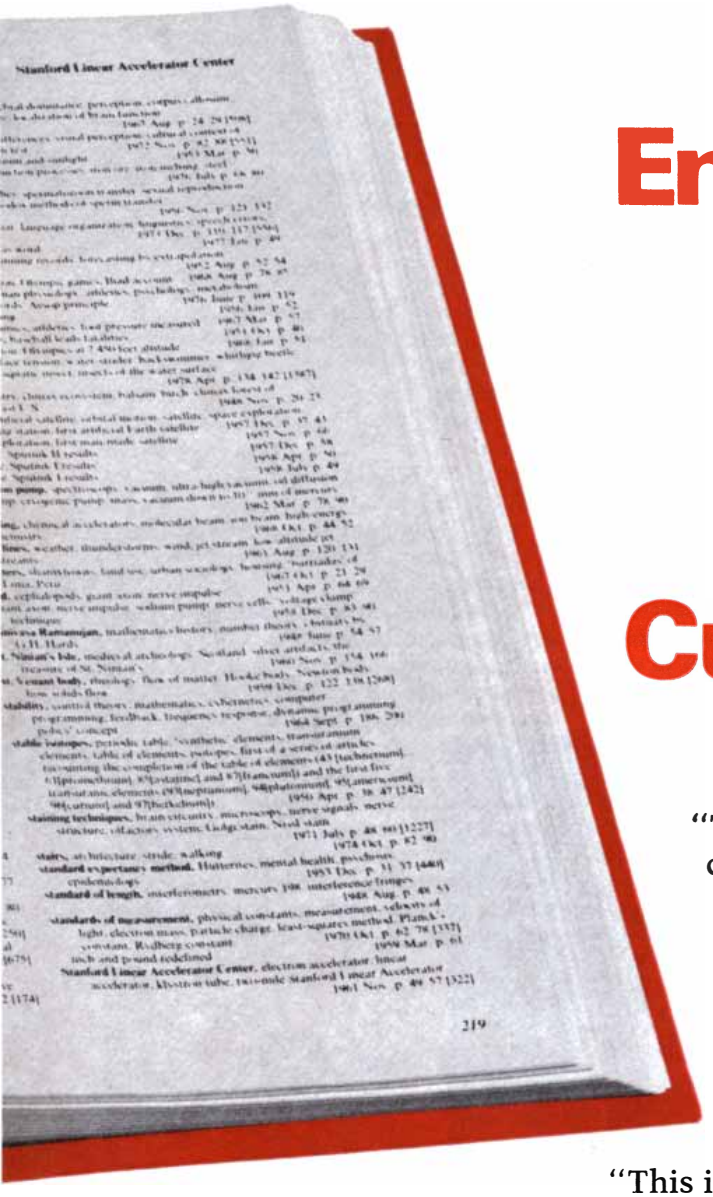
Index to Authors—an alphabetical listing of authors of all articles.

Index to Titles of Articles—an alphabetical listing of all articles by the first key word in each title, and by the other key words in the title.

Index to Book Reviews—a listing of the longer book reviews in three parts: authors of books, titles and reviewers.

Index to Mathematical Games—a listing of the puzzles, games and diversions presented in this department since its inception in January 1957, under the editorship of Martin Gardner.

Index to The Amateur Scientist—a listing of all the projects, experiments and demonstrations presented in this department from April 1952 through February 1976 under the editorship of the late C. L. Stong and from July 1977 through June 1978 under the editorship of Jearl Walker.



Enjoy direct entry to 30 years of SCIENTIFIC AMERICAN in the new Cumulative Index 1948-1978

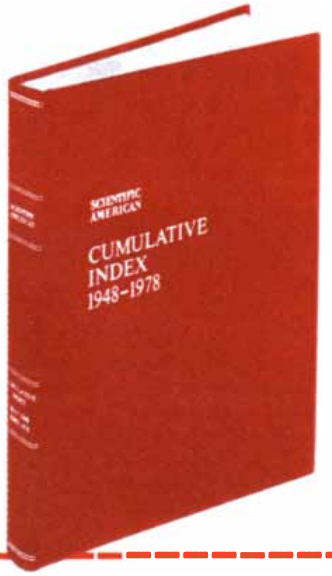
"This kind of meticulous cross-referencing inspires confidence. Only a very persistent gremlin could lose something from SCIENTIFIC AMERICAN down a black hole . . ." —**The New Library World,** London, June 1979

"You people sure know how to organize an index!" —**Donn P. Brazier, Director,** St. Louis Museum of Science and Natural History

"This index is an invaluable complement to thirty years of SCIENTIFIC AMERICAN. Aside from the usual author and title indexes, there is a topic index—a kind of permuted key-word index—which takes the reader-researcher below the level of title index." —**New Technical Books,** The New York Public Library, July 1979

SCIENTIFIC AMERICAN

415 Madison Avenue, Dept. CI, New York, N.Y. 10017



Yes, send me _____ copies of the new SCIENTIFIC AMERICAN Cumulative Index to the 362 issues from May 1948 to June 1978, at the price of $45 each plus $2.50 for shipping. If I enclose payment, you will absorb the shipping charge.

☐ **My check for $_____ is enclosed, saving me $2.50 shipping charges.**

☐ **Bill me for $45 each copy, plus shipping.**

Name _____

Address _____

City _____

State _____ Zip _____

(Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York.)

1/80

included two episodes of wolf bites involving nine people and six episodes of dog bites involving 36 people. The biting animals were shown to have rabies. The exposed individuals received serum and a total of only six doses of the new vaccine rather than the 21 doses that were standard for the old one. The treatment of some of the victims did not start until more than three days after exposure (for one it was 14 days), a delay that past experience has shown will almost certainly result in death from rabies in some cases. None of the 45 people died of rabies. The vaccine is now employed routinely in many countries.

A critically important step in the prevention of rabies after exposure is to treat the wound itself as soon as possible with virus-destroying solutions such as soap and water, alcohol, iodine and quaternary-ammonium disinfectants. If antirabies serum is available, it can be infiltrated around the wound. Suturing the wound should be delayed if possible.

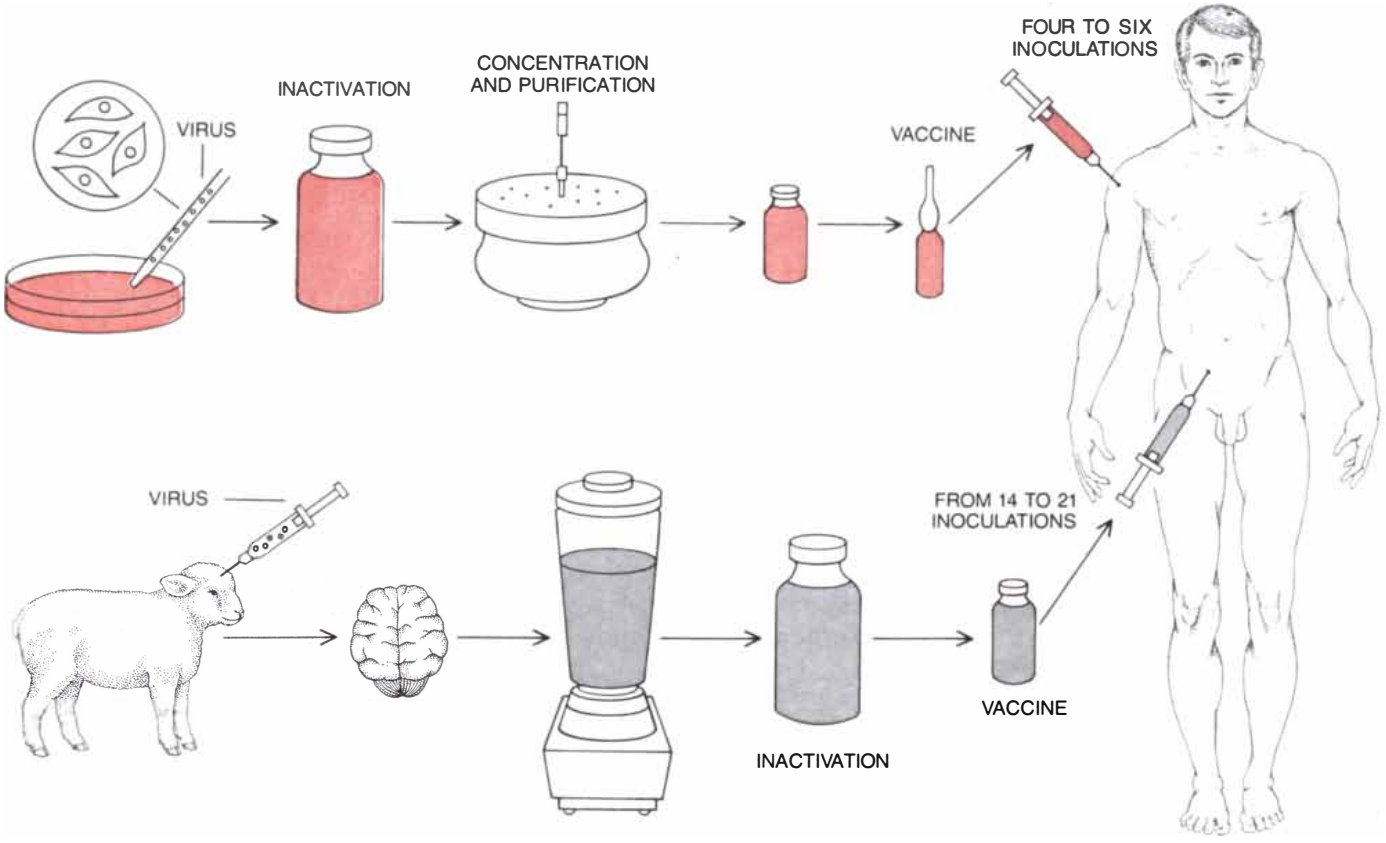
Although spontaneous recovery from rabies once the disease takes hold has been reported, it is extremely rare. In the past two decades only two patients have survived; they were treated by around-the-clock monitoring of vital functions (the respiratory, cardiovascular and nervous systems) and prompt intervention to restore them when the disease reached its critical stages of excitement and coma. Apart from these rare exceptions, Moses Maimonides' observations in his *Treatise on Poisons* (1198) are still valid: "Everything mentioned in the literature against the bite of a mad dog is useful, if at all, only when applied before the Rabies sets in. When such is the case, I have as yet seen nobody who escaped with his life."

Another way to attack rabies is to try to control it in animals. Many of the large rabies-free islands have successfully invoked quarantine periods of several months for dogs, cats and related carnivores being brought into the country. Effective and inexpensive vaccines are available to protect animals against rabies. One inoculation before exposure is sufficient to protect most animals for three years. Dog and cat rabies in cities can be easily controlled by vaccinating the animals and not allowing them to run free. Vaccines are also employed to protect livestock.

The control of rabies in wildlife is extremely difficult. Efforts to curb fox rabies by reducing the fox population have usually met with only temporary success. A plan to vaccinate foxes orally by means of bait (chicken heads containing capsules of attenuated-virus vaccine) is being tested in Europe. It is a questionable procedure because little is known about how such a strain of virus will behave under natural conditions of multiplication and possible passage from fox to fox. The virus may revert to virulence.

The control of bat rabies presents even more formidable difficulties. Efforts to reduce the vampire-bat population in Latin-American areas heavily afflicted by rabies have been fairly successful, but the procedures employed (ingestion of blood anticoagulants) seem to be feasible only on a local scale and are inapplicable with other species of bats. The commoner method of protection against bat rabies is the vaccination of cattle.

Even though rabies is one of the oldest recognized diseases of man, it remains a puzzle in many ways. Many of its factors (and their interrelation) need elucidation: the latency and prolonged incubation periods of the virus and the respective roles of antibody, of lymphocyte cell-mediated immunity and of interferon. The answers to such questions will be of importance not only for rabies but also for many other infections of the central nervous system that remain a medical and biological mystery.

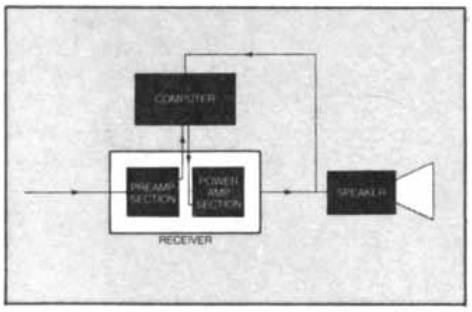


VACCINE PROCEDURES are depicted, beginning (*bottom*) with the conventional technique that originated with Louis Pasteur. An animal such as a sheep is injected in the brain with a fixed strain of rabies. On becoming sick the animal is killed, and its brain tissue is removed and homogenized in a blender. The virus is then inactivated by physical or chemical means. The treatment of a person exposed to rabies consists of from 14 to 21 injections of the inactivated virus under the skin of the abdomen. The treatment is painful and sometimes causes undesirable reactions. The new vaccine (*top*) is made from tissue cultures infected with an adapted fixed strain of rabies. The virus is inactivated and the resulting preparation is concentrated and purified. The vaccine is more potent than the Pasteur vaccine.

How to get an honest 30Hz from a 1.25 cubic foot speaker system.

The story behind the new KLH Computer Controlled Loudspeakers.™

With the introduction of acoustic suspension more than 20 years ago, the loudspeaker industry took an impressive step forward. This technology allowed speaker manufacturers to achieve full-range frequency response in a cabinet substantially smaller than any previous full-range loudspeaker.



Since then, breakthroughs have come and gone in the industry. But none that has significantly reduced the size of a true, full-range system.

The reason is actually quite simple. Accurate bass reproduction requires a woofer to displace a large volume of air. In a small system with a small woofer, the woofer cone must therefore travel a long way to reproduce the lower frequencies.

Although a small woofer is perfectly adequate most of the time, occasional high level, low frequency signals can drive the cone well beyond its intended excursion, causing severe overload distortion.

To avoid this, it has been necessary to attenuate lower frequencies in smaller systems. Which is why small speakers have always had compromised bass.

The KLH Analog Bass Computer.™*

To solve this problem, we developed a completely new approach — computer control.

We designed a separate component, the KLH Analog Bass Computer, as an integral part of the entire speaker system. This component sits next to the receiver or amplifier and constantly monitors its output. The computer derives an electronic analog of cone motion, and controls the woofer at the precise instant at which overload distortion would otherwise occur.

With this kind of accurate, reliable control, our designers were free to extract the optimum theoretical performance for any given cabinet size. And develop a line of loudspeakers that can deliver extended bass response in cabinets that are substantially smaller than ever before possible.

The KLH-1 is one example. From a 1.25 cubic foot cabinet, it delivers bass to 30 Hz (−3dB) at 105 dB s.p.l. with absolutely no possibility of overload distortion.

Beyond the Computer.

Since the Analog Bass Computer and the speakers must be designed as a single, integrated system, we started from scratch with the objective of optimizing our new technology.

To achieve the widest possible bandwidth with acceptable efficiency, we employed sixth-order equalized systems. Combined with the Analog Bass Computer, these systems provide a −3dB point equal to conventional acoustic suspension systems of at least four times their volume.


The new KLH-1 with Analog Bass Computer

In keeping with our objectives, we also refused to compromise other elements of the design.

For our cones, we selected polypropylene, a material first developed for use in studio monitors by BBC engineers. The movement of polypropylene reflects the electrical signal more faithfully than either paper or bextrene. The result is a remarkably clear, transparent, uncolored midrange.

For our speaker baskets, we used die-cast aluminum rather than stamped steel.

And we used massive magnet assemblies, optimized for the sixth-order design.

Three Applications.

Finally, we applied all we had learned to accomplish three distinct objectives.

Our first objective was to produce a speaker that raises the absolute level of low-frequency response in a cabinet that is still practical for the home environment. The new KLH-1 does exactly that. It delivers flat bass to 30 Hz (−3dB) from a floor standing unit just 11″ x 30½″ x 10¼″. At a price per pair of $1100** including Analog Bass Computer.

Our second objective was to provide the best possible combination of price and performance. Our solution is the KLH-2. At $660** per pair with computer, the KLH-2 can deliver flat bass to 38 Hz (−3dB) at 102 dB s.p.l. from a cabinet that measures 10¼″ x 21″ x 8½″.

Our third and final objective was to design a moderately priced speaker with performance equal to or better than anything near the cost, in a cabinet one fourth the size. This is the KLH-3. It measures 8½″ x 12½″ x 6″, delivers bass to 40 Hz (−3dB) at 95 dB s.p.l. and costs $450 ** per pair including computer.

The new line of KLH Computer Controlled Loudspeakers.

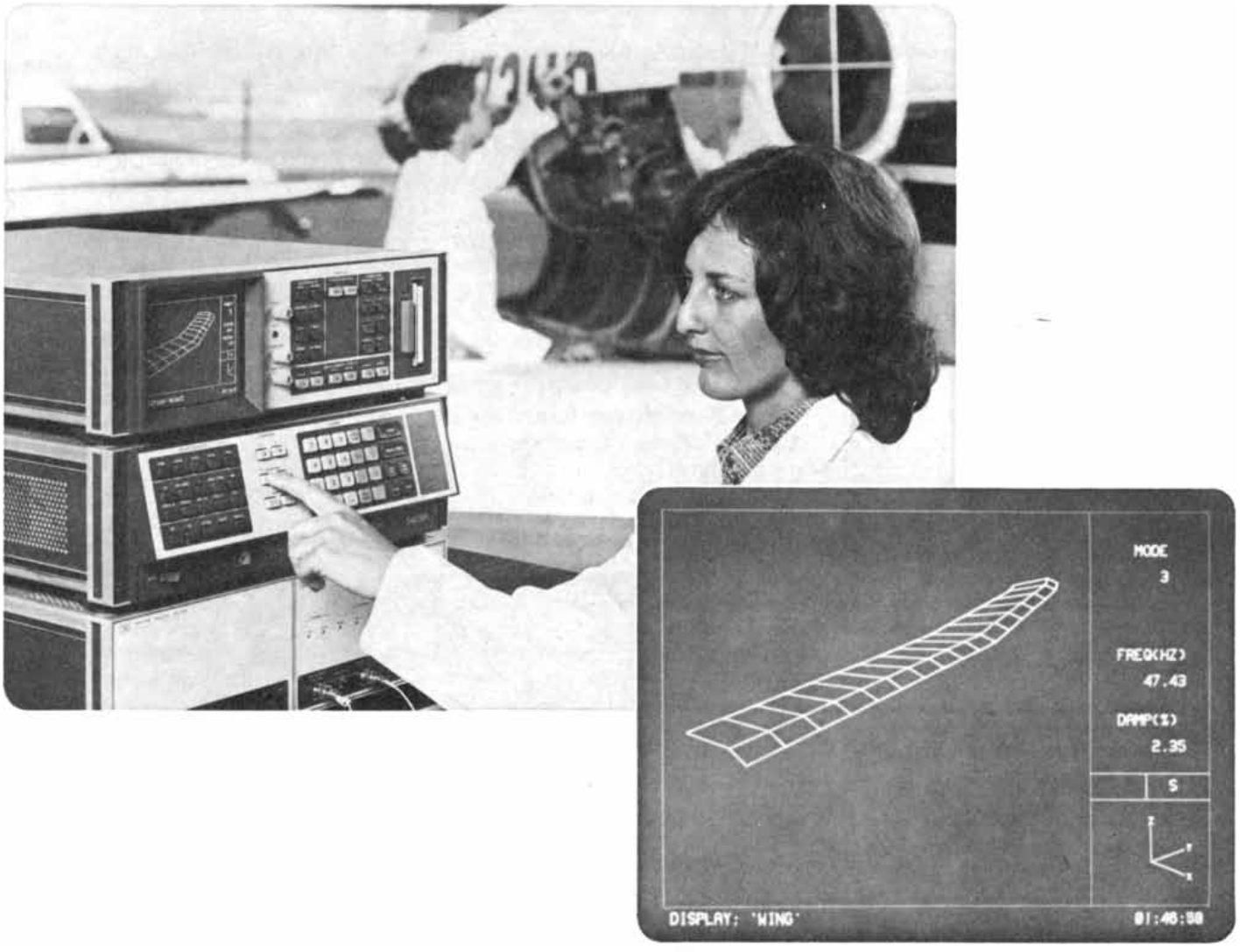
Listen to them.

For more information, call toll-free 800-225-6042 (in Mass. 1-800-532-9566). Or write KLH Research and Development Corp., 145 University Avenue, Westwood, MA 02090. In Canada: The Pringle Group, Ontario.

KLH

*Patent applied for.
**Manufacturer's suggested retail price.

HP computer and measurement advances



To help you design more efficient dynamic structures—from lawn mowers to aircraft—more efficiently.

Almost every structure vibrates. Analyzing the characteristics of its vibrations can tell you a great deal about how something—say an airplane wing or a power mower blade—will behave in use. And whether design changes can improve its performance, lengthen its useful life, or let you select alternative materials or avoid overdesigning with traditional ones.

The HP 5423 structural dynamics analyzer is the newest, smallest, and least expensive of Hewlett-Packard's family of modal (Fourier) analyzers. As such, it puts the benefits of modal analysis within the grasp of small manufacturers who previously thought they couldn't afford such a tool.

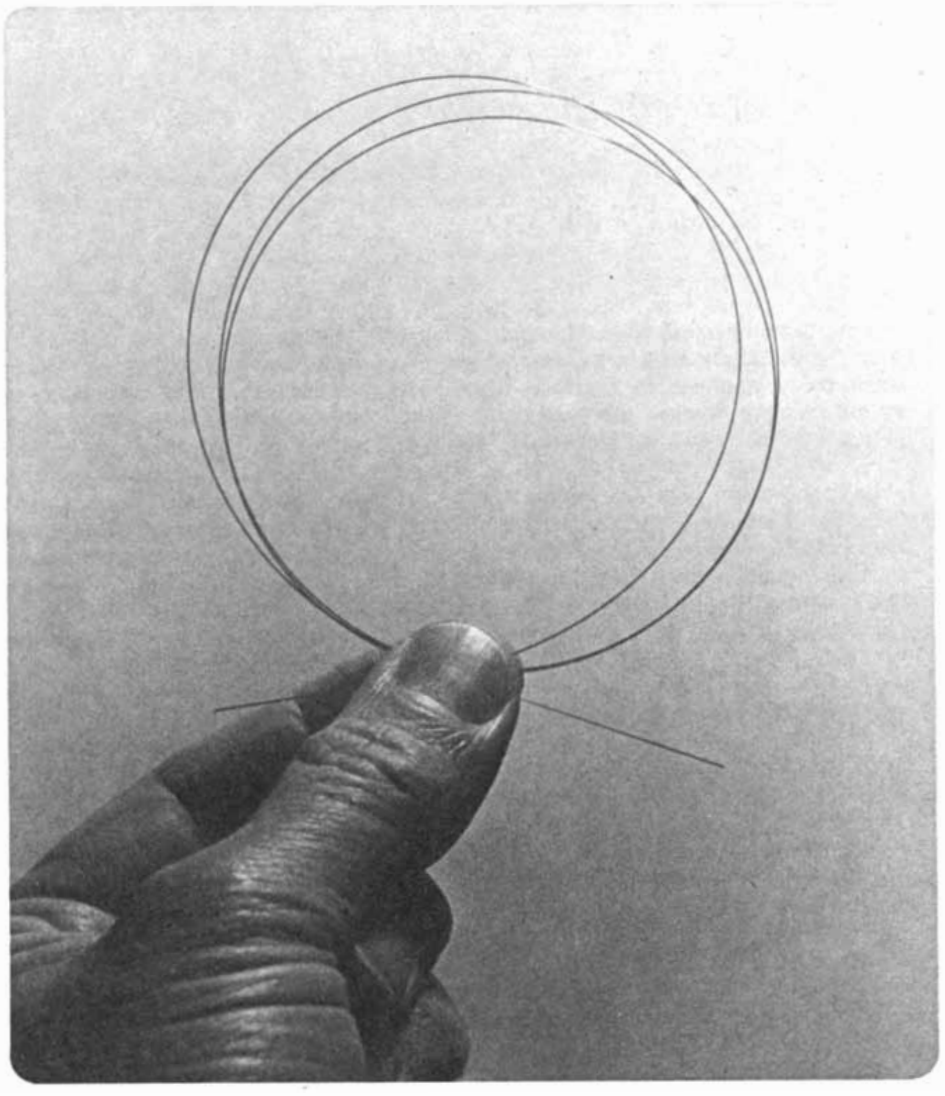
Although compact enough to be conveniently transportable, the HP 5423 has an impressive repertiore, with broad computed measurement capabilities in the d.c. to 25 kHz range that deliver answers upon which you can base design decisions.

Vibration data from a mechanical structure can be easily analyzed after some training by your present design staff without the services of a specialist. The instrument delivers answers in the form of *animated displays*, showing the deformation of the structure under test for each mode of vibration. The structure can be viewed from virtually any vantage. It can be rotated about any desired axis and viewed in perspective.

Results can also be displayed as tables or graphs. All data can be stored for further study on a built-in digital magnetic tape cartridge.

Price of the structural dynamics analyzer is $36,000*.

extend your possibilities.



HP's new flexible capillary column: a technological answer to a gas chromatographer's prayer.

Any chromatographer can tell you: the ideal GC column would be highly inert, thermally stable, mechanically strong, and easy to use. And the chromatographer will then probably point out that satisfying all these requirements at once has never been possible.

Now HP's flexible, fused silica capillary column can claim all these attributes and a few more. In developing this revolutionary column, HP drew on its extensive research in fiber optics. Close communication between this technical branch and our chromatography pros had suggested the possibility of modifying fiber optic production techniques for the manufacture of silica GC capillaries.

This fusion resulted in a thin-walled glass column, inherently straight yet very flexible, which can be wound into a coil without heating. The column's inherent straightness delivers the chromatographer from the tedious, delicate task of straightening the ends, and permits direct column connections.

And because there are virtually no metal oxides in the silica, analyses previously difficult or impossible can be performed with negligible sample adsorption or column deterioration.

Prices range from $100* for a 12-metre length to $275* for 50 metres.

For more information, mail the coupon to Hewlett-Packard, 1502 Page Mill Road, Palo Alto, CA 94304. Or call the HP regional office nearest you: East (301) 258-2000, West (714) 870-1000, Midwest (312) 255-9800, South (404) 955-1500, Canada (416) 678-9430.

HEWLETT PACKARD

Mail to: Hewlett-Packard, 1502 Page Mill Road, Palo Alto, CA 94304
Please send me further information on

( ) HP 5423 structural dynamics analyzer
( ) HP flexible GC capillary column

Name _____
Company _____
Address _____
City _____ State _____ Zip _____

*Domestic U.S. prices only.

00040

The Origins of the Water Turbine

How did the water wheel evolve into the water turbine, the machine that converts a head of water into shaft power and provides a fourth of the world's electricity?

by Norman Smith

The water turbine currently produces about a fourth of the world's electric power. Hydroelectricity is and will continue to be an immensely important component of energy production. In Europe a century and a half ago, however, the role of the water turbine was quite different. It was harnessed not as a means of generating electricity but as an improvement over the conventional hydraulic prime mover, the water wheel, and as an alternative to the steam engine for driving machinery.

In the 18th century the harnessing of the steam engine in British mines and mills was a crucial factor in achieving increased and more profitable production of coal, textiles, iron and other commodities. The view came to be accepted that one of the keys to growth in manufacture, that is, to bringing about an "industrial revolution," lay in a country's capacity to generate power.

The steam engine epitomizes Britain's break with the traditional European "wood and water" technology in favor of one based on coal and iron. By about 1820 this transition was, if not complete, advanced well beyond the point of no return. It was by no means, however, the result of shrewd planning on the part of British engineers and industrialists. They had little choice. First wood supplies and then water-power resources became so overused, depleted and eventually exhausted, at least in Britain's key areas, that alternatives had to be found.

At the beginning of the 19th century not every country with aspirations to industrialize could emulate Britain's exploitation of easily won and cheap coal. One such country was France, where the most abundant energy resource was water. Indeed, *houille blanche* (literally "white coal") is to this day the French term for water power. Harnessing its potential with more and better hydraulic motors was seen to be vital to achieving industrialization. In 1826 the Société d'Encouragement pour l'Industrie Nationale, whose formation 24 years earlier was a tacit recognition of the generally backward state of French industry, offered a prize of 6,000 francs to anyone who "would apply on a large scale, in a satisfactory manner, in factories and manufacturing works, the water turbines or wheels with curved blades of Belidor."

Belidor was an 18th-century hydraulic and military engineer, Bernard Forest de Belidor (1693–1761), who had described water wheels with curved blades in a monumental work, *Architecture hydraulique,* published in four volumes between 1737 and 1753. The wheels were located on the Garonne River at Bazacle. They belonged to a class of water wheels known as *roues à cuve,* or tub wheels. They were horizontal wheels, as opposed to the more familiar vertical water wheel. The tub wheel was about a meter in diameter and was set in the base of a tall cylindrical chamber. A tapered sluice supplied large quantities of water at a tangent to the chamber, so that the water entered the chamber with an appreciable rotational velocity. The weight of water above the wheel, together with the water's flow through the curved blades, made the wheel rotate through a combination of pressure and kinetic energy.

The tub wheels at Bazacle, and others like them elsewhere, had been derived from a simpler open form of wheel, turned by a water jet, through a process of steady evolution incorporating many increments of change and refinement. This wheel was a horizontal one known as a Norse mill; it was first illustrated early in the 15th century but was alluded to in an Irish document of the eighth century. The Norse mill achieved a wide distribution on the continent of Europe and in Ireland and Scotland but curiously was unknown in England and Wales. In the evolved form of the tub wheel it was particularly characteristic of southwestern France, the Pyrenees and Spain.

Belidor's reputation was such that his favorable description of the tub wheel must have persuaded the society, even more than half a century after his death, that this machine was the most amenable to improvement and was therefore a good basis for developing the water wheel of the future. Tub wheels did, in fact, have much in their favor. They were not difficult to build and install, and they were reliable and convenient in operation. The tub wheel was not, however, particularly efficient; at its best it reached an efficiency of 20 percent, and 15 percent was more typical.

By the 1820's much more efficient hydraulic motors were already at work in the Western world. There was the classic vertical water wheel, called the Vitruvian wheel after Marcus Vitruvius Pollio, the Roman engineer-architect of the first century B.C. It had evolved into three basic types: overshot, breast-shot and undershot. In the 1750's a series of cleverly contrived and perceptively observed experiments by the English instrument maker and engineer John Smeaton (1724–92) had demonstrated that overshot water wheels were superior, being at least 60 percent efficient and therefore twice as good as the undershot type. These conclusions, skillfully drawn by Smeaton from tests on working models, disengaged any further British interest in hydraulic-motor theory. Overshot water wheels quickly came to predominate in Britain until steam power took over completely.

French inquiries into water-power theory advanced on a much broader front, and French engineers succeeded in identifying two fundamental requirements of the ideal hydraulic motor. First, the water must always enter the motor without impact. Second, the water should have lost its initial velocity by the time it flows out of the motor. Both criteria insist in effect that no hydraulic energy be wasted in the form of turbulence or unused kinetic energy. Another French military engineer, J. V. Poncelet (1788–1867), applied these precepts to the undershot vertical water wheel Smeaton had shown to be an inferior motor half a century earlier. All that was needed to effect a marked improvement in the undershot wheel, Poncelet found, was an adequate number of correctly shaped blades. His improved

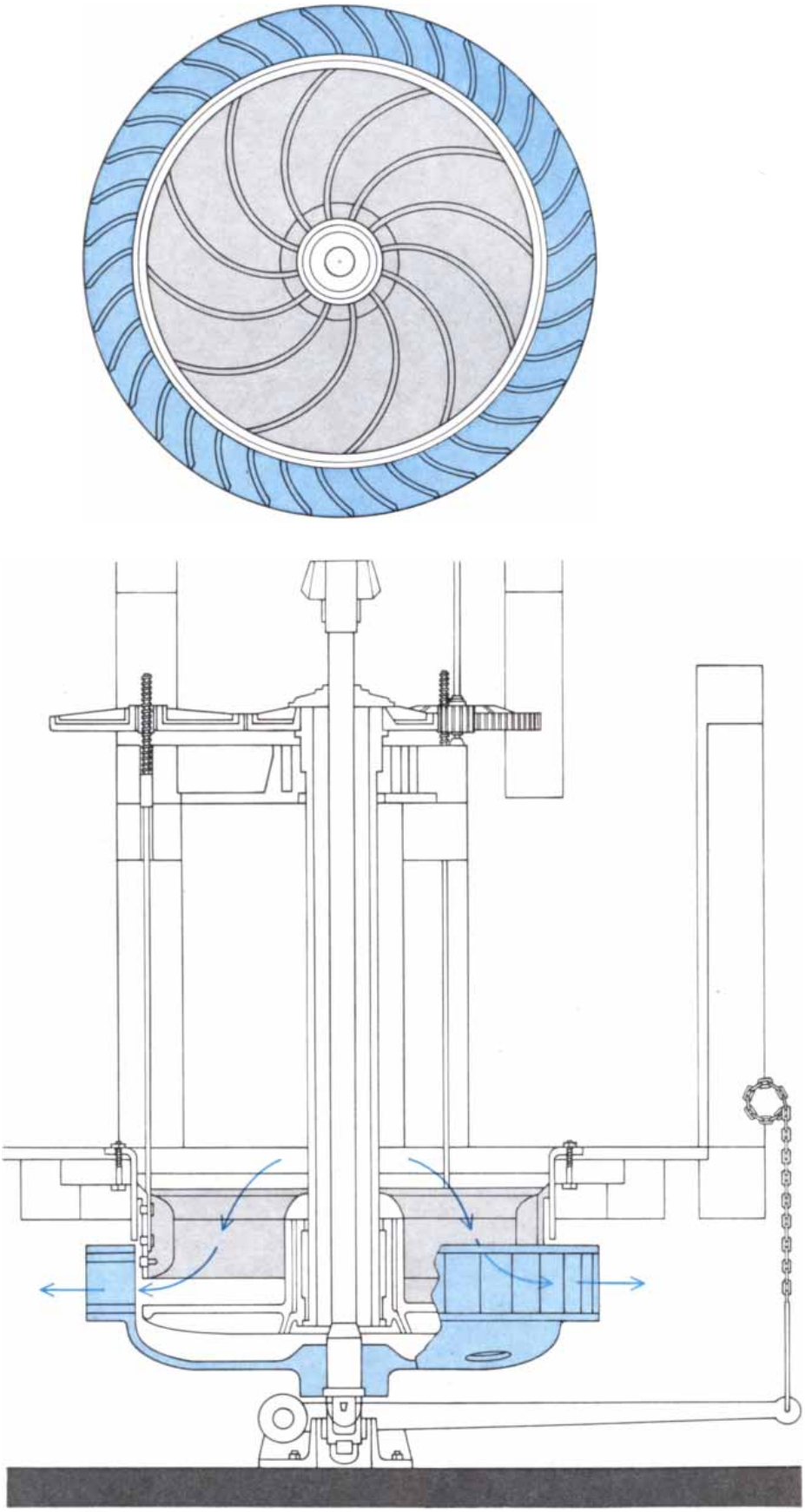
designs readily achieved efficiencies of from 60 to 70 percent, and the Poncelet wheel came into wide service.

By the beginning of the 19th century a third type of hydraulic motor figured prominently in the French development of water power. This motor was the reaction wheel. The origins of the machine are obscure, but the principle was simple enough: it was the same as that of the modern lawn sprinkler. Jets of water issued from the tip of two or more arms, the pressurized water being supplied to the center of the motor through a watertight bearing. The reaction caused the arms to rotate. The great Swiss mathematician Leonhard Euler (1707–83) investigated the theory of reaction wheels. With his son Albert he also experimented with working models.

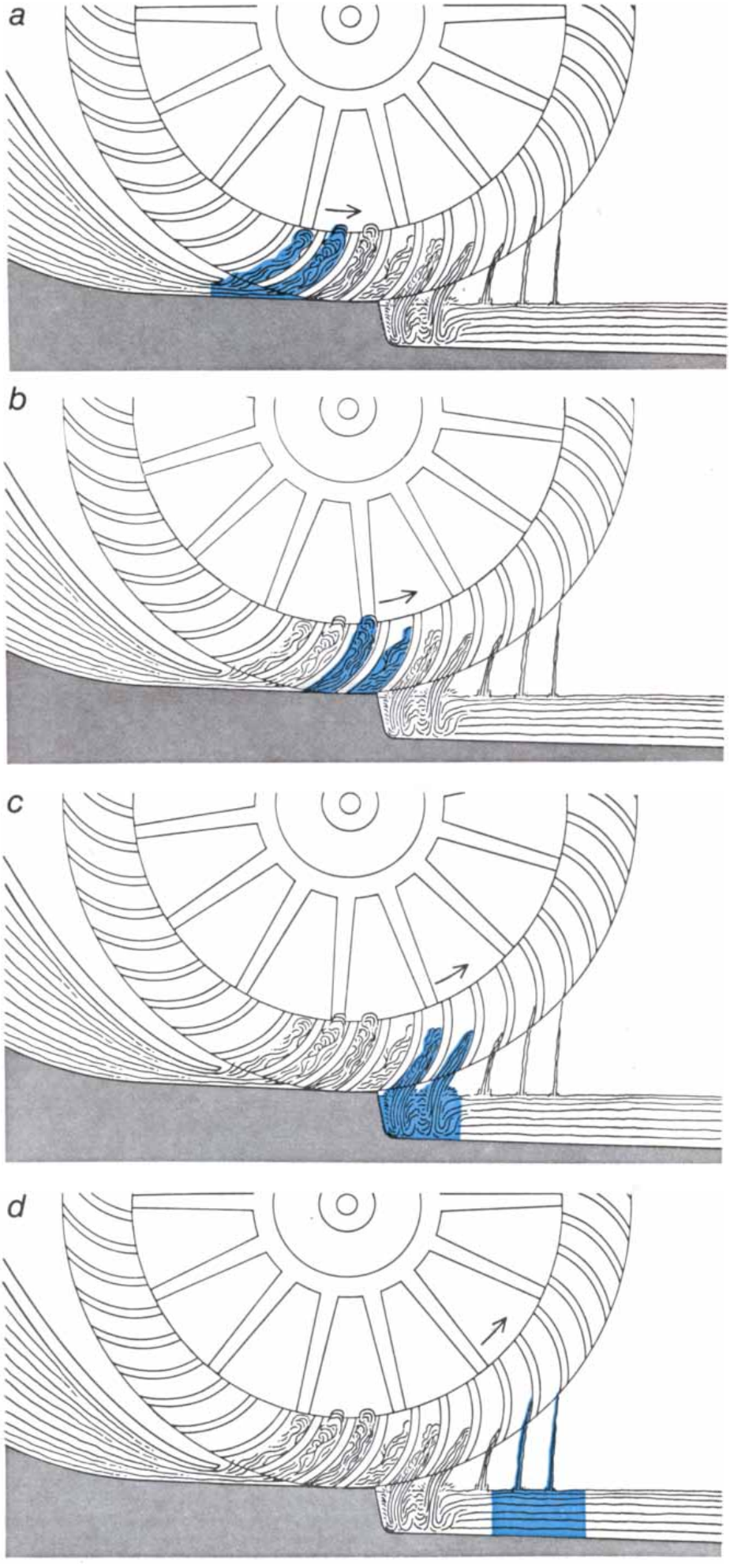
More influential, and much more significant as a practical proposition, was the reaction wheel introduced in 1807 by the Marquis Jean Charles Alexandre François de Mannoury d'Ectot (1777–1822). It had the graphic name *levier hydraulique* (hydraulic lever), and within a few years more than a dozen of Mannoury d'Ectot's wheels were at work in Normandy driving forges and ironworks. Mannoury d'Ectot next extended the principle of the reaction wheel to a multiduct version. The wheel took in water around the inner edge of a ring of flat or curved blades and discharged it at the outer edge. He constructed the first of these wheels at a site in Normandy, the Montaigu mill near Caen, in about 1804. It was the subject of a report by Lazare Carnot (father of the famous French physicist Nicolas Léonard Sadi Carnot) to the Academy of Sciences on June 21, 1813. The crown of the wheel, in the shape of a bell, carried the blades and was fixed to a vertical axle that transmitted the motion. The motor ran submerged.

In engineering terms Mannoury d'Ectot's motor was a radial outward-flow machine, the whole of the rotor being supplied with water all the time. This arrangement is termed complete admission. Poncelet had appreciated the same possibility; he viewed it as a way to overcome a fundamental weakness in his own rationally designed vertical water wheel. In that undershot wheel the water was fed to the blades at one position only; moreover, the flow entered and left the wheel at the same point. Therefore a reversal of flow within the wheel was unavoidable. Such a reversal was also undesirable. In 1826, in a lecture given in Metz, Poncelet proposed to deal with this difficulty simply by shifting his wheel from vertical to horizontal, admitting the water all around the periphery and discharging it without velocity through the interior. This machine was in effect a radial inward-flow one.

When the society offered its prize in



FIFTY-HORSEPOWER TURBINE that provided the air blast for a French ironworks, built in 1832 by Benoît Fourneyron, is seen in plan (*top*) and section (*bottom*). The rotor, a large one-piece metal casting (*color*), ran submerged; the water flow, guided by a series of fixed vanes (*gray*), was from the center outward. The rotor was 2.4 meters in diameter and the head of water was 1.3 meters high. The largest turbine then known, it helped to win Fourneyron a prize of 6,000 francs for the satisfactory application of water motors to large-scale industry. All of the more than 100 water motors that were built by Fourneyron were outward-flow designs in spite of the disadvantage of the large, heavy rotor at their circumference. The bearing at the bottom of the shaft was also subject to excessive wear and had to be replaced frequently.



1826, the challenge was framed in a curiously specific form. It was true that Belidor's wheels were amenable to improvement, but they were not the only type, or even the most promising type, on which to base future development. They did not represent the most advanced concepts or practices, and they were only one of many modes of extracting power from water. By 1826 the efficient all-metal overshot wheel, the Poncelet wheel set either vertically or horizontally and the *levier hydraulique* and its variants were also available. Indeed, these later machines constituted the bulk of the diverse contemporary repertory. One reason their efficiency was high was that their design had been guided by theoretical principles.

In the search for the origins of the water turbine the breadth and diversity of the types of water wheel in the early 19th century have all too often been overlooked. When one encounters a new term, in this instance the term turbine, one's impulse is to search for the name of an inventor to go with it. Here there is none. No one person invented a fundamentally new kind of hydraulic motor and thereby consigned all the old-fashioned motors to the scrap heap overnight. The real picture is one of various parallel lines of development converging by degrees. From this convergence gradually emerged new and better concepts and techniques. One French engineer in particular, Benoît Fourneyron (1802–67), played a crucial role in the evolution of the water turbine. Another, Claude Burdin (1790–1873), invented the term.

Burdin was a military engineer (an officer in the Royal Corps of Mines) and a teacher at the School of Mines in Saint-Étienne. In 1822 he submitted to the Academy of Sciences a paper titled "Des turbines hydrauliques ou machines rotatoires à grande vitesse." This was the first time the word turbine was used. Burdin had coined it from the Latin *turbo:* that which spins, for example a top. When in 1824 the academy reported on Burdin's paper, interest was expressed but only qualified approval. It

TWO THEORETICAL CRITERIA for an ideal hydraulic motor are diagrammed. The first (*a, b*) is that the water must enter the motor without wasting energy in turbulent impact. The effect is shown for two curved paddles of an undershot water wheel developed by the French military engineer J. V. Poncelet; the water (*color*) speeding from the sluice glides uphill along the concave surface of the paddle with minimal turbulence. The second criterion is that by the time the water flows out of the motor (*c, d*) it should have lost all its initial velocity. The criterion is seen to be met when the water begins to drain into the millrace, deprived of further forward motion and impelled only by gravity. Efficiency of Poncelet's water wheels reached 70 percent.

Patience.
It's part of our character.

A paper crane brings joy.
Origami, the art of paper folding,
cannot be rushed. It takes time,

and to us time well spent, when it
brings such happiness.
Patience, Japan Air Lines style,
means that you will never feel
rushed or overlooked or forgotten.
To our passengers,
we give the gift of time.
Time to beguile a child with *origami*;
time to tuck-in a blanket; to care.
Nintai. Patience. Because we never
forget how important you are.

The way we are is the way we fly.

JAL
JAPAN AIR LINES

What kind of man owns his own computer?

Rather revolutionary, the whole idea of owning your own computer? Not if you're a diplomat, printer, scientist, inventor... or a kite designer, too. Today there's Apple Computer. It's designed to be a *personal* computer. To uncomplicate your life. And make you more effective.

It's a wise man who owns an Apple.

If your time means money, Apple can help you make more of it. In an age of specialists, the most successful specialists stay away from uncreative drudgery. That's where Apple comes in.

Apple is a real computer, right to the core. So just like big computers, it manages data, crunches numbers, keeps records, processes your information and prints reports. You concentrate on what you do best. And let Apple do the rest. Apple makes that easy with three programming languages—including Pascal—that let you be your own software expert.

Apple, the computer worth not waiting for.

Time waiting for access to your company's big mainframe is time wasted. What you need in your department—on *your* desk—is a computer that answers only to you... Apple Computer. It's less expensive than timesharing. More dependable than distributed processing. Far more flexible than centralized EDP. And, at less than $2500 (as shown), downright affordable.

Visit your local computer store.

You can join the personal computer revolution by visiting the Apple dealer in your neighborhood. We'll give you his name when you call our toll free number **(800) 538-9696**. In California, **(800) 662-9238**.

Apple Computer, 10260 Bandley Drive, Cupertino, CA 95014.

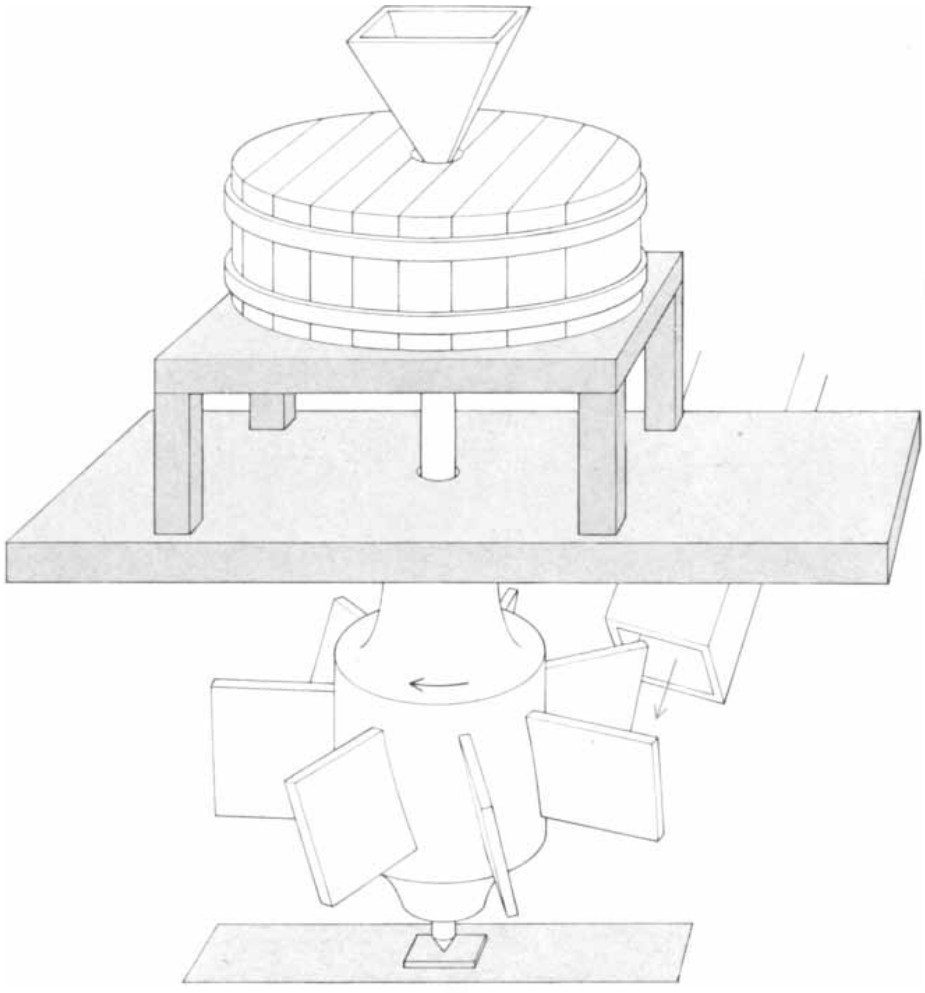
apple computer

was remarked that although Burdin's theoretical analysis rightly emphasized the twin requirements of water entry without shock and exit without velocity, these conditions were by 1822 hardly original. Nor were the evaluators satisfied that Burdin's tests of his high-speed rotary machines had been "sufficiently numerous to give his report all the precision desirable."

Actually the most original and potentially the most promising of Burdin's several ideas was a design for an outward-flow radial turbine driven by the free efflux of water from a centrifugal runner. According to the theoretical concepts embodied in Burdin's approach to turbine design, the stationary guide vanes and runner blades (six of each) were appropriately shaped and angled to ensure shockless entry and zero exit velocity. Evidently he was familiar with contemporary water-wheel theory, and his outward-flow radial design was conceptually important: it was the fundamental arrangement employed by Fourneyron in every turbine he built. For Burdin, however, practical proof of the validity of his design was another matter. It appears that he never demonstrated even a working model, let alone a prototype.

In 1827 Burdin attempted to win the society's prize of 6,000 francs. To his earlier theoretical analyses he added some discussion and description of working machines that had been built and used. These machines were not, however, practical manifestations of the radial outward-flow concept; they were no more than elementary reaction wheels, in most cases variations on Eulerian machines. Burdin was nonetheless quite at ease in designating every one a turbine. His reaction wheels may have been quite practical devices in the forges and sawmills of Puy-de-Dôme that they served, but they were unexceptional in terms of power, speed, efficiency and novelty of concept. They were certainly not considered a sufficient improvement on existing hydraulic motors to meet the criteria of the society. Burdin was, however, rewarded for his theoretical designs with encouraging words and a consolation prize of 2,000 francs. The big prize remained intact.

From 1817 to 1819 Fourneyron was a teenage pupil at the School of Mines; for a time in 1818 he studied under Burdin. To what degree these studies influenced Fourneyron's interest in water power is uncertain, even though some confident assertions have been made on this point. In any event before Fourneyron had reached his majority he had built a water-powered plate-rolling mill on the Ognon River, which flows into the Saône east of Dijon. The success of the mill prompted him to undertake a more thorough study of the design and construction of water wheels.



NORSE MILL, a water wheel with a horizontal array of paddles rather than the vertical array characteristic of Vitruvian wheels such as Poncelet's, was the precursor of the fast-turning horizontal water motors that came into use in the western Mediterranean and elsewhere in Europe during the 18th century. The wheel illustrated here drove the grindstone of a grist mill.

Fourneyron's first experimental water turbine was developed and tested at the same location, Pont-sur-l'Ognon, between 1823 and 1827. It was a small machine driven by a 1.4-meter head of water, and it could operate submerged or unsubmerged. At its maximum efficiency, which was more than 80 percent, it developed about six horsepower at 60 revolutions per minute. The water flow was radial and outward and admission was full. This was an arrangement from which Fourneyron never deviated. In 1827 he had to move his flourishing engineering practice to Besançon, on the Doubs River, a stream in Franche-Comté that runs parallel to the Ognon. There between 1827 and 1832 he further developed his ideas in the construction of two more turbines. Both were harnessed to provide the air blast at the ironworks of Dampierre and Fraisans. The Fraisans motor was the largest turbine of its day: its wheel, 2.4 meters in diameter, delivered 50 horsepower under a 1.3-meter head of water.

Those nine years of painstaking design and testing had enabled Fourneyron to bring not only his ideas but also three working turbines to the point where he felt confident enough, at age 31, to submit his own memoir to the society. In 1833 his detailed account of the theoretical and actual performance of the turbines won him the 6,000-franc prize. After nearly a century Belidor's tub wheel had been "officially" improved.

Over the remaining 30-odd years of his life Fourneyron designed and installed more than 100 turbines in France and elsewhere in Europe. A few of his turbines even went to the U.S., the first in about 1843, and their influence there was considerable. Fourneyron's motors could accommodate to widely differing conditions. For example, in 1837 he installed two of them in a spinning mill at Saint Blaisien in the Black Forest. One operated under a 108-meter head of water and the other under a 114-meter head. These were pressures of unprecedented magnitude, and they brought Fourneyron a fair share of difficulties. Not the least of the difficulties was that the bottom bearing of the turbine shaft was so heavily loaded that it had to be renewed every 10 to 14 days. The performance of the Saint Blaisien turbines, however, was a revelation: they ran at

We pay these scientists to make our production lines obsolete.

As soon as there's a better way to do something, the old way becomes obsolete.

And these scientists do nothing *but* design new and better ways for Western Electric to manufacture products for the Bell System. In the last two years, they and their colleagues at Western Electric's Engineering Research Center have helped cut costs by over $156 million.

They are a vital part of Western Electric's commitment to helping the Bell System keep the cost of communications service down, while keeping the quality up.

The Factory's Laboratory

The people at our Engineering Research Center work full time on manufacturing research. Over 50% of them are PhD's.

At home in the factory as well as the lab, they turn theoretical experiments into real savings. For example, while exploring ways to improve production of telephone circuit boards, they invented a totally new material with which to make the boards and a whole new way to mass-produce them.

Another breakthrough is already saving more than $2 million a year. It's a new soldering technique that makes 20,000 connections in less than 60 seconds.

The list of money-saving developments goes on. A new computer program that makes technical drawings in one-fourth the time of manual methods. An automatic system for sorting telephone parts by color. A process that measures the production of optical fiber 1000 times per second, keeping its diameter accurate within 30-millionths of an inch. And there's a new system that takes only 30 seconds to help align tiny integrated circuits that are 200 times thinner than the thinnest human hair.

These are just a few examples of what the Engineering Research Center does throughout *all* of Western Electric's manufacturing operations — developing new ideas in technology to improve telephone products and produce them at lower cost.

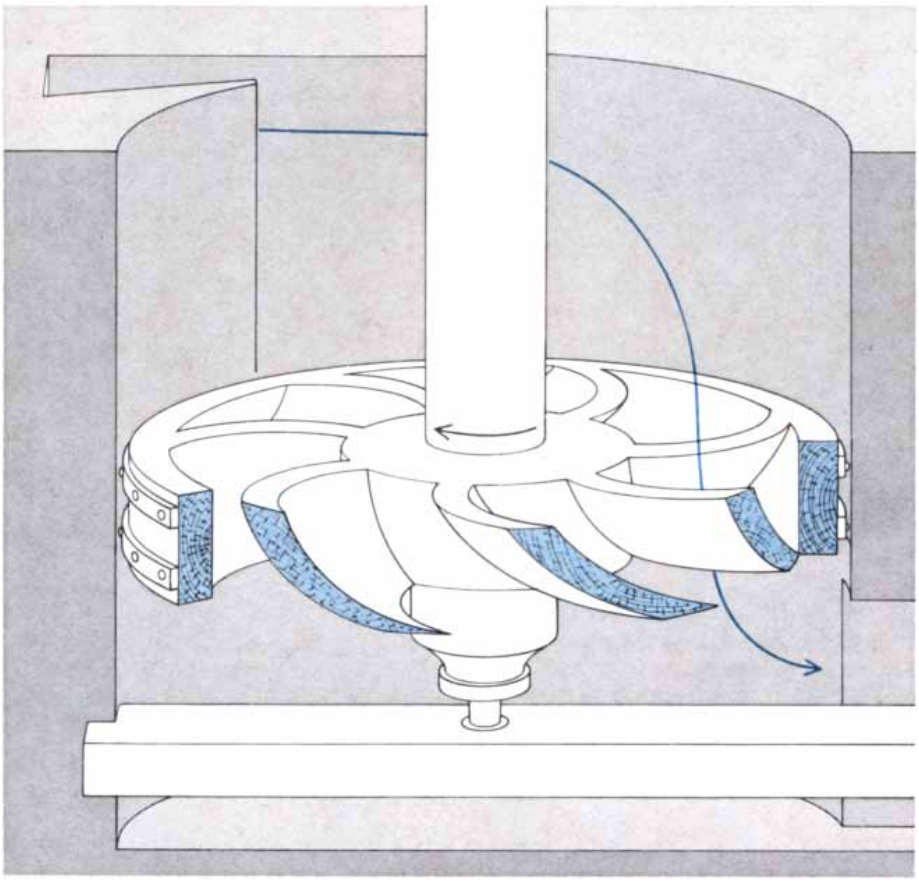
Anticipating the Future

Much of the work of the Engineering Research Center is designed to keep Western Electric ahead of technology, by anticipating what new processes will be needed for tomorrow's products.

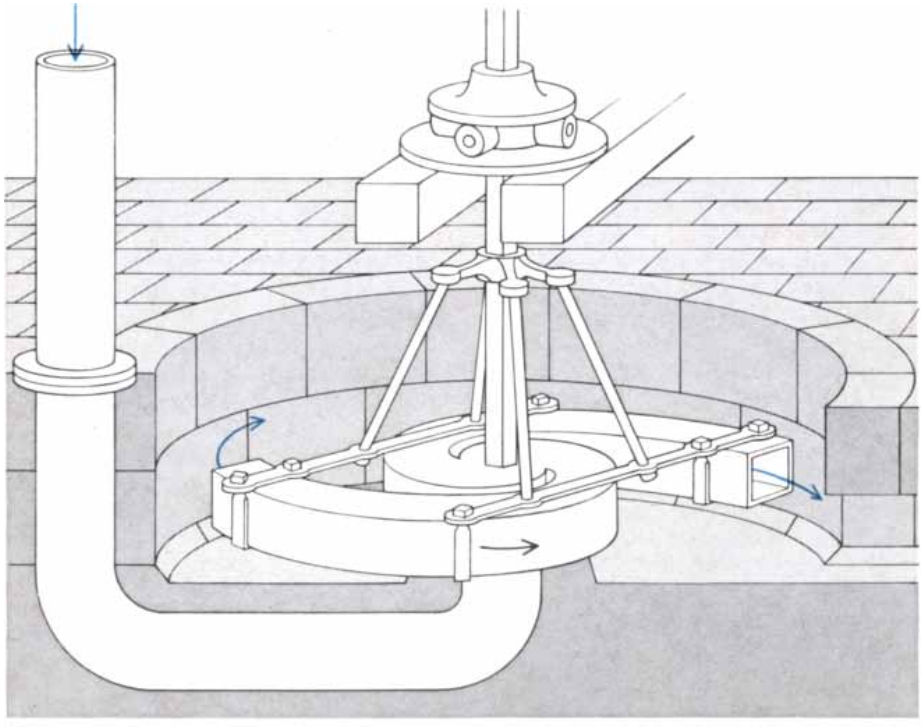
It's a total and ongoing commitment. A commitment to provide you with the highest quality telephone service at the lowest possible cost.

Keeping your communications system the best in the world.

Western Electric



ADVANCED NORSE MILL, described by Bernard Forest de Belidor, ran entirely submerged. The ironbound wood wheel has been cut away to reveal its curved horizontal blades. Water entered the wheel chamber from the deep, tapered sluice at the left; the pressure of the water above the wheel and its rotary motion through the array of blades combined to turn the wheel.



NOVEL WATER MOTOR, the "hydraulic lever" of the Marquis Jean Charles Alexandre François de Mannoury d'Ectot, was a reaction wheel identical in principle with rotary lawn sprinklers. Its origin is not known, but the great mathematician Leonhard Euler experimented with such devices. Mannoury d'Ectot's water motor was introduced in 1807 and then gained popularity as a prime mover for forges and ironworks. Making the connection between supply pipe and the twin-armed rotor watertight was a problem that was never satisfactorily solved.

2,300 r.p.m. and developed 60 horsepower at an efficiency of more than 80 percent. For a textile mill in Augsburg, Fourneyron built two very-low-head turbines, each of which developed 220 horsepower. They were his most powerful turbines: between them they drove 30,000 spindles and 800 looms.

Just as Fourneyron's work was instrumental in drawing together decades of evolving ideas, so did it become a starting point for future developments. Initially a number of engineers, most of them in France, attempted to develop Fourneyron's outward-flow concept. They were not successful. Inherent in the design were defects that were soon identified. For example, it turned out that the performance of Fourneyron's turbines was impressive only under certain limited conditions. Running submerged at full flow and at the pressure for which they were designed, they were highly efficient. If conditions changed, and in particular if the water flow was reduced in order to reduce the power output, their efficiency dropped sharply. Operating experience revealed another problem that was potentially even more serious. If the load on the turbine was suddenly removed, then unless the turbine was shut down promptly it would pick up speed and race, conceivably until it flew to pieces.

Not surprisingly, then, the Fourneyron turbine was quickly outmoded. The most fundamental revision was the abandonment of the outward-flow configuration. Three alternatives were investigated; after due time and effort all of them yielded good results. In the 1840's European engineers addressed themselves to the concept of axial flow, a category of hydraulic turbines that is represented today by "propeller" turbines of the Kaplan type. In the U.S., Samuel Howd and Uriah A. Boyden (1804–79) laid the foundations that supported the precision work of James B. Francis (1815–92). Francis established the inward-flow motor and subsequently the mixed-flow one. These motors represent a second modern category of hydraulic turbines, a category that still carries, not altogether appropriately, Francis' name.

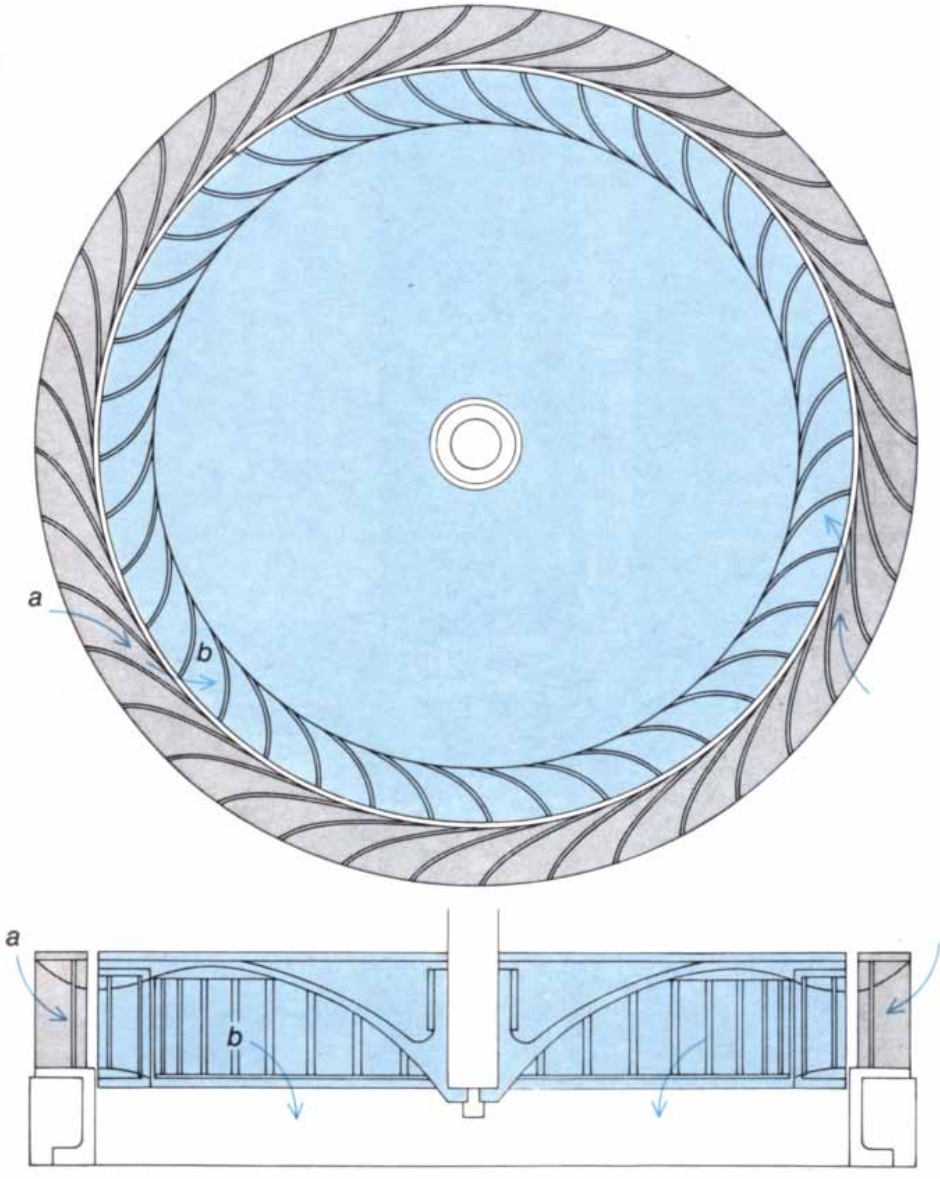
A third line of development, owing little or nothing to Fourneyron, followed the concept of a vertical wheel driven by a jet of water applied at one point on its circumference. The flow of water through the blading was inward and radial in the manner described by Poncelet in 1826. The approach led ultimately to the Pelton wheel: a jet or jets of water impinge on an array of specially shaped buckets closely spaced around the rim of the wheel. The Pelton wheel was developed at the end of the 19th century by a group of California engineers, among them Lester A. Pelton (1829–1908). Pelton was by no means

the only one in the group with good technical ideas, but he appears to have had the best commercial acumen.

In short, the evolution of hydraulic prime movers advanced by degrees, by increments of change and improvement continuously monitored and evaluated for their worth. The good ideas accumulated; the bad ones were discarded. Given this perspective, who is to be credited with inventing the water turbine? Should it be Fourneyron or can one make out as good a case, or a better one, for Francis or Pelton, or Burdin or Mannoury d'Ectot? Actually it is not a very rewarding line to follow. What turns out to be more profitable is to examine the transition represented by the introduction of the word turbine. By taking this approach one comes to see how some historians of technology have deluded themselves.

Performance figures for 18th-century water wheels are difficult to derive. For the rotational speed of the wheels, however, at least some orders of magnitude can be given. An undershot wheel driven by a head of 16 feet of water would ideally have had at its periphery a velocity of 16 feet per second. If the wheel was 20 feet in diameter, its rotational speed would have been 15 r.p.m. Depending on the diameter of the wheel and variations in the available head of water, one can postulate a range of speeds between 10 and 20 r.p.m. Overshot wheels were in general decidedly slow-turning. Smeaton advised peripheral speeds of no more than two to three feet per second, although usually a rather higher figure (four to eight feet per second) was favored in order to have the wheel act to some extent as a flywheel. Even so, for a 20-foot wheel such speeds represent no more than 8 r.p.m.

Horizontal wheels turned faster; from



INWARD DIRECTION OF FLOW is shown in plan (*top*) and section (*bottom*). The turbine is of the type first proposed by Poncelet in the 1820's and patented in the U.S. in the 1830's. The water enters the motor through an array of fixed blades (*a*) and makes a runner, or central rotor (*b*), spin. The water then enters the race close to the center of the motor. This is a turbine of the kind first built by James B. Francis in 1849 and subsequently installed in Lowell, Mass.

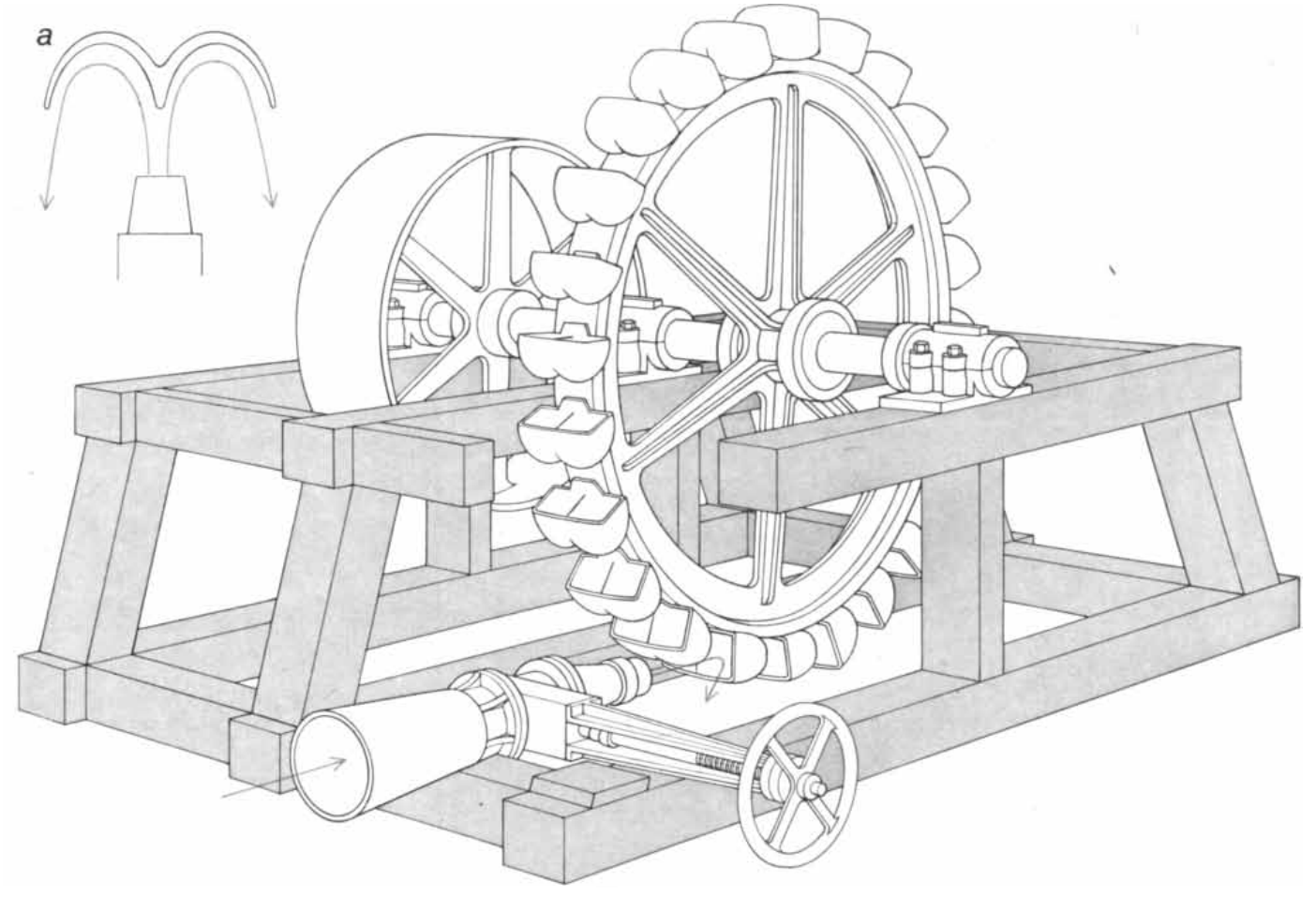
the first a speed of 50 r.p.m. would seem to have been typical. Later this speed was considered low. Some 18th-century wheels, of a type called *rouets volants* (flying wheels), reportedly turned at a rate of from 40 to 60 r.p.m. In the Balkans today simple horizontal water wheels typically turn at from 70 to 90 r.p.m. In France tests conducted in and around Toulouse in the 1820's showed that the speeds were commonly more than 100 r.p.m. and occasionally as high as 130 r.p.m. Even one of Burdin's Eulerian machines is said to have turned at 80 r.p.m.

Hence it is not simply a convenience to classify water wheels according to whether they are set vertically or horizontally; the two types also occupy two distinct ranges of speed. That is evidently how water wheels were differentiated in the opening decades of the 19th century. It therefore emerges that when Burdin titled his 1822 paper "Des turbines hydrauliques ou machines rotatoires à grande vitesse," he was using the word turbine not to designate any new category of hydraulic motor but merely to designate the category of water wheels whose speed of rotation was comparatively high.

Beyond that Burdin did not discriminate. He had no hesitation about describing all his machines as turbines; they all turned rapidly. As it happens, all Burdin's machines were also set horizontally. Soon thereafter the orientation of the wheel became just as much a criterion as its shaft speed. For example, in 1847 it was stated that "horizontal wheels are usually called by the general name of turbines." Eleven years later the French engineer Jacques-Eugène Armengaud (1810–91), in a comprehensive study of hydraulic motors, titled one chapter, "Des turbines ou roues horizontales." In the same year another French engineer, Arthur Morin, referred to "des nouvelles roues à axe vertical, appelées turbines." He went on to note: "The name turbine is new and was introduced by M. Burdin, who applied it to one such wheel that he had constructed, and since then it has been applied indifferently to all vertically axled wheels that derive some advantage by running immersed in their tail race."

The distinction in nomenclature soon became unimportant. By the second half of the 19th century the Vitruvian vertical water wheel was a thing of the past and all hydraulic motors were called turbines. Actually they might just as well have been called water wheels (and they not infrequently were, particularly in America), or perhaps "improved water wheels" or "superior water wheels." And late in the 19th century the Pelton configuration was called the Pelton wheel and only rarely the Pelton turbine.

Nevertheless, a certain preoccupation with definitions has persisted. Water wheels and turbines are often differentiated by stating that turbines are smaller, run at higher speed, will work submerged, can utilize a wide range of heads of water and finally are more powerful and/or more efficient. That, however, is merely compiling a list of the superior characteristics of later water wheels rather than fundamentally different characteristics. Thereby a point is missed. For Burdin's own reasons, and they were not necessarily bad ones, he invented a word. He did not invent a new machine, and no one else did either. Rotating hydraulic motors, formerly called water wheels, were improved by degrees, over the years and through the work of many engineers. As part of this extended sequence of development and improvement that kind of hydraulic motor became known as a turbine.



VERTICAL MOTOR, the Pelton wheel, was designed in the 1880's. A water jet from a nozzle at the bottom struck a succession of paired cups fixed on the rotor. Curved design of the cups (*a*) served to extract maximum energy from the rushing water. One 220-pound Pelton wheel, a portable two-man load for gold miners, could provide 125 horsepower. The configuration is used in some modern water turbines.

SCIENCE/SCOPE

A ground-based radar now being developed will be so "quiet" that anti-radiation missiles will be less likely to home in on its beam. The new radar will have two antennas, one to transmit low-energy beams continuously and the other to listen for returns. (Conventional radars differ by transmitting high-energy pulses so that one antenna can alternately transmit and receive.) A new antenna technique reduces the radar's side lobes -- the secondary patterns of energy that enemy missiles can home on. Hughes is building a prototype quiet radar antenna for evaluation by the U.S. Army.

Using special temperature-controlled chambers, NASA scientists will create clouds for study aboard Space Shuttle flights in the early 1980s. To properly form clouds in the weightlessness of space, the chambers, which are flat-plate heat pipes, must be extremely level over a large area (2'x3'x3/4") and uniform in temperature to within .01°C. Neither requirement has ever before been met in a heat pipe of this size. Hughes, under contract to General Electric, is developing eight isothermal vapor chambers to form the inner walls of the Atmospheric Cloud Physics Laboratory. The project is managed by NASA's Marshall Space Flight Center at Huntsville, Alabama.

A new microwave sensor will allow military weather satellites for the first time to "see" through clouds to monitor meteorological conditions below. The instrument will scan a 1300-kilometer field of view every 1.9 seconds from an altitude of 450 nautical miles, sensing microwave radiation energy reflected from ice, land, and clouds. It will detect rainfall, ice masses, ocean wind speeds, soil moisture content, and other conditions, and relay compiled data within minutes nearly anywhere in the world. The information will help commanders of land, sea, and air forces in their planning of operations that depend on accurate weather forecasts. Hughes, under a U.S. Air Force contract, is to build one prototype and develop computer software for ground processing.

The Manufacturing Division of Hughes Missile Systems Group in Tucson has immediate openings for engineers. Typical openings range from digital logic, analog, and IF/RF circuit design to electro-optical and IR system design. Also needed are industrial engineers and manufacturing production engineers to conceive and design production methods and tooling for the fabrication of complex hardware for missile systems. For immediate consideration, send your resume to Roy McCalla, Dept. SE, Hughes Aircraft Company, P.O. Box 11337, Tucson, AZ 85734. Or call (602) 746-6200. Equal opportunity M/F/HC.

Three Intelsat IV communications satellites have exceeded their design life of seven years and are still operating. The eldest of the trio, positioned in geostationary orbit over the Atlantic Ocean, was launched in December 1971. The second was launched a month later and placed over the Pacific. The third, now over the Indian Ocean, was launched in June 1972. Today, more than 30 Hughes communications satellites have combined for more than 150 years of service.

THE AMATEUR SCIENTIST

A homemade spectrophotometer scans the spectrum in a thirtieth of a second

by Jearl Walker

Over the past dozen years this department has presented several designs for spectroscopes, spectrophotometers and related instruments for spreading light into its component colors (wavelengths) so that it can be analyzed. The instruments were of two basic types, one for studying the light-absorbing properties of a transparent liquid, solid or gas and the other for analyzing a source of light such as a lamp or a star. A new design for a special type of spectrophotometer was sent to me recently by Dean Morelli of Rye, N.Y. Its primary advantage over the instruments described here previously is its speed in scanning the visible spectrum: it can obtain a complete set of measurements in as little as a thirtieth of a second.

I can best set the stage for describing Morelli's instrument by reviewing the two basic types of spectroscope and spectrophotometer. In the first of them white light from a lamp is directed through a cell containing the liquid, solid or gas to be investigated. The sample absorbs the light at certain wavelengths and so diminishes it. The transmitted light is dispersed (separated according to wavelength) by a prism or a diffraction grating before it reaches a detector. In this way the detector can determine which colors have been removed from the initially white light by the sample. With reference works on atomic and molecular spectroscopy the observer can identify the constituents of the sample and can distinguish the sample from others that might look similar to the unaided eye even though they differ in composition.

The second basic type of instrument is the one most frequently employed in astronomical work. The spectroscope is attached to a telescope so that the observer can identify the colors in the light from a celestial object. Sunlight has been analyzed in this way by many amateurs. The sun radiates light throughout a range of wavelengths, with the peak intensity falling near the middle of the visible spectrum. When that basically white light passes through the sun's atmosphere, the elements in the atmosphere absorb the light at their characteristic wavelengths. The absorbed wavelengths appear as thousands of dark lines (the Fraunhofer lines) superposed on the colors of the visible spectrum. With the appropriate reference works the observer can identify the lines with the elements in the sun's atmosphere.

A spectroscope relies on the human eye for examination of the spectrum. If a light-sensitive device such as a photomultiplier is substituted for the eye, the instrument is a spectrophotometer. The experimenter may still have to take readings of the intensity of the light point by point across the spectrum. A more convenient system is one where the photomultiplier is coupled to a chart recorder, so that as the photomultiplier moves across the spectrum its response is automatically recorded on paper as a curve of the intensity of the light at each wavelength.

Recording the intensity of light point by point across the spectrum can take a good deal of time. Wavelengths are measured in angstrom units, and the visible spectrum is 3,000 angstroms wide. Therefore if a measurement every 100 angstroms is needed, it will take 31 measurements and at least that many minutes to cover the spectrum. Morelli's rapid-scan spectrophotometer does the job in a fraction of a second and so gives the experimenter the opportunity to investigate phenomena occurring much faster than the ones that can be investigated with slower instruments.

As an example of the potential of rapid scanning, Morelli told me how he examined the color oscillations of the chemical oscillator I described in this department for July, 1978. Although the period of the oscillations depends on the concentrations of the chemicals, it is usually less than a minute. With a conventional spectrophotometer one would have to be content with following the oscillations at a single wavelength, since the instrument would never be able to keep pace with the reactions.

Although the design for Morelli's instrument can be modified according to the experimenter's needs, it basically consists of a lamp, a curved diffraction grating, a photomultiplier tube and an oscilloscope. Light from the lamp is passed through a cell containing the sample that is to be investigated. The diffraction grating disperses the light into its component wavelengths. The grating is oscillated horizontally to sweep the visible spectrum across the photomultiplier tube, and the intensity of the light is displayed by the oscilloscope. The horizontal axis of the oscilloscope trace represents wavelength and the vertical axis intensity. One therefore sees a graph on which the intensity of the light is plotted against the wavelengths in the visible range. When the sample cell is absent, the oscilloscope shows a straight horizontal line. When the cell is in place, the resulting curve quickly identifies the wavelengths at which the sample has absorbed light.

For a lamp Morelli used an automobile light bulb (such as the bulb for a ceiling light) with a straight filament. The lamp is mounted in its housing with the filament oriented vertically. Although the lamp is rated for operation at 12 volts, Morelli runs it at about 21 volts in order to obtain the necessary intensity and uniformity across the visible spectrum. (The higher voltage does, of course, shorten the life of the lamp.)

The housing for the lamp is fashioned out of thin-gauge sheet steel cut and bent into shape. Morelli made the housing in two pieces so that he could replace the lamp more easily. The rear part of the housing is fastened to the main base of the spectrophotometer with a *U* bracket. The front part has attached to it a short tube through which the light passes. At the far end of the tube is a flat plate with two razor blades fastened to it by four screws. The tube and the razor blades serve as a baffle to keep stray light from entering the photomultiplier. More baffles are set up inside the instrument for the same purpose. The entire instrument is also covered with a box to keep out room light.

Morelli made the sample cell by cutting a *U*-shaped piece out of quarter-inch Plexiglas and sandwiching it between two glass plates. He glued the pieces together with silicone sealant. As a support for the cell he suggests cutting and bending sheet metal into a slide and screwing it onto a wood block. The cell is of course positioned in front of the light output from the lamp. In the normal procedure of measuring the absorption characteristics of a solution two cells are needed, one cell for the solvent alone so that the spectrophotometer can be standardized and one for the solution

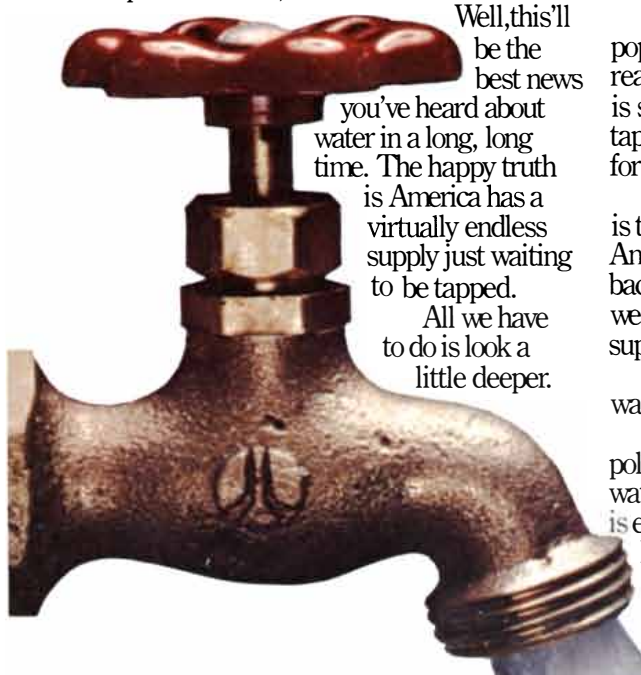
The Happy Truth About Th Tales Of The Unexpected

Droughts. Water shortages. Deserts encroaching on once-fertile land.

It's about time America had some good news about water. It's about time we found a supply that could take the burden off the country's overtaxed and often polluted lakes, rivers and reservoirs.

Well, this'll be the best news you've heard about water in a long, long time. The happy truth is America has a virtually endless supply just waiting to be tapped.

All we have to do is look a little deeper.



AMERICA'S BEST WATER IS ALWAYS GETTING UNDERFOOT.

The earth has as much water today as it had in prehistoric times. The water may sometime change forms, but the supply itself is constant. So where is all the water today?

An amazing percentage is where it's always been. Underground.

An incredible 95% of all the fresh water on earth is ground water. In the United States, there's 20 or 30 times as much water underground as there is in all the lakes, streams and rivers *combined.*

What's more, it's a completely renewable supply: Although the rate of replenishing varies greatly, water is seeping into the ground constantly, every hour of the day, every day of the year.

WELL WATER IS A NATURAL DELIGHT.

The American public is using more than 90 *billion* gallons of well water a day.

Approximately 700,000 new wells are drilled every year. Wells now provide almost 35 percent of the water used by municipalities—and over 80% of the water used in rural areas for homes and livestock.

Why is it so incredibly popular now? Part of the reason is the very fact that it is so plentiful—and that once tapped, the water is available for as long as it's sensibly used.

But part of the reason, too, is that well water is part of America's healthy drive to get back to its roots. After all, wells tap America's purest supply of natural water.

It's called the hydrologic cycle. It means that billions of gallons of water are entering underground reservoirs every day.

Does that mean this water's 100% pure? Absolutely not.

Although it rarely suffers from the pollution that so often afflicts surface water, nine times out of ten, well water is even better than totally pure water. It's enriched with dissolved natural minerals. Those minerals give well water in every part of the country its distinct, regional flavor. (If those minerals impart too strong a flavor, a simple filtering device can usually purify the water automatically.)

AMERICA'S MOST FERTILE FIELDS ARE ALL WET.

Many of the irrigation systems in the arid and semi-arid regions of the United States rely exclusively on huge underground reservoirs that have been collecting water for centuries. Without those rich water supplies, many of the most fertile fields in the country would soon become barren. In the 17 western states, for example, about 70 percent of the water used for irrigation comes from wells.

e Endless Well And Other From The Underground.

As long as that water isn't pumped faster than it's naturally replaced, there will always be plenty for the fields—and for the factories, homes and entire cities that rely on our underground resource.

WHY YOU SHOULD ALWAYS LOOK BELOW THE SURFACE.

Let's say you're looking for a new water supply. Maybe you're building a new home, factory or office building. Maybe you're planning a multi-acre farm or a quarter-acre garden. Maybe you're on the board of a regional or municipal water authority. Or maybe you'd just like the secure feeling that comes with knowing that no one can turn off or limit your water supply.

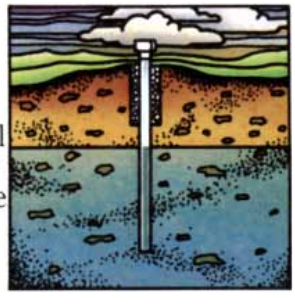
The first place to look is your own backyard.

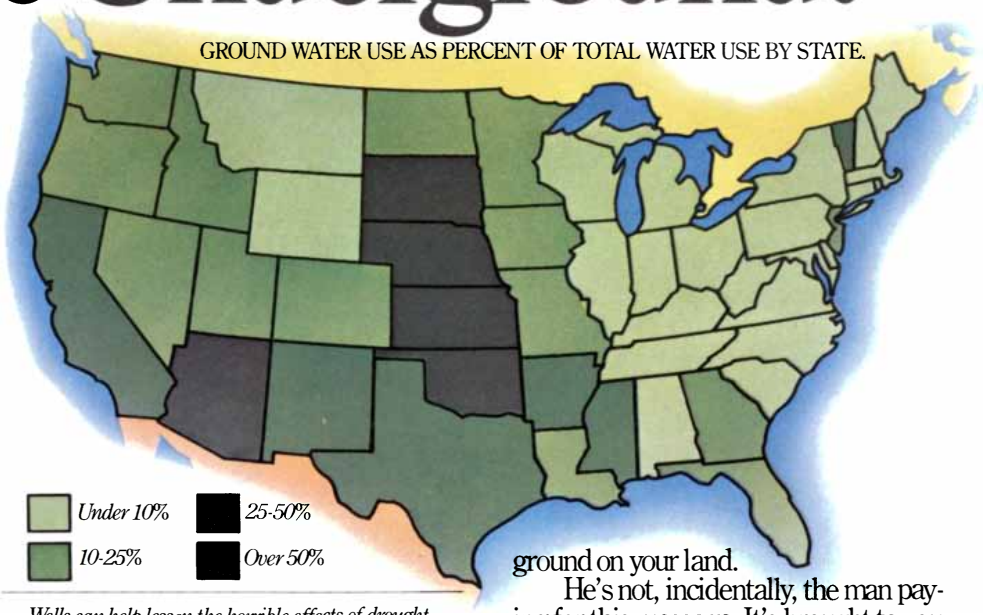
Ground water in some quantity can be found under almost every tract

America grows on ground water.

of inhabitable land on earth. The cost of getting that water is almost always lower than the cost of relying on any other water supply.

If you have a professional, experienced water well contractor inspect your site to determine the feasibility of using ground water, you'll be doing yourself and your locality a favor, because you'll be relying on a water storage system that doesn't drain the public reservoirs. A system that doesn't require major


There's water under almost every inhabitable tract of land on Earth.



Wells can help lessen the horrible effects of drought.

Bettman Archive

construction. A system that never becomes clogged with silt or weeds. One that never loses water by evaporation. One that doesn't remove productive land from other beneficial uses.

The fact is, ground water is a virtually untapped, unfailing resource. A priceless resource that's so vast, its potential has just begun to surface.

IT'S TIME YOU WENT TO THE PROS.

A professional water well contractor is a water specialist, an engineer and a craftsman all rolled into one. He knows the right techniques for finding the best site and building the best kind of well. He knows what equipment to use.

He's the man you should call to find out more about what's going on underground on your land.

He's not, incidentally, the man paying for this message. It's brought to you instead by Franklin Electric, the company that makes the motors for most of the submersible pump manufacturers.

We're doing it because we think it's time you knew the good news about one American resource that's not in trouble.

The story's been kept underground long enough.

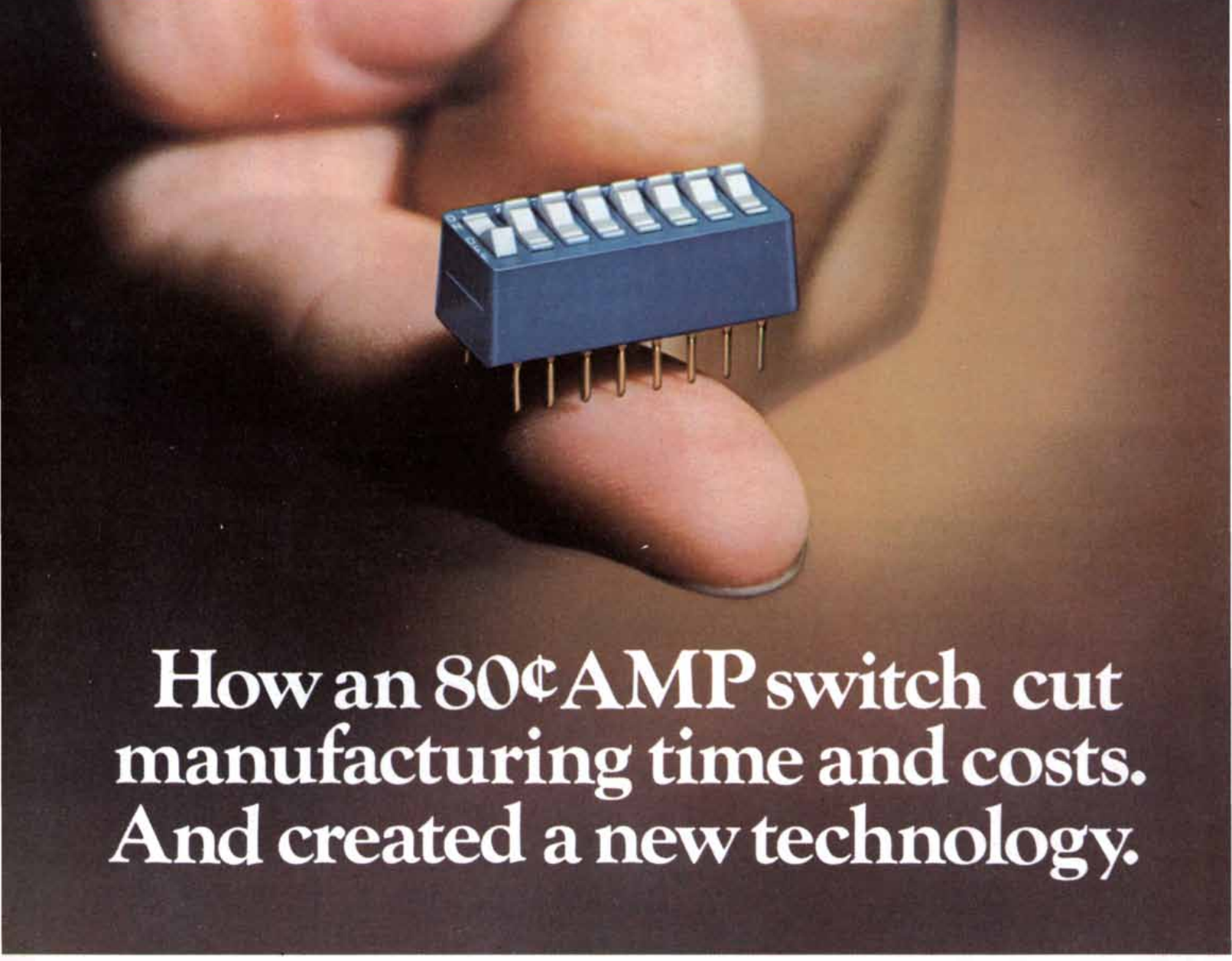
To find out more about well water, ground water contractors and ground water pumping techniques, just send this coupon to Franklin Electric, 400 E. Spring St., Bluffton, IN 46714. We'll work with the National Water Well Association to send you more good news from the underground.

Name _____

Address _____

City _____ State _____ Zip _____

Franklin Electric
Bluffton, Indiana 46714



How an 80¢ AMP switch cut manufacturing time and costs. And created a new technology.

Many companies don't call AMP until they need electrical connectors or switches. But why wait till then? If you involve us early enough, we can help you save money. A lot of money.

Here's an example:

Programming printed circuit boards used to involve installing an intricate network of jumper wires and plugs.

Until a customer asked us to find a better way.

We invented the DIP switch, which made programming p-c boards as easy as pushing number-coded actuators.

As a result, our customer was able to produce standardized boards that they could easily and quickly customize according to orders.

They saved hundreds of thousands of dollars in manufacturing and inventory costs. And a whole new technology was born.

Can we produce the same kind of savings for you?

If your company manufactures electrical/electronic products, the answer is probably yes. But it's important for you to call us in early.

That's when our experience in developing connectors, switches and application equipment for hundreds of industries worldwide can help you the most.

Early Involvement.

At AMP, we call this approach to solving our customers' problems "Early Involvement." It's our better way. And it's what makes an 80¢ AMP switch worth so much more than 80¢.

Ask for a copy of our brochure, "AMP Has a Better Way." Call (717) 564-0100, Ext. 8420.

Or write to AMP Incorporated, Harrisburg, PA 17105.

AMP

AMP has a better way.

Barcelona • Brussels • Buenos Aires • Frankfurt • Harrisburg • Helsinki • s-Hertogenbosch • London • Luzern • Mexico City • Montreal
Paris • San Juan • Sao Paulo • Stockholm • Sydney • Turin • Toronto • Tokyo

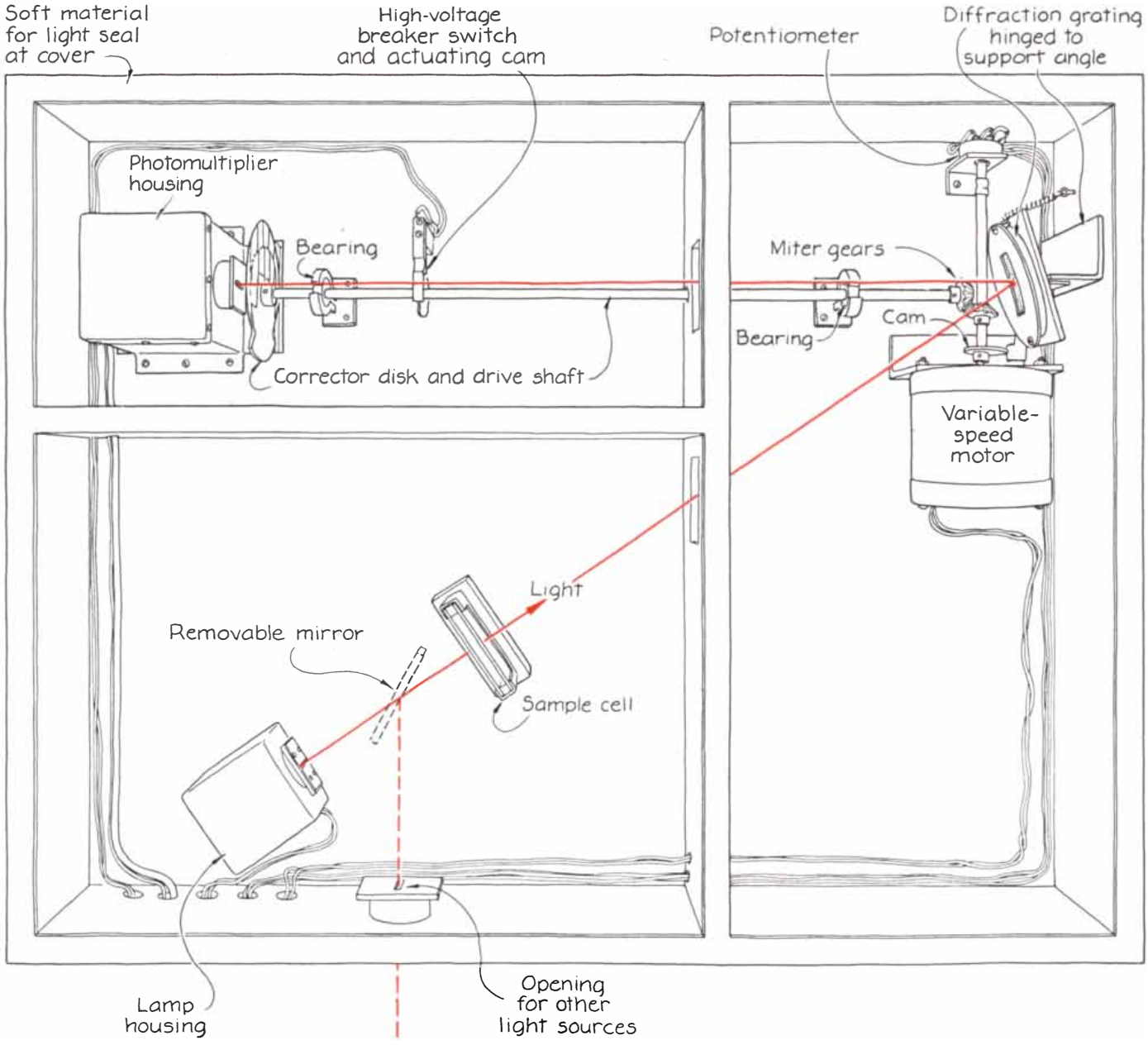
that is to be compared with the standard.

The diffraction grating disperses the light transmitted through the sample cell. Morelli chose a curved diffraction grating of the kind found in certain astronomical spectroscopes. (For retailers of the gratings check the pages of an astronomy magazine.) Morelli's grating has a radius of curvature of 50 centimeters, a ruled area of 3.175 by 4.445 centimeters and a line density of 1,200 lines per millimeter. The lines are spaced so that light with a wavelength of 5,000 angstroms has its maximum reflection in the first order of the diffraction pattern. (Such a grating is said to be blazed at 5,000 angstroms.) Other curved gratings could be substituted, although Morelli warns that gratings with a higher *f* number will make the instrument less sensitive. If a flat grating is used, a lens will be needed in order to project the image of the slit of the lamp housing onto the slit of the photomultiplier (by way of reflection from the grating).

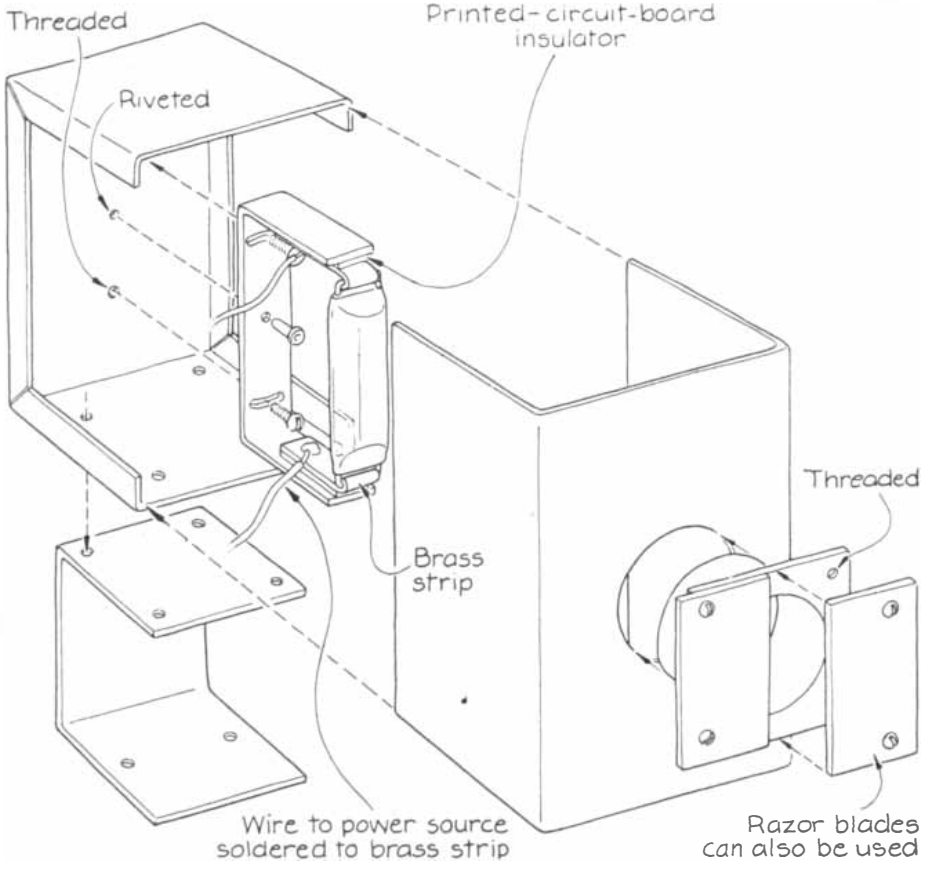
Morelli's grating is mounted on an aluminum plate to which a hinge is riveted. The idea is to have the grating swing horizontally around the hinge in order to sweep the spectrum across the photomultiplier. The hinge is attached to an aluminum angle that is fastened to the main base of the spectrophotometer. The hinge should turn smoothly with no free play, but it should not be so firm that it offers too much resistance to being swung. Morelli says he tried several hinges before he found a suitable one.

For the highest possible scanning rate the grating should be rigged to oscillate around its center. This pivot point will not, however, provide the best focusing for all the wavelengths in the visible spectrum, because the best focusing distance for a grating varies with wavelength. For the optimum optical performance the pivot point should be offset by about six centimeters from the center of the grating. For such a pivot point the scanning rate is lower (because the grating has a greater moment of inertia than it would have with a central pivot point), but the improvement in focusing is worth the sacrifice.

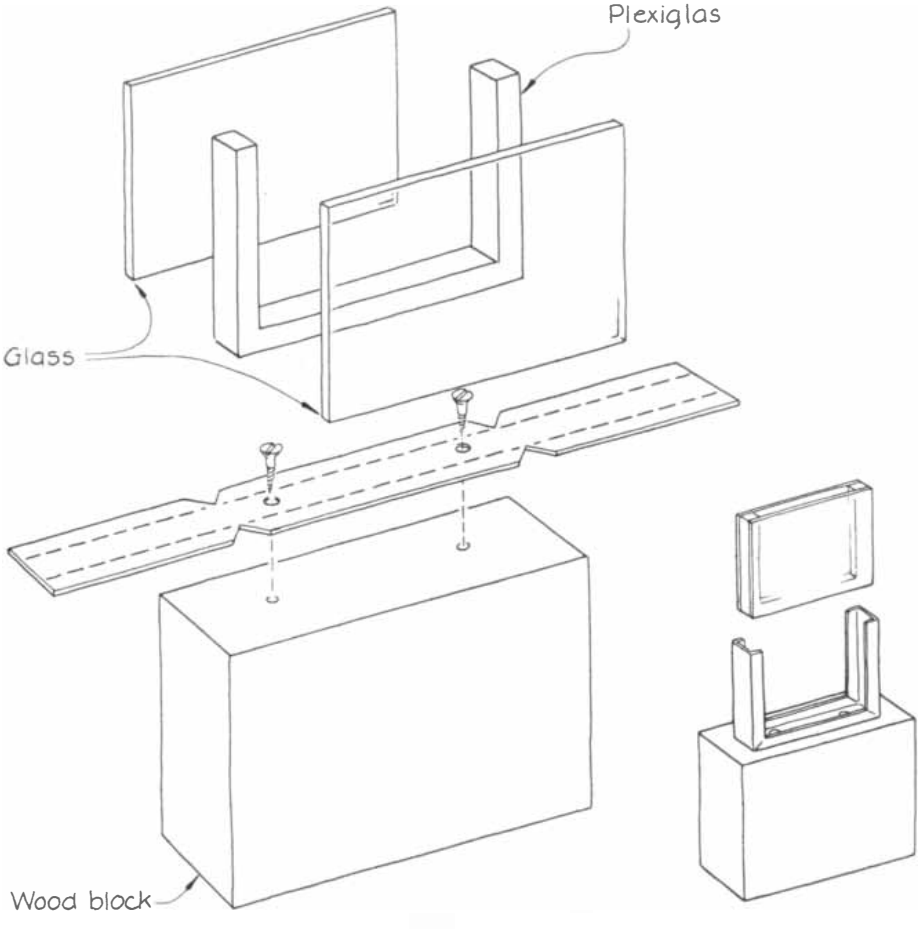
The grating is made to sweep by means of a cam against which a small tab from the mounting plate of the grating is held. The tab is cut from right-angle stock and firmly attached to the plate. One end of the tab is held against the cam by the tension from a spring attached to the plate. When the cam turns, the tab slides along the edge of it



The spectrophotometer designed and built by Dean Morelli



The housing for the lamp



The sample cell and its holder

and forces the grating to sweep horizontally. The spring ensures that the tab is held against the cam.

The cam is cut from 3/32-inch sheet brass. Its shape and size are first mapped out on paper. The paper mask is laid on the metal sheet and traced, after which the metal is cut and filed smooth around its perimeter. Ideally the cam should sweep the grating in a sawtooth motion, uniformly driving it horizontally until the end of the sweep is reached. This motion is obtained when the radius of the cam varies uniformly with the angle around its center. The center hole for the cam is drilled with a quarter-inch bit to fit on a motor shaft of the same diameter. Morelli soldered onto the cam a brass insert he had removed from a panel knob (Radio Shack Knob 274–415). The insert is normally inside the plastic covering. The screw into the insert enabled Morelli to attach the insert and the cam firmly to the motor shaft.

The upper limit to the sweep rate of the spectrophotometer is set by the oscillation rate of the grating. If the grating is run too fast, the spring cannot hold the tab against the cam and the sweep becomes erratic. Substituting a stronger spring may not improve the performance if the resulting increase in friction between the cam and the tab distorts the running speed of the motor.

Opposite the grating on the base of the spectrophotometer is the photomultiplier and its housing. The housing is made from thin-gauge sheet steel. A steel tube is fastened to the front of the housing with a weld or epoxy glue. Inside this tube is slid a smaller tube. Razor blades are attached to the smaller tube to make a narrow slit. Since the smaller tube is easily removed, the experimenter can readily install slits of different widths.

The photomultiplier tube is mounted on a shelf in the housing. The resistors necessary to provide the proper range of voltage to the photomultiplier are soldered directly onto the socket on the shelf. Like all the other exposed metal parts in the spectrophotometer, the housing should be painted a dull black to absorb stray light.

Morelli has tried several types of photomultiplier tube. The one you would buy may depend primarily on what is to be had on the surplus market. Model 931-A is often available. Morelli also has a 1P21, which is the same type of tube except that it has a higher quantum efficiency and a lower dark current. (I recently saw an advertisement offering two of these tubes for $75.) The quantum efficiency has to do with how many electrons the photocathode of the tube will emit for each photon it absorbs. If the ratio is fairly high, the photomultiplier is said to be sensitive and to have a high quantum efficiency. Dark current is the current that flows from the tube

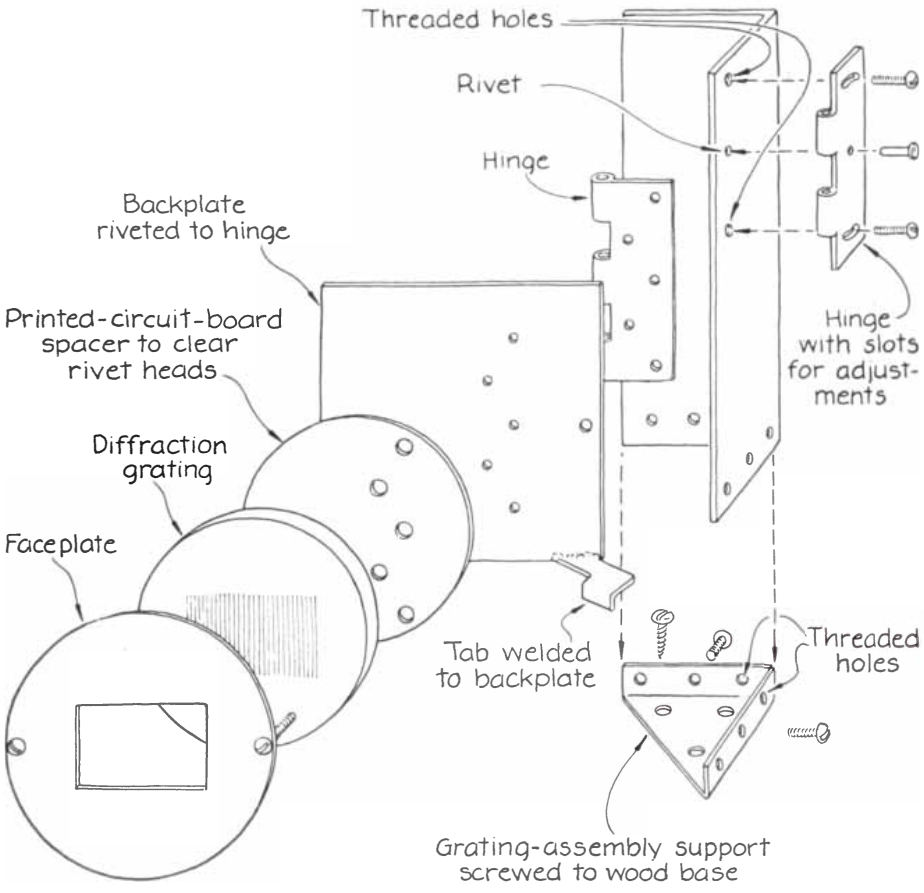
when no light is falling on the photocathode. It should be low so that it does not mask the desired signal and prevent the measurement of low levels of light. The 931-A and 1P21 tubes both have the type of photoelectric surface designated S-4; it is composed of cesium and antimony. Although the tubes have a relatively high quantum efficiency over the greater part of the visible spectrum, they are somewhat insensitive toward the red end, reaching a practical limit near 6,700 angstroms.

A photomultiplier tube must be biased properly, that is, it must have a suitable range of voltage. Morelli employs a regulated power supply to provide about 100 volts per stage for each of the nine stages in his 931-A photomultiplier. One should consult *Radio Amateur's Handbook* for information on power-supply designs.

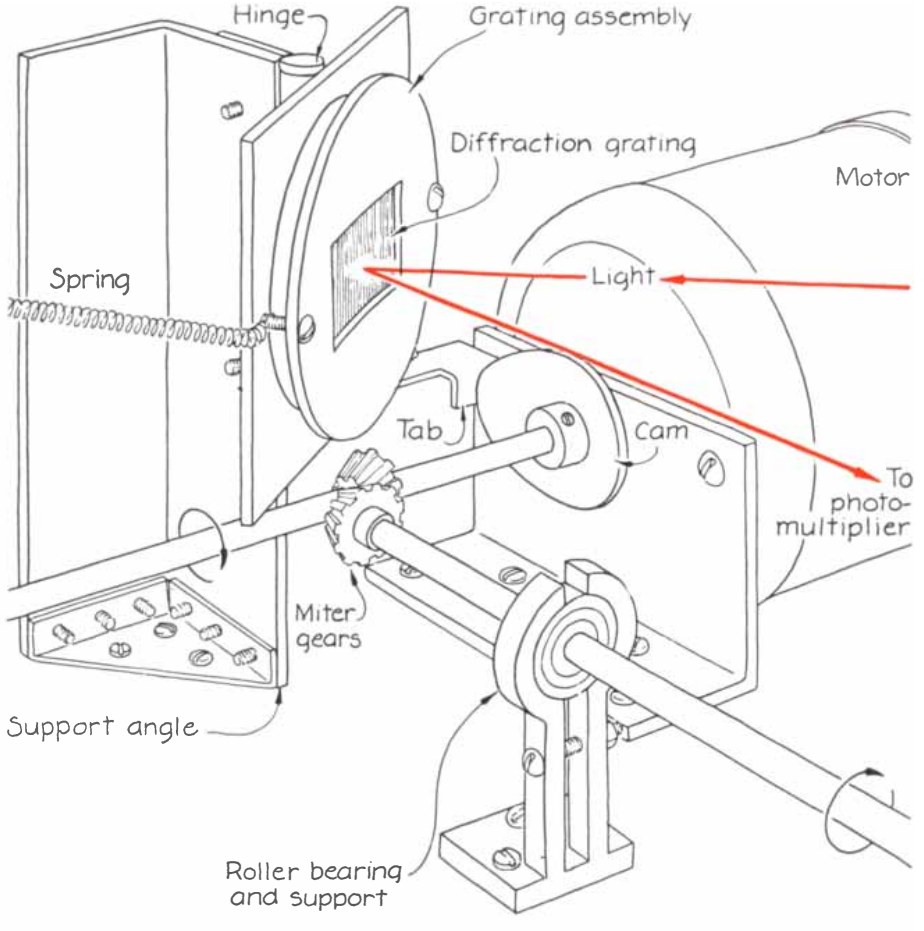
Recently Morelli tested an RCA 4840 photomultiplier. Its multialkali photoelectric surface (designated S-20) had a spectral range larger than that of his 1P21 tube, being sensitive to wavelengths as long as 8,000 angstroms. The tube also has an envelope that passes ultraviolet wavelengths. With this tube Morelli was able to get full-scale vertical deflections on the oscilloscope as he measured the intensity of the light in the range from 3,400 angstroms to 7,000.

The output from the photomultiplier is connected to the vertical input of the oscilloscope, which should have a vertical amplifier for direct current. The oscilloscope trace should move across the screen from left to right as the grating sweeps the spectrum over the entrance slit of the photomultiplier. How the oscilloscope is triggered to begin each sweep across the screen depends somewhat on the type of oscilloscope. If it has an internal provision for a triggered sweep, the triggering signal can be provided by a small light, a slit disk and a silicon photocell. The disk is mounted on an extension of the shaft of the motor that oscillates the grating. The light is on one side of the disk and the photocell is on the other. When the slit in the disk is properly positioned, the light from the small bulb falls on the photocell each time the grating begins another sweep of the spectrum across the photomultiplier. The signal from the photocell then starts the sweep of the oscilloscope trace. (Morelli cautions that inexpensive oscilloscopes do not trigger reliably at the low sweep frequencies of the spectrophotometer.)

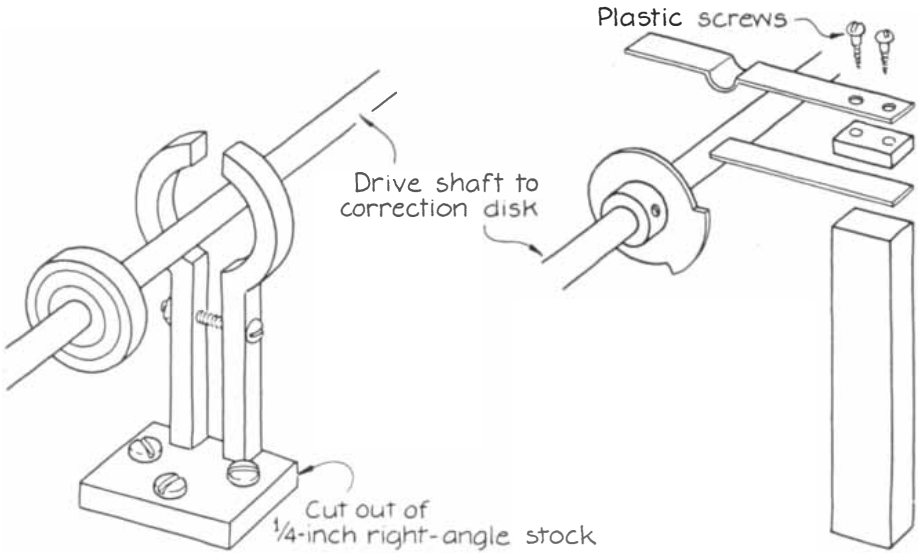
A better system for displaying the output of the photomultiplier can be set up if the oscilloscope has a horizontal amplifier for direct current. A small projector is mounted in a metal tube so that it projects an image of a slit onto a silicon photocell (SM4) on the other side of an appropriately shaped cam. The



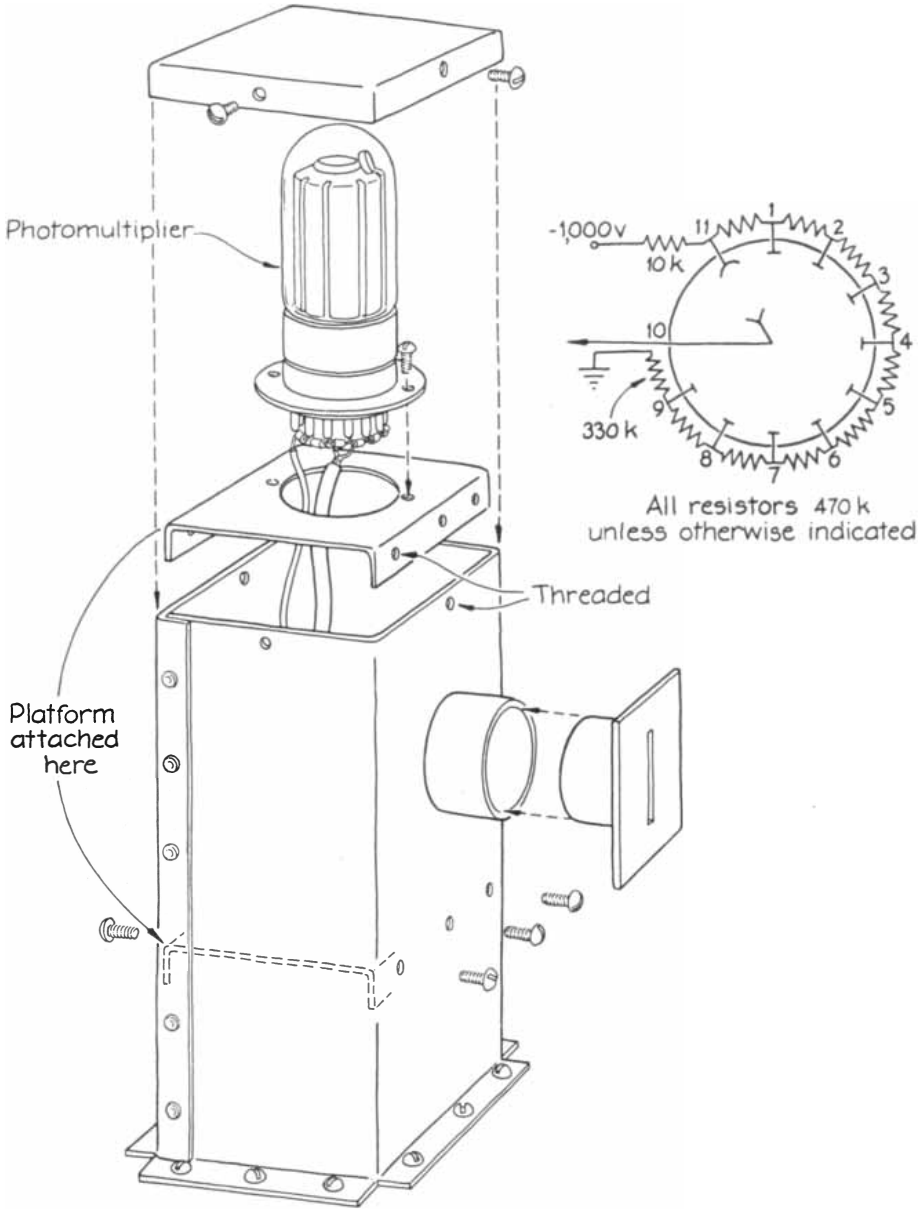
Details of the grating assembly



How the grating is mounted



Arrangement of the high-voltage breaker switch



The photomultiplier and its housing

cam is similar to the one that drives the grating. When this cam turns, it causes progressively more light to fall on the photocell. The signal from the photocell is amplified at the horizontal input of the oscilloscope, sending the trace from left to right across the screen. The second cam could be mounted on an extension of the shaft of the grating's motor to synchronize the horizontal sweep of the oscilloscope with the movement of the grating.

If the amplifier for the horizontal input of the oscilloscope is not sensitive enough for this system, the projector and the photocell can be replaced with a potentiometer, which will supply a larger signal to the input. In this role the potentiometer must turn continuously, so that the small indentation in its housing that keeps it from rotating fully must be removed. The potentiometer is mounted on the shaft of the motor with a short length of rubber tubing. A 1.5-volt battery provides the voltage across the potentiometer, the output of which then varies continuously from zero to 1.5 volts in a sawtooth fashion as the shaft causes the inner connection of the potentiometer to rotate. The signal will drive the oscilloscope trace smoothly across the screen.

One may want to eliminate the return trace of the oscilloscope, which moves across the screen from right to left before each new measurement of intensity. To blank out the return trace Morelli put a switch on the drive shaft between the mount for the grating and the corrector plate for the photomultiplier. The switch consists of two metal strips that are pushed apart by a cam on the drive shaft during the period when the oscilloscope trace is to be blanked out. The switch is in series with the wire supplying the high voltage to the oscilloscope. One may, however, want to keep the return trace to serve as a reference line on the screen indicating the base line of zero light intensity. Such a base line may be particularly desirable if the oscilloscope has an unregulated power supply and so has a trace that tends to float.

If the oscilloscope has low sensitivity on its vertical input, a preamplifier is needed to boost the signal from the photomultiplier. Morelli employs an inexpensive uA741 operational amplifier chip in a simple feedback circuit powered by penlight batteries (enough to give ± 15 volts).

Morelli uses his instrument as follows to determine the absorption characteristics of a solution compared with a pure solvent. He first gets an oscilloscope trace with the pure solvent in the sample cell. Even if the solvent transmits perfectly, the trace will not be flat, because the intensity of the light emitted by the lamp is not uniform at all wavelengths across the visible spectrum. Moreover,

the response of the photomultiplier is not uniform across the spectrum. Hence the trace on the oscilloscope is approximately bell-shaped, with the top of the bell disappearing off the top of the screen.

A corrector disk is made to compensate for these nonuniformities (and any others that distort the response of the instrument across the spectrum). The disk is mounted on a shaft coupled through miter gears to the shaft from the motor that causes the grating to oscillate. The purpose of the disk is to selectively diminish the light falling on the photomultiplier at the wavelengths that give a higher trace on the oscilloscope. A suitably fashioned corrector disk will block the light at each wavelength in the spectrum in such a way that the trace is approximately horizontal across the top of the screen. With a pure solvent one then gets a flat trace on the oscilloscope.

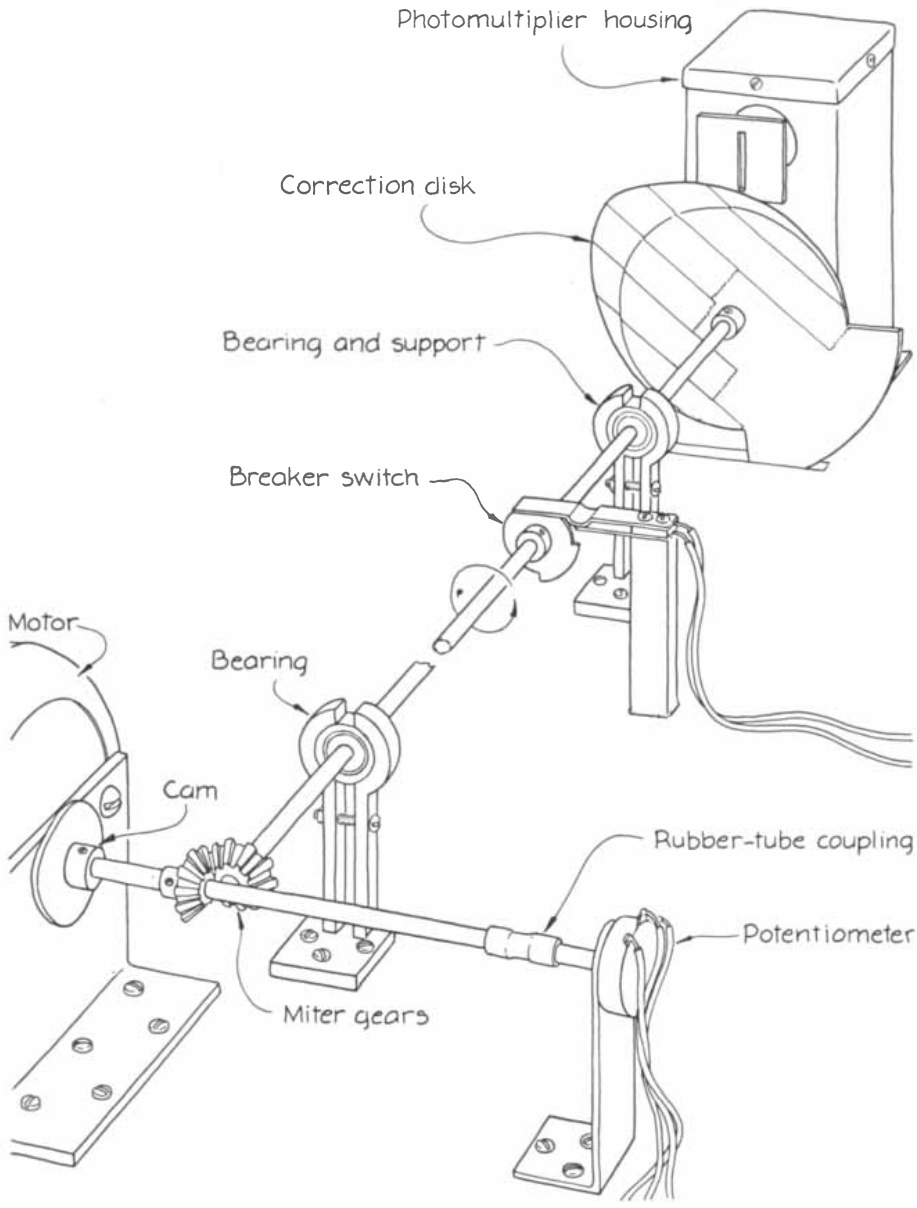
Morelli cuts the corrector disk out of a piece of sheet brass. When the disk is mounted on the drive shaft, it blocks part of the light falling on the slit in front of the photomultiplier. Next Morelli places pieces of black electrical tape at the edge of the disk as he manually turns the shaft of the motor. For each position of the motor the grating casts a certain wavelength onto the slit, and the vertical deflection of the spot on the oscilloscope screen corresponds to the response of the photomultiplier at that wavelength. At each wavelength the spot is initially deflected off the top of the screen. Morelli adds a small piece of tape to the corrector disk to block enough light for the spot to be lowered to the top of the screen. Then he rotates the shaft on the motor slightly and repeats the process. He says that although the system is somewhat crude, the resulting trace across the top of the screen is flat to within 5 percent.

When the cell containing only the solvent is replaced with a cell containing a substance dissolved in the same solvent, the absorption of light by the solution at certain wavelengths gives rise to spikes that extend downward on the oscilloscope screen. If the solution were able to totally absorb the light at a certain wavelength, the corresponding spike would reach the bottom of the screen.

The resolution of the spectrophotometer is governed partly by the quality of the grating and partly by the width of the slit in front of the photomultiplier. A narrower slit samples a narrower band of wavelengths in the spectrum the grating casts across the slit. At the position of the slit the width of the full visible spectrum is 125 millimeters. Morelli normally uses a slit with a width of about four millimeters, which therefore admits 4/125 (roughly 100 angstroms) of the full spectrum to the photomultiplier. One can increase the resolution by replacing the four-millimeter slit with a narrower one, but a price must be paid in a loss of light intensity. Then the photomultiplier will respond only to light at wavelengths in its more sensitive range (for Morelli's 931-A tube approximately 3,700 angstroms to 6,000) and the red end of the spectrum will be lost. With a four-millimeter slit a larger range (3,500 angstroms to 6,300) can be monitored. The lower limit to the range is then set by the ultraviolet absorption of the glass in the lamp, the photomultiplier tube and the sample cell. Another problem with slits narrower than four millimeters is that they may call for off-center pivoting of the grating to prevent defocusing at the ends of the spectrum.

Morelli calibrates his spectrophotometer by putting a hydrogen Geissler emission tube in place of the usual lamp. Several of the visible spectral lines emitted by hydrogen (the Balmer series) can be identified on the oscilloscope screen. When the four-millimeter slit is in place, the lines appear as sharp spikes. Since the wavelengths of the lines are known precisely, Morelli can calibrate the horizontal axis of the trace on the screen in angstroms.

I have purposely not been precise about the dimensions of the spectrophotometer because Morelli emphasizes the need for the experimenter to modify and adjust the basic design to suit whatever equipment he has gathered for building the instrument. The housings for the photomultiplier and the lamp should be as close together as possible. Morelli suggests a way to determine the proper distances of the lamp, the photomultiplier and the grating. Set up on a table a Geissler tube, the grating and a slit. With the room lights off and the tube on move



Assembly of the scanning mechanism

the three objects on the table around until the spectrum falls on the slit in good focus. The distances so determined can serve as a guide for the construction of the spectrophotometer.

To use the spectrophotometer for examining the spectral characteristics of external sources of light, such as a lamp or a flame, a mirror is mounted in front of the lamp in the instrument. Light from the external source is admitted through a slit in the side of the box housing the spectrophotometer and is directed by the mirror to the grating. The source must be relatively bright or the red end of the spectrum will be lost because of the insensitivity of the photomultiplier at that end.

The possible applications of Morelli's rapid-scan spectrophotometer are numerous. In addition to the oscillating chemical reactions I have mentioned, there are almost countless other optical phenomena that change too rapidly for any spectrophotometer requiring a point-by-point measurement across the visible spectrum. Morelli himself, for example, wants to examine the phosphorescence of certain crystals.

Some amateur astronomers may be interested in analyzing the spectral composition of sunlight. Morelli employs a telescope and a mirror to project an image of the sun onto a pinhole that he has installed temporarily in place of the slit in the side of the box housing the spectrophotometer. He moves the image of the sun across the pinhole and thereby gets tracings on his oscilloscope for spectra of different areas on the sun. Similar observations could be made of other celestial objects, but because of the much lower intensity of their light the spectrophotometer of Morelli's basic design would respond primarily to the wavelengths at which the photomultiplier is the most sensitive.

Morelli has not stopped tinkering with his instrument. His next big change will be to replace the mechanical scanning system of the present design with an entirely electronic scanning system based on a Vidicon (an image-storage tube). Such a system will be more sensitive because the oscilloscope trace will be built up through the storage of information from many scans. It would therefore be better for the spectral analysis of light from faint sources such as phosphorescing crystals. With the electronic system Morelli will also be able to examine ultrarapid phenomena such as the flash from a high-speed xenon flash lamp. The Vidicon is also more sensitive at the red end of the spectrum, so that the instrument will not have one of the shortcomings of an instrument based on a photomultiplier. Morelli is now building a rapid-scan spectrophotometer with a Vidicon tube that is employed in closed-circuit black-and-white television cameras.

WHAT'S BETTER THAN SPEED READING?

SPEED LEARNING

(SPEED PLUS COMPREHENSION)

Speed Learning is replacing speed reading because it's easy to learn . . . lasts a lifetime . . . applies to everything you read . . . and is the only fully accredited course with the option of college credits.

Do you have too much to read and too little time to read it? Do you mentally pronounce each word as you read? Do you frequently have to go back and re-read words or whole paragraphs you just finished reading? Do you have trouble concentrating? Do you quickly forget most of what you read?

If you answer "yes" to any of these questions — then here at last is the practical help you've been waiting for. Whether you read for business or pleasure, school or college, you will build exceptional skills from this major breakthrough in effective reading, created by Dr. Russell Stauffer at the University of Delaware.

Not just "speed reading" — but speed reading-thinking-understanding-remembering-and-learning

The new *Speed Learning Program* shows you step-by-proven-step how to increase your reading skill and speed, so you understand more, remember more and use more of everything you read. The typical remark made by the 75,000 slow readers who completed the *Speed Learning Program* was: "Why didn't someone teach me this a long time ago?" They were no longer held back by the lack of skills and poor reading habits. They could read almost as fast as they could think.

What makes Speed Learning so successful?

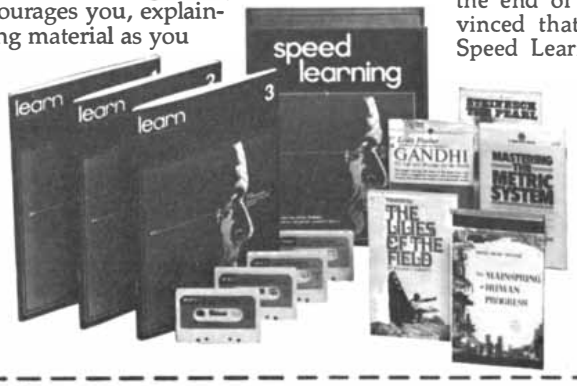
The new *Speed Learning Program* does not offer you a rehash of the usual eye-exercises, timing devices, costly gadgets you've probably heard about in connection with speed reading courses or even tried and found ineffective.

In just a few spare minutes a day of easy reading and exciting listening, you discover an entirely new way to read and think — a radical departure from anything you have ever seen or heard about. Research shows that reading is 95% *thinking* and only 5% eye movement. Yet most of today's speed reading programs spend their time teaching you rapid eye movement (5% of the problem) and ignore the most important part (95%) *thinking*. In brief, *Speed Learning* gives you what speed reading *can't*.

Imagine the new freedom you'll have when you learn how to dash through all types of reading material *at least* twice as fast as you do now, and with greater comprehension. Think of being able to get on top of the avalanche of newspapers, magazines and correspondence you have to read . . . finishing a stimulating book and retaining facts and details more clearly and with greater accuracy than ever before.

Listen-and-learn at your own pace

This is a practical, easy-to-learn program that will work for you — no matter how slow a reader you think you are now. The *Speed Learning Program* is scientifically planned to get you started quickly . . . to help you in spare minutes a day. It brings you a "teacher-on-cassettes" who guides you, instructs, encourages you, explaining material as you read. Interesting items taken from *Time* Magazine, *Business Week, Wall Street Journal, Family Circle, N.Y. Times* and many others, make the program stimulating, easy and fun . . . and so much more effective.



Executives, students, professional people, men and women in all walks of life from 15 to 70 have benefited from this program. Speed Learning is a fully accredited course . . . costing only 1/5 the price of less effective speed reading classroom courses. Now you can examine the same, easy, practical and proven methods at home . . . in spare time . . . without risking a penny.

Examine Speed Learning FREE for 10 days

You will be thrilled at how quickly this program will begin to develop new thinking and reading skills. After listening to just one cassette and reading the preface you will quickly see how you can achieve increases in both the speed at which you read and in the amount you understand and remember.

You must be delighted with what you see or you pay nothing. Examine this remarkable program for 10 days. If, at the end of that time you are not convinced that you would like to master Speed Learning, simply return the program and owe nothing. See the coupon for low price and convenient credit terms.

Note: Many companies and government agencies have tuition assistance plans for employees providing full or partial payment for college credit programs.

In most cases, the entire cost of your Speed Learning Program is Tax Deductible.

EARN PROFESSIONAL & COLLEGE CREDITS

Speed Learning is approved for credit by the following professional and educational institutions:

- **Foundation for Accounting Education** 20 CPE Credit Hours
- **American Management Association** 1.9 Continuing Education Units
- **National Society of Public Accountants** 20 Continuing Education Hours
- **College Credit** 3 credits from the National College of Business

Details and registration forms included with each program.

OFFERED INTERNATIONALLY BY

- **Institute of Electrical & Electronics Engineers**
- **American Chemical Society**
- **National Association of Life Underwriters**

learn INCORPORATED 113 Gaither Drive, Mount Laurel, N.J. 08054

31SAP

Please send me the Speed Learning Program at $89.95 plus $3.00 for handling and insured delivery.

Please check the method of payment below:

☐ Check or money order enclosed for items ordered. New Jersey Residents add 5% sales tax.

☐ Please charge my credit card under the regular payment terms: ☐ Bank Americard ☐ Master Charge Interbank No. _____ ☐ American Express ☐ Diners Club

Card No. _____ Exp. Date _____

I understand that if after 10 days I am not delighted in every way, I may return the materials and obtain a full refund with no questions asked.

Name _____

Address _____

City _____ State _____ Zip _____

x Signature _____

If you don't already own a cassette player, you may order this Deluxe Cassette Recorder for only $49.95. (Includes handling and delivery.)

Check here to order ☐

Outside U.S.A. $99.95 + $3 surface mail – airmail extra

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

MATHEMATICAL GAMES

SOME STUDIES IN MACHINE LEARNING USING THE GAME OF CHECKERS. A. L. Samuel in *IBM Journal of Research and Development,* Vol. 3, No. 3, pages 210–229; July, 1959.

SOME STUDIES IN MACHINE LEARNING USING THE GAME OF CHECKERS, II: RECENT PROGRESS. A. L. Samuel in *IBM Journal of Research and Development,* Vol. 11, No. 6, pages 601–617; November, 1967.

THE COMPLEXITY OF CHECKERS ON AN $N \times N$ BOARD—PRELIMINARY REPORT. A. S. Fraenkel, M. R. Garey, D. S. Johnson and Y. Yesha in *19th Annual Symposium on the Foundations of Computer Science.* Institute of Electrical and Electronics Engineers, 1978.

THE NEXT GENERATION OF PARTICLE ACCELERATORS

THE BATAVIA ACCELERATOR. R. R. Wilson in *Scientific American,* Vol. 230, No. 2, pages 72–83; February, 1974.

PROTON-PROTON INTERSECTING STORAGE ACCELERATOR FACILITY ISABELLE AT THE BROOKHAVEN NATIONAL LABORATORY. H. Hahn, M. Month and R. R. Rau in *Reviews of Modern Physics,* Vol. 49, No. 3, pages 625–679; July, 1977.

DESIGN STUDY OF A 22 TO 130 GeV e^+e^- COLLIDING BEAM MACHINE (LEP). The LEP Study Group, European Organization for Nuclear Research, 1979.

FANTASIES OF FUTURE FERMILAB FACILITIES. R. R. Wilson in *Reviews of Modern Physics,* Vol. 51, No. 2, pages 259–273; April, 1979.

THE CELL BIOLOGY OF HUMAN AGING

THE STRATEGY OF SENESCENCE. Leonard Hayflick in *The Gerontologist,* Vol. 14, No. 1, pages 37–45; February, 1974.

CURRENT THEORIES OF BIOLOGICAL AGING. Leonard Hayflick in *Federation Proceedings,* Vol. 34, No. 1, pages 9–13; January, 1975.

THE CELL BIOLOGY OF HUMAN AGING. Leonard Hayflick in *The New England Journal of Medicine,* Vol. 295, No. 23, pages 1302–1308; December 2, 1976.

HANDBOOK OF THE BIOLOGY OF AGING. Edited by Caleb E. Finch and Leonard Hayflick. Van Nostrand Reinhold Co., 1977.

TIME, CELLS, AND AGING. Bernard L. Strehler. Academic Press, 1977.

THE BIOLOGY OF SENESCENCE. Alex Comfort. Elsevier North-Holland, Inc., 1979.

WORLD URANIUM RESOURCES

THE ENERGY RESOURCES OF THE EARTH. M. King Hubbert in *Scientific American,* Vol. 224, No. 3, pages 60–70; September, 1971.

A SECOND IRON AGE AHEAD? Brian J. Skinner in *American Scientist,* Vol. 64, No. 3, pages 258–269; May–June, 1976.

SHORT COURSE IN URANIUM DEPOSITS: THEIR MINERALOGY AND ORIGIN. Edited by M. M. Kimberley, Mineralogical Association of Canada. University of Toronto Press, 1978.

URANIUM DISTRIBUTION IN MINED DEPOSITS AND IN THE EARTH'S CRUST. Kenneth Deffeyes and Ian MacGregor. U.S. Department of Energy, Grand Junction, Colo., August, 1978.

THE GALILEAN MOONS OF JUPITER

MELTING OF IO BY TIDAL DISSIPATION. S. J. Peale, P. Cassen and R. T. Reynolds in *Science,* Vol. 203, No. 4383, pages 892–894; March 2, 1979.

INFRARED OBSERVATIONS OF THE JOVIAN SYSTEM FROM VOYAGER 1. R. Hanel, B. Conrath, M. Fraser, V. Kunde, P. Lowman, W. Maguire, J. Pearl, J. Pirraglia, R. Samuelson, D. Gautier, P. Gierasch, S. Kumar and C. Ponnamperuma in *Science,* Vol. 204, No. 4396, pages 972–976; June 1, 1979.

THE JUPITER SYSTEM THROUGH THE EYES OF VOYAGER 1. Bradford A. Smith, Laurence A. Soderblom, Torrence V. Johnson, Andrew P. Ingersoll, Stewart A. Collins, Eugene M. Shoemaker, G. E. Hunt, Harold Masursky, Michael H. Carr, Merton E. Davies, Allan F. Cook II, Joseph Boyce, G. Edward Danielson, Tobias Owen, Carl Sagan, Reta F. Beebe, Joseph Veverka, Robert G. Strom, John F. McCauley, David Morrison, Geoffrey A. Briggs and Verner E. Suomi in *Science,* Vol. 204, No. 4396, pages 951–972; June 1, 1979.

VOYAGER 1 ENCOUNTER WITH THE JOVIAN SYSTEM. E. C. Stone and A. L. Lane in *Science,* Vol. 204, No. 4396, pages 945–948; June 1, 1979.

GEOMETRICAL ILLUSIONS

A DEPTH PROCESSING THEORY OF THE POGGENDORFF ILLUSION. B. Gillam in *Perception and Psychophysics,* Vol. 10, No. 4A, pages 211–216; 1971.

THE PSYCHOLOGY OF VISUAL ILLUSION. J. O. Robinson. Hutchinson University Library, 1972.

ILLUSION IN NATURE AND ART. Edited by R. L. Gregory and E. H. Gombrich. Charles Scribner's Sons, 1973.

SEEING IS DECEIVING: THE PSYCHOLOGY OF VISUAL ILLUSIONS. Stanley Coren and Joan S. Girgus. Laurence Erlbaum Associates, 1978.

THE ANCESTRY OF CORN

THE ORIGIN OF CORN. Walton C. Galinat in *Annual Review of Genetics,* Vol. 5, pages 447–478; 1971.

CORN: ITS ORIGIN, EVOLUTION AND IMPROVEMENT. Paul C. Mangelsdorf. Harvard University Press, 1974.

ZEA DIPLOPERENNIS (GRAMINEAE): A NEW TEOSINTE FROM MEXICO. Hugh H. Iltis, John F. Doebley, Rafael Guzmán and Batia Pazy in *Science,* Vol. 203, No. 4376, pages 186–188; January 12, 1979.

RABIES

LABORATORY TECHNIQUES IN RABIES. Edited by Martin M. Kaplan and Hilary Koprowski. World Health Organization, 1973.

WHO EXPERT COMMITTEE ON RABIES, SIXTH REPORT. World Health Organization, 1973.

THE NATURAL HISTORY OF RABIES. Edited by George M. Baer. Academic Press, 1975.

SUCCESSFUL PROTECTION OF HUMANS EXPOSED TO RABIES INFECTION. Mahmoud Bahmanyar, Ahmad Fayaz, Shokrollah Nour-Salehi, Manouchehr Mohammadi and Hilary Koprowski in *The Journal of the American Medical Association,* Vol. 236, No. 24, pages 2751–2754; December 13, 1976.

MONOCLONAL ANTIBODIES AGAINST RABIES VIRUS PRODUCED BY SOMATIC CELL HYBRIDIZATION: DETECTION OF ANTIGENIC VARIANTS. T. J. Wiktor and H. Koprowski in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 75, No. 8, pages 3938–3942; August, 1978.

THE ORIGINS OF THE WATER TURBINE

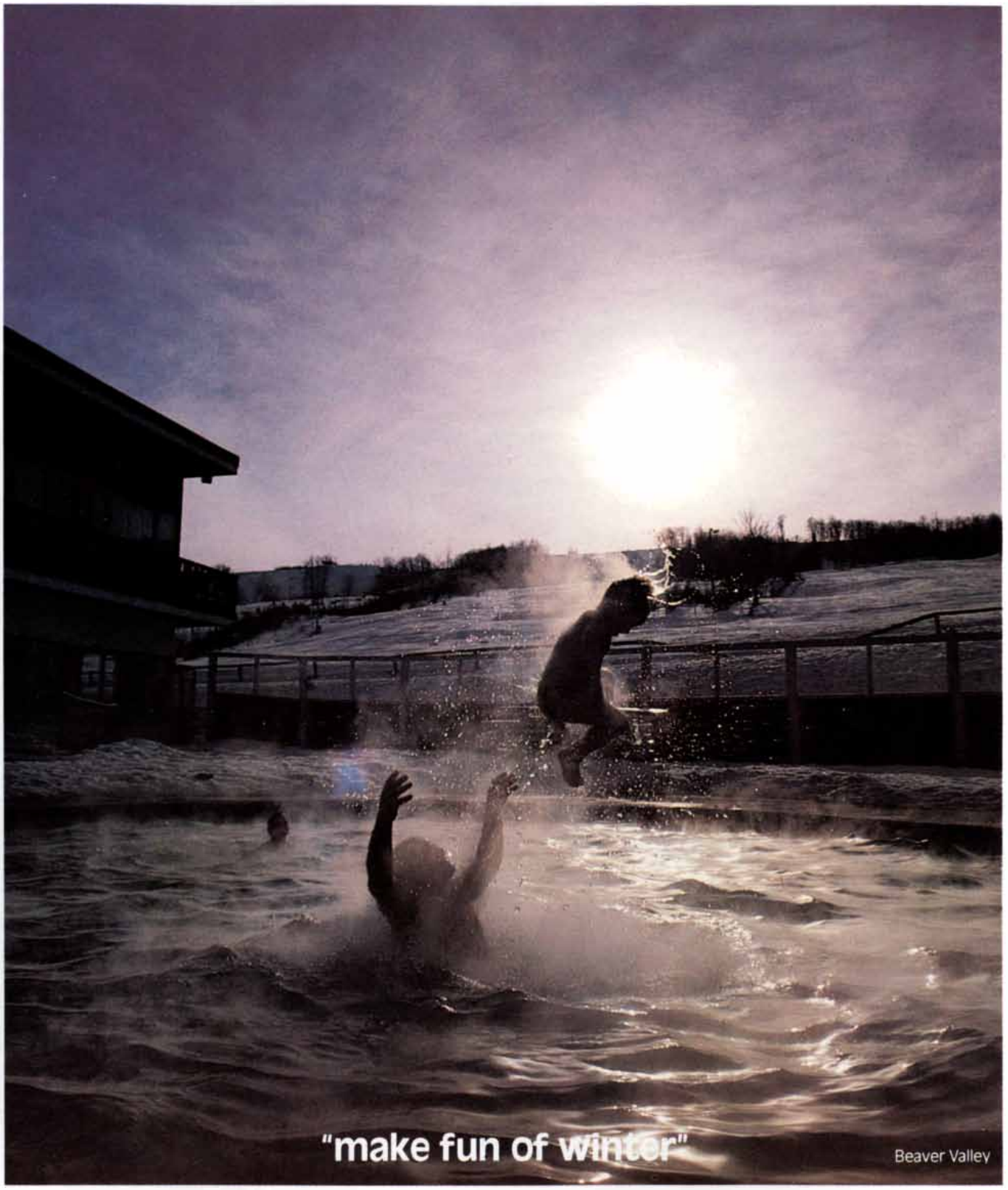
TRAITÉ THÉORIQUE ET PRATIQUE DES MOTEURS HYDRAULIQUES. Armengaud Ainé. Published by the author, 1858.

BENOÎT FOURNEYRON (1802–1867). Frederic W. Keator in *Mechanical Engineering,* Vol. 61, No. 4, pages 295–301; April, 1939.

MAN AND WATER: A HISTORY OF HYDROTECHNOLOGY. Norman Smith. Peter Davies, Ltd., 1976.

THE AMATEUR SCIENTIST

ANALYTICAL ABSORPTION SPECTROSCOPY: ABSORPTIMETRY AND COLORIMETRY. Edited by M. G. Mellon. John Wiley & Sons, Inc., 1950.



Head for the hills or go cross-country – we've all kinds of white stuff to test your mettle and après ski, some first class resorts, mulled wine and good times to help you unwind. For more about our winter holidays, call us COLLECT (416) 965-4008, or write: Ontario Travel, Dept. M.F., Queen's Park, Toronto, M7A 2E5.

Ontario Canada

We treat you royally

MOTOROLA HAS DEVELOPED A

It isn't as if we had invented the wheel, but at Motorola we appreciate some of the feelings of the person who did. Because we have developed a new microprocessor—the 16-bit MC68000—of such exceptional speed and capacity that, quite frankly, we don't know all the uses that equipment designers will find for it.

Of course, we know a lot of its applications. And its capabilities are so impressive that we think the MC68000 is authentically a new frontier in electronic innovation. And because we have developed the technology for making the MC68000 in great quantities, we think it may indeed be the great problem-solver of the eighties.

ELECTRONIC HORSEPOWER.

We developed the MC68000 to handle jobs that benefit from more speed and efficiency than even the best available microprocessors could supply. For instance, the additional electronic horsepower of the MC68000 can more cost-effectively turn a typewriter work-station into a complete desktop computer, one with a capacity available only in a room-size computer just a few years ago.

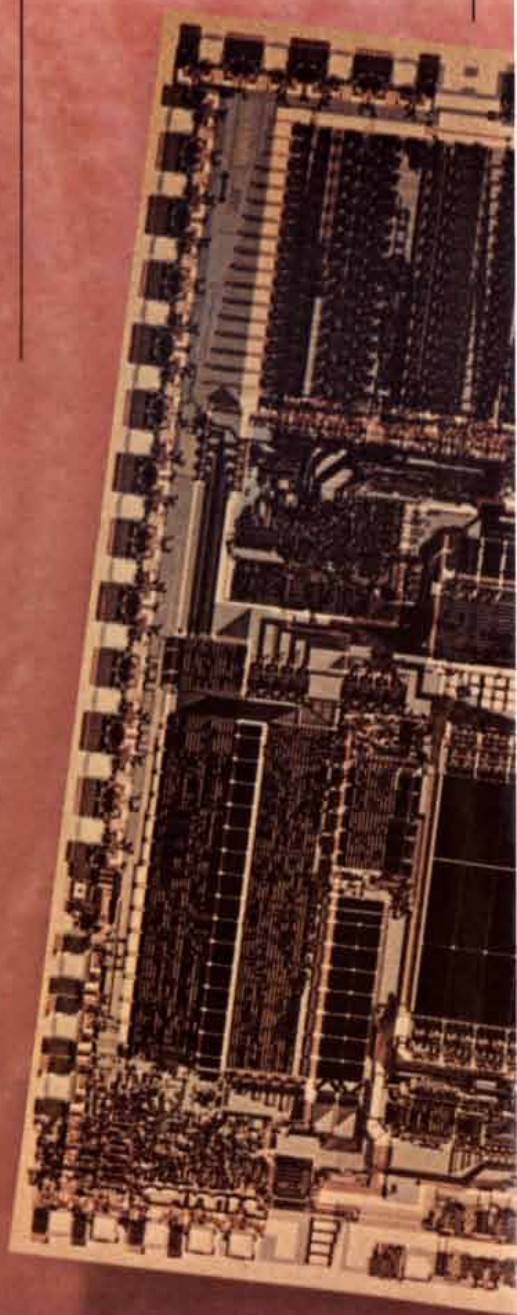
It can give a telephone switchboard the discretion to direct many thousands of phone calls simultaneously, each by the cheapest route.

It can give the auto industry an important increase in the ability to control engine efficiency by giving engines the ability to respond to measurements of such minutiae as air temperature, barometric pressure, altitude, even changes in the chemical composition of gasoline.

It can give the utility industries the ability to distribute electricity in response to second-by-second changes in customer demand.

It can sleeplessly tend scores of patients in intensive care, automatically and instantly responding to critical signs (it can cope with two million instructions per second) with all the relevant corrections.

Even more exciting, however, are the applications we know nothing about.



PROBLEM-SOLVER FOR THE '80s.

ELECTRONIC HISTORY IN THE MAKING.

What are the applications of a superchip that measures only one-quarter inch square, but has the ability to control a memory containing 128 million bits of data, and can complete a job in microseconds instead of minutes? The mind flies.

Perhaps a worldwide medical diagnostic network that recognizes the symptoms and knows the prescribed treatment for every identified disease on earth. Perhaps computerized machinery to disassemble a space satellite automatically the moment its orbit begins to decay.

Perhaps an industrial computer that can handle everything, from extraction of raw materials, to packing and shipping to retail outlets. Or an automated national allocation system for home heating fuel, with computers in every city and town to regulate fuel distribution by local temperature fluctuations.

PUSHING THE LIMITS OF MICROELECTRONICS.

Motorola, of course, makes only the microprocessor. But the MC68000 is a tool of enormous sophistication. In a very legitimate sense, it is a tool for expanding minds, extending possibilities, exciting exactly that sense of practical wonderment that is the real source of all progress.

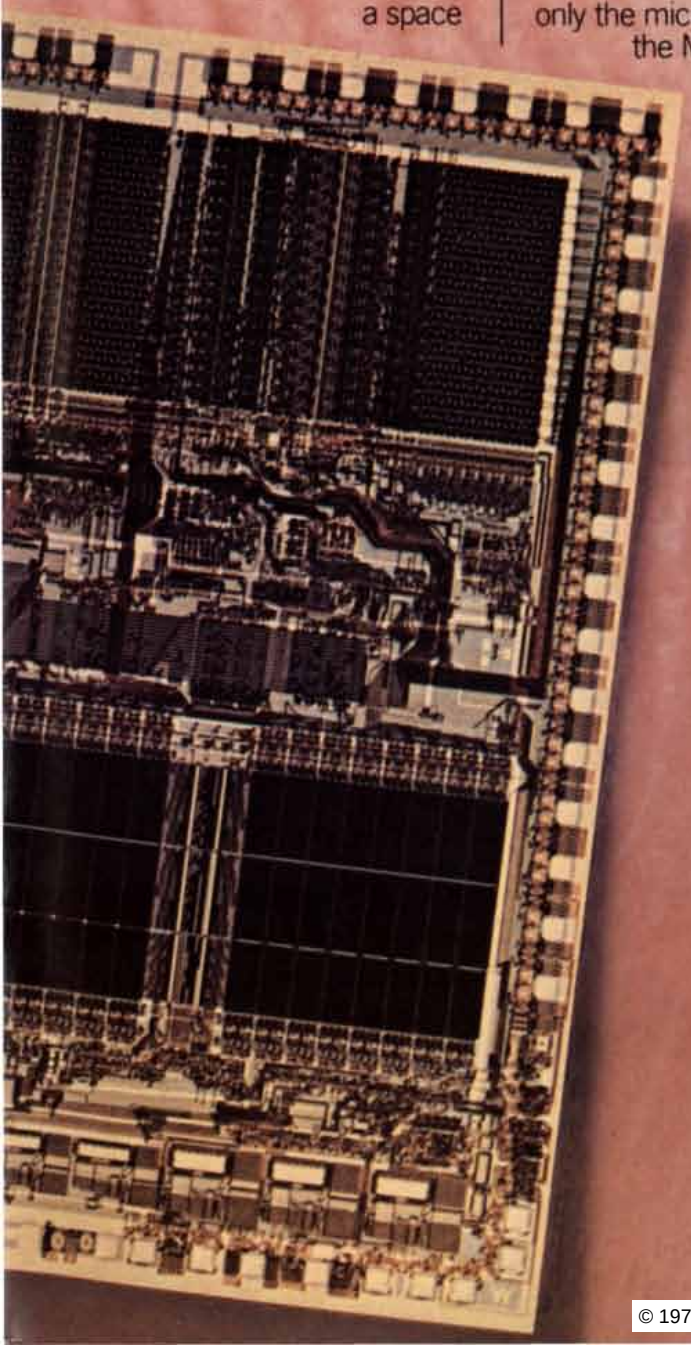
Motorola's MC68000 is a giant stride even for a company that systematically explores the very limits of microelectronics technology; that has become a foremost producer of semiconductors; that has become one of the world's largest manufacturers devoted exclusively to electronics.

We have been making electronics history since 1928. With the MC68000, we have outdone ourselves.

For further information, write Public Affairs Office, Corporate Offices, Motorola, Inc., 1303 E. Algonquin Road, Schaumburg Illinois 60196.

MOTOROLA

Making electronics history since 1928.



Technics

From the grandest opera to the Grand Ole Opry. A lot of FM stations play a lot of different music yet still have one thing in common: The need for uncommonly accurate turntables. That's why so many FM stations use Technics direct drive turntables.

That professionals use Technics direct drive turntables is really not surprising. What is, is that now you can get professional performance in Technics quartz-synthesizer MK2 Series: The SL-1800 manual, the SL-1700 semi-automatic and the SL-1600 fully automatic.

Wow & Flutter	Rumble	Speed Accuracy	Start-up Time
0.025% WRMS	−78 DIN B	±0.002%	1/4 rotation

As you can see, they all have impressive performance. But with Technics MK2 Series, you also get impressive advances in electronics. Like a quartz-synthesizer pitch control. As you vary the pitch it's instantaneously displayed by 13 LED's in exact 1% increments. That makes life easy.

So does the SL-1600 MK2's infrared disc-size sensor. Just place a disc on the platter, press the start button and immediately an infrared ray activates the micro-computer. Then the Technics precision gimbal-suspension tonearm automatically sets down in the lead-in groove.

And for double protection against acoustic feedback, Technics precision aluminum diecast base has a double-isolated suspension system. One damps out vibration from the base, the other from the tonearm and platter.

The MK2 Series. You don't have to be a radio station to afford performance good enough for a radio station.

Your next turntable should be as accurate as the ones many radio stations use.