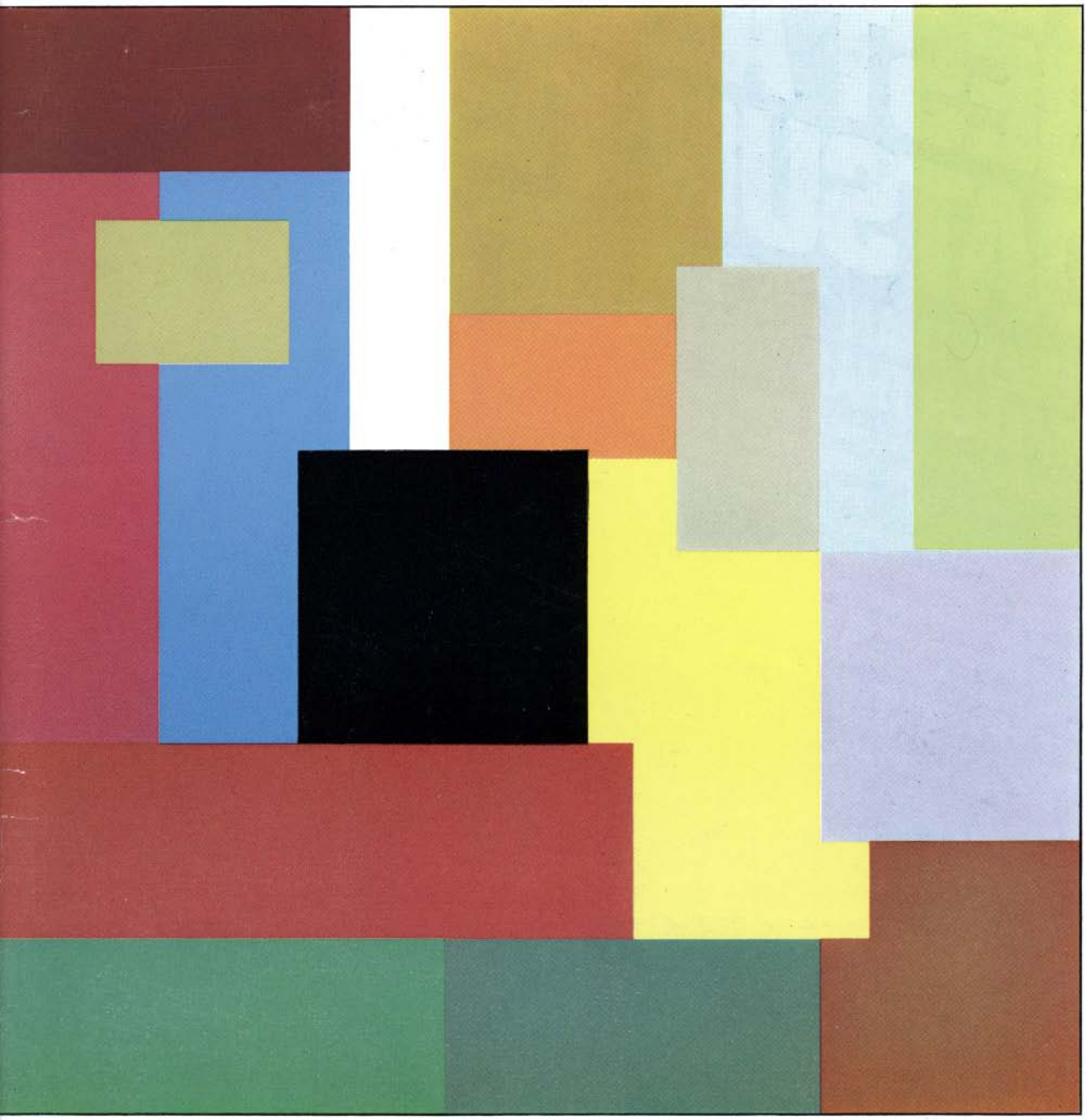
SCIENTIFIC AMERICAN



COLOR-THEORY EXPERIMENT

$1.50

December 1977

THE NEW 510. BEST ALL-AROUND DATSUN YET.

BIGGER THAN THE SMALL CARS. SMALLER THAN THE BIG CARS.

A whole new class of economy car. We drove our engineers to the limits. We demanded a car whose performance, room and ride were bigger than the small cars. A car whose price and handling were smaller than the big cars. We demanded a lot. And we got it.

The new 510 is big where you need it. In leg room, hip room, family room. Datsun engineers even extended the front seat track for extra stretch-out room.

The new 510 is small where it counts. On the outside, which means it can slip into the snuggest of parking spaces. New light-touch steering helps make it easy.

A. 5-Door Wagon.
B. 3-Door Hatchback.
C. 4-Door Sedan.
D. 2-Door Sedan.

Levers, switches, gauges galore. The new 510 abounds in conveniences you won't find in every economy car. Stalks mounted on the steering column control lights, turn signals, wipers and washer. A multi-level air-flow system cools you all over and, wonder of wonders, defrosts and defogs even the side windows. An electric defroster helps keep the rear window clear.

Built to endure. Every new 510 hitches an all-steel, all-welded unibody to a 2-liter overhead cam power plant we've raced to test endurance, and which won to prove it. A MacPherson strut front suspension helps iron out road wrinkles. And beefed-up insulation helps keep the noisy outside outside. Just like the big cars.

First and foremost, it's a Datsun. The new 510 is Datsun ingenuity at its best. A car that's big where you want it big, and small where you want it small. A whole new class of economy car, in Wagon, 2- and 4-Door Sedans and a Hatch-back that got 35 mpg highway, 25 mpg city. These are 1978 EPA estimates based on standard manual transmission. Your mileage may differ, depending on how and where you drive, your car's condition and optional equipment. California/high-altitude mileage lower.

The new 510. It's Datsun, dedicated to perfection; driven to excellence. Drive the new 510 for yourself. You'll find it's not only the best all-around Datsun yet. It just may be the best all-around economy car yet.

NOBODY DEMANDS MORE FROM A DATSUN THAN DATSUN.

DATSUN. WE ARE DRIVEN.

"You're perfect, 510!"

If you know several people who deserve Crown Royal, they're very fortunate. So are you.

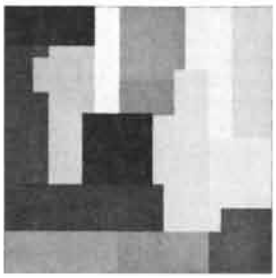
SEAGRAM DISTILLERS CO., N.Y., N.Y. BLENDED CANADIAN WHISKY. 80 PROOF.

Established 1845 **SCIENTIFIC AMERICAN** December 1977 Volume 237 Number 6

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $22 PER YEAR, ALL OTHER COUNTRIES.



THE COVER

The pattern on the cover was used in experiments testing Edwin H. Land's retinex theory of color vision, an account of which begins on page 108. Because the pattern bears a resemblance to the works of the Dutch painter Piet Mondrian, Land refers to this display and similar ones as Mondrians. In more elaborate examples (*see top illustration on page 111*) perhaps 100 pieces of paper of various colors and sizes are mounted on large boards and so arranged that each piece of paper is surrounded by at least five or six other pieces of different colors. In a typical demonstration the Mondrian is illuminated by projectors that provide adjustable amounts of radiant energy in three wave bands: long ("red"), middle ("green") and short ("blue"). With the proper selection of the mixture of illuminants falling on the Mondrian the radiant flux reaching the eye from any selected area can be made to match the flux that had previously reached the eye from a totally different area. In the first instance the selected area could have been red; in the second instance it could have been green. With the same flux of energy reaching the eye the two areas will still be seen as red and green.

THE ILLUSTRATIONS

Cover photograph by Julius J. Scarpetti

Page	Source
17	Culver Pictures, Inc.
18–53	Ilil Arbel
54	G. Nigel Godson, Yale University School of Medicine
55	John Finch, Medical Research Council Laboratory of Molecular Biology
56	G. Nigel Godson, Yale University School of Medicine (*top*); Jack D. Griffith, University of North Carolina School of Medicine (*bottom*)
57–67	Jerome Kuhl
68	U.S. Army
70–78	Alan D. Iselin
89	Culver Pictures, Inc.
90	Gopal Murti, St. Jude Children's Research Hospital
91–103	Bunji Tagawa
109	Julius J. Scarpetti
110	Jerome Kuhl
111	Julius J. Scarpetti (*top*), Jerome Kuhl (*bottom*)
112–119	Jerome Kuhl
120	Julius J. Scarpetti
122–128	Jerome Kuhl
131	Yoichi Ishida, University of Tokyo
132–138	George V. Kelvin
139	Ted Lodigensky
140–144	George V. Kelvin
147–153	Turid Hölldobler
154	Type Systems Inc.
156	Aerofilms Ltd.
158	Cambridge University Collection (*top*); Department of the Environment, Fortress House (*bottom*)
159	Department of the Environment, Fortress House
160–166	Tom Prentiss
168	Department of the Environment, Fortress House
172–178	Michael Goodman

THE ONLY THING IT HAS IN COMMON WITH OTHER 7-INCH TAPE DECKS IS THE SIZE OF ITS REELS.

Pioneer's new RT-707 has a lot more in common with today's most sophisticated 10-inch tape decks than it does with most 7-inch tape decks.

Because unlike other 7-inch tape decks, the RT-707 isn't filled with 15 year old ideas.

Take the drive system of the RT-707.

Instead of the old fashioned belt-drive system, the RT-707 is driven by a far more accurate and efficient AC Servo direct-drive motor. This motor generates its own frequency to help correct even the slightest variation in tape speed. Which all but eliminates wow and flutter. And because it doesn't generate heat like the belt-driven "dinosaurs" it doesn't need a fan. So all you'll hear is music with a clarity and crispness not possible on any 7-inch, or many 10-inch tape decks.

Our direct-drive system also makes pitch control possible. To help you regulate the speed of the tape and give you greater control over your recordings.

With technology like this it shouldn't surprise you that our super-sensitive heads will deliver frequencies from 20 to 28,000 Hertz. And our pre-amp section is built to handle 30 decibels more than any other 7-inch tape deck without distorting.

But great sound isn't everything.

As you can see, the RT-707 is smaller and more compact than other tape decks. It's also rack-mountable. And unlike *any* other tape deck, it's stackable. So it'll fit right in with the rest of your components.

But frankly, all the revolutionary thinking that went into the RT-707 wouldn't mean much if it weren't also built to fit comfortably into your budget. It is.

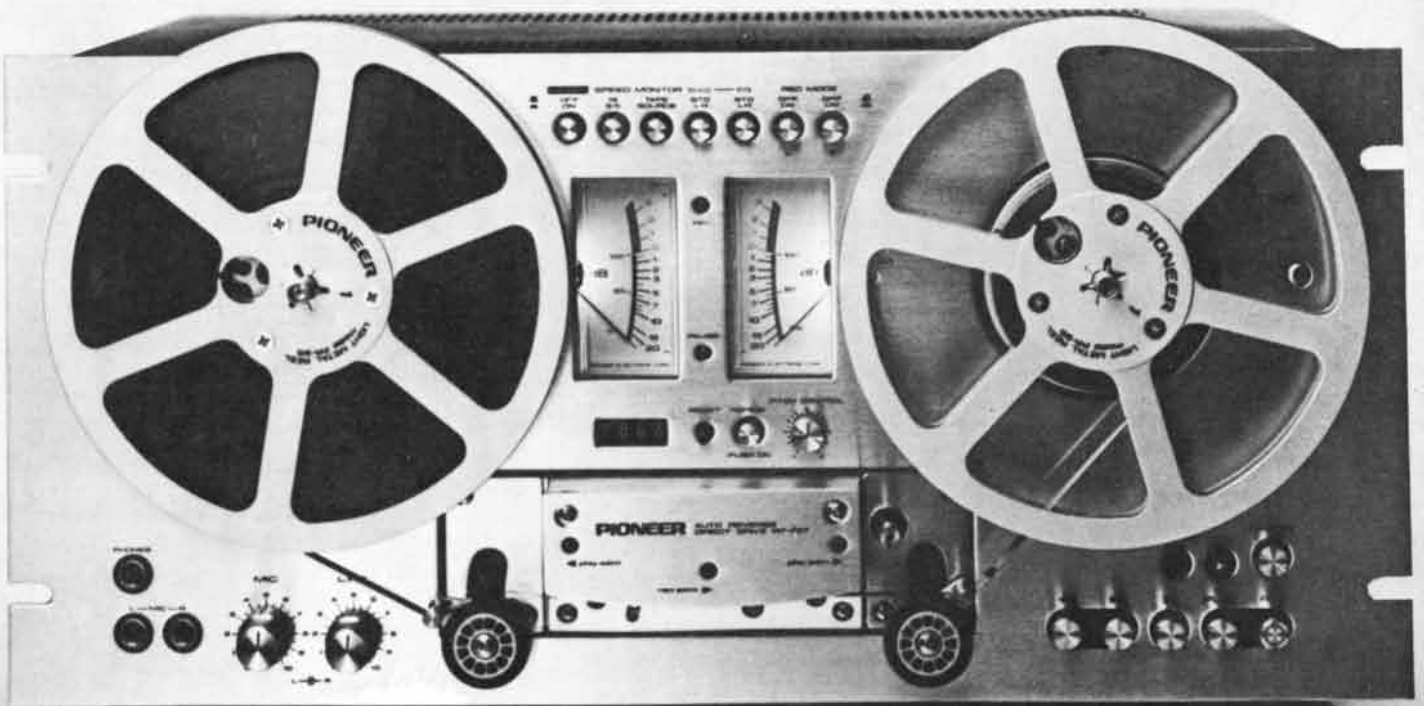
See your Pioneer dealer for a closer look at this extraordinary 7-inch tape deck.

We think you'll find the only things that the RT-707 has in common with other 7-inch tape decks is the size of the reels. And the size of the price.

High Fidelity Components

PIONEER

WE BRING IT BACK ALIVE.



THE RT 707.

LETTERS

Sirs:

William A. Silverman's article "The Lesson of Retrolental Fibroplasia" [SCIENTIFIC AMERICAN, June] was noted with interest by the members of the National Commission for the Protection of Human Subjects. Indeed, the lesson of which Dr. Silverman speaks is one that has been cited several times during the commission's deliberations as an example of the need for controlled clinical trials in order to protect children from harms that may attend the use of unvalidated or untested therapies. The need to do research, however, must be balanced against a concern for the rights and welfare of children who may participate in research. This commission has recommended guidelines for attaining that balance in a report on research involving children that was recently transmitted to the Secretary of Health, Education, and Welfare.

I should like to note for your readers that Dr. Silverman's account of Federal regulations governing research is inaccurate. This commission is not aware of any regulation or policy that would prevent the participation of children (including newborn infants) in properly designed clinical trials when the safety and

Scientific American, December, 1977; Vol. 237, No. 6. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to 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 35 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_____

efficacy of an accepted treatment is called into doubt, or when it is unclear which of two interventions is preferable. Review of the protocol by a local Institutional Review Board is required to ensure that any risk to the subjects is outweighed by the benefit to them and the importance of the knowledge to be gained, that the rights and welfare of the subjects will be adequately protected and that informed consent will be obtained from those legally authorized to give it. The requirements of local review and informed consent are an implicit recognition of the value nature of the decisions that must be made.

The commission agrees with Dr. Silverman that use of the terms "therapeutic" and "nontherapeutic" research may obscure these decisions. We have substituted an approach that employs determinations regarding the nature of the risk and anticipated benefit that are presented by proposed research. Those interested in the issues raised at the close of Dr. Silverman's article are invited to write to the commission for copies of its report on research involving children.

KENNETH J. RYAN

Chairman
National Commission
for the Protection
of Human Subjects
of Biomedical and Behavioral
Research
Bethesda, Md.

Sirs:

It is encouraging to learn that the National Commission is now ready to discard the figmentary labels "therapeutic" and "nontherapeutic" research (see "Nontherapeutic Research in Children: An Ethical Dilemma," by Charles U. Lowe, D. Alexander and B. Mishkin; *Journal of Pediatrics,* Vol. 84, page 468, 1974). The rest of the news from Bethesda is less cheering.

The qualifying adverb "however" in the third sentence of Dr. Ryan's letter serves to perpetuate the myth that the "need to do research" and the "rights and welfare of children" are antithetical, when in fact they are complementary. Surely the need to be protected against unexpected hazards of unevaluated innovations and treatments should rank high on a list of children's rights. Dr. Ryan's statement of negation ("...not aware of any regulation that would *prevent* the participation of children...") sheds light on an important problem. Until the commission undertakes positive actions to *promote* use of low-risk strategies of randomized clinical trials in studies involving children (see "The Epidemiology of the Gastrointestinal Randomized Clinical Trial," by E. Juhl, E. Christensen and N. Tygstrup; *New England Journal of Medicine,* Vol. 296, page 20, 1977), their rights and welfare are guarded by a paper tiger.

WILLIAM A. SILVERMAN

Greenbrae, Calif.

Sirs:

The stimulating article "The History of the Airflow Car," by Howard S. Irwin [SCIENTIFIC AMERICAN, August], contains a misleading paragraph on streamlining and drag. Streamlining does not deal with the maintenance of laminarity (as opposed to turbulence) but with the minimizing of the effects of the "separation" of the flow of air around a vehicle, particularly at the vehicle's rear. Separated-flow regions (formed when the oncoming fluid ceases to follow the body surface) can be made smaller and less wasteful of energy by streamlining regardless of whether the flow is laminar or turbulent. In fact, Ludwig Prandtl demonstrated that the drag of smooth, bluff bodies can be dramatically decreased by artificially causing the flow to become turbulent just upstream of the maximum body cross section. In other words, improved streamlining is achieved through turbulence.

Three-dimensional separated flows remain one of the least understood domains of aerodynamics in spite of the availability of powerful computers for studying them. Both the basic and the applied issues were discussed recently at an informative international symposium convoked by the General Motors Research Laboratories, titled "Aerodynamic Drag Mechanisms of Bluff Bodies and Road Vehicles." (The proceedings, edited by G. Sovram, T. Morel and W. Mason, were published by Plenum Press this year.) These days decreasing the drag of automobiles and trucks has less to do with increasing speed than it does with conserving energy and fuel. It is therefore gratifying that the Energy Research and Development Administration and the Department of Transportation have just launched an impressive five-year effort in aerodynamic research on vehicles of the automotive type. The General Motors symposium showed that there is still much to be learned today, 43 years after the birth of the Airflow car.

MARK V. MORKOVIN

Professor of Mechanical and
Aerospace Engineering
Illinois Institute of Technology
Chicago, Ill.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

DECEMBER, 1927: "An important victory in radio was won by Dr. Lee De Forest when the United States Circuit Court of Appeals at Philadelphia recently held that he is the inventor of the radio feed-back circuit and the oscillating audion. The court overruled the claims of the Government that Alexander Meissner, a German inventor, discovered the feed-back (regenerative) circuit and audion (vacuum tube), the claims of the Westinghouse Electric and Manufacturing Company that Major Edwin H. Armstrong was the inventor and those of the General Electric Company that these important radio discoveries were made by Irving Langmuir. The controversy has been in litigation for several years. Dr. De Forest has won four decisions, but it is probable that before the issue is finally settled it will be passed on by the United States Supreme Court."

"At the International Conference of Cancer Control held at Lake Mohonk, N.Y., Dr. L. I. Dublin considered the chance of death from cancer for persons living today. His investigations were based on the records of the Metropolitan Life Insurance Company, and they showed that cancer as a cause of disability and death is increasing. Of course, much of this increase is due to the fact that people nowadays are living longer than they used to and die from the diseases of advanced years rather than from the infectious diseases that formerly carried them off at an early age. Currently the probability at age 10 of dying ultimately from cancer is exceeded only by the probability of dying from three other causes: hemorrhage into the brain, chronic inflammation of the kidneys and heart disease. Among women cancer is third on the list, being exceeded only by heart disease and brain hemorrhage. For each of the ages beyond 10 up to the age of 90, over the past 15 years there has been an increase in the cancer hazard running from 40 to 50 per cent. The increase for women has been less than that for men, approximately 21 per cent. Cancer of the mouth and throat is far commoner in people who smoke than in those who do not."

"At an address to the American Chemical Society at State College, Pa., Miss Elizabeth Wagner was presented as a model of a modern bride clad from head to toe, except for the soles of her slippers, in synthetic materials. Her dress was made of rayon fibers trimmed with rayon lace. The sleeves were of cellulose acetate fibers. Her tulle bridal veil was a nitrocellulose product. The orange blossoms were precipitated calcium carbonate coated with paraffin. Her stockings were of rayon. Her slippers were of rayon and metal threads, the metal a tin-copper alloy. Her beads were made of collodion with fish-scale essence as the iridescent material. Her prayer book had a celluloid back and the paper and ink were both chemical products. Even the traditional garter, embodying 'something old, something borrowed and something blue,' was made of rubber rendered adaptable by chemical treatment, covered with rayon and ornamented with rayon and metal roses. And as for the remainder of the bride's trousseau, she now has an unlimited number of synthetic textile fabrics from which to choose. Brocaded rayon velvets, rayon fabrics for sport wear, charming color effects made by combining two kinds of rayon with different dyeing properties and dresses and shawls embroidered with lustrous rayon threads, hats of rayon, of rayon plush or of cellophane, beads of collodion, of glass, of casein or of Bakelite—all these and more are to be had."

SCIENTIFIC AMERICAN

DECEMBER, 1877: "Mr. Thomas A. Edison recently came into this office, placed a little machine on our desk and turned a crank. The machine inquired as to our health, asked how we liked the phonograph, informed us that *it* was very well and bade us a cordial goodnight. These remarks were perfectly audible not only to ourselves but also to a dozen or more persons gathered around. The principle on which the machine operates is as follows. There is, first, a mouth piece across the inner orifice of which is a metal diaphragm. To the center of this diaphragm is attached a point, also of metal. A brass cylinder is supported on a shaft that is screw-threaded and turns in a nut for a bearing, so that when the cylinder is caused to revolve by the crank, it has a horizontal travel in front of the mouth piece. It will be clear that the point on the metal diaphragm must describe a spiral trace over the surface of the cylinder. On the cylinder is cut a spiral groove of a pitch corresponding to that on the shaft, and round the cylinder is attached a piece of tinfoil. When sounds are uttered into the mouth piece, the diaphragm is caused to vibrate, and the point thereon is caused to make contact with the tinfoil at the portion where the latter crosses the spiral groove. Hence the foil, not being there backed by the solid metal of the cylinder, becomes indented, and these inden-

tations are necessarily an exact record of the sounds that produce them. The reading mechanism is nothing but another diaphragm held in a tube on the opposite side of the machine and a point of metal that is held against the tinfoil on the cylinder by a delicate spring. The metal point is caused to vibrate as it is affected by the passage of the indentations. These vibrations, transmitted to a second membrane, must cause the latter to vibrate like the first membrane. The result is a synthesis of the sounds that in the beginning were analyzed."

"According to recent careful computations the population of the world is 1,423,917,000. The populations of the great divisions of the earth are as follows: Europe, 309,178,300; Asia, 824,548,500; Africa, 199,921,600; Australia, 4,748,600; America, 85,519,800. London has 3,489,428 souls, Paris 1,851,792, New York and Brooklyn 1,535,622 and Berlin 1,045,000."

"The number of oil wells going down in Pennsylvania is simply enormous. The principal field of production is the eastern belt, or what is known by that title in the district. A large number of 'wildcats,' or test wells, have gone down off the eastern edge of the defined line, but with very few exceptions they have proved to be dusters. The most productive territory appears to be in the vicinity of Slam Bang City, a new town on the 22nd degree line. The Davis well, struck at this place about three months ago, started off at 400 barrels a day, and it has averaged about 125 barrels a day ever since the 'head' was pumped off."

"We cannot recall in our time so gross an infringement on the rights of the people in relation to their property as is now being perpetrated in the erection of the elevated railway in New York. The effect of this outrage is to disfigure and otherwise damage several thoroughfares throughout the length of the city for the benefit of a clique of stock speculators and out-of-town land owners. We do not believe that this railway corporation has any legal right to erect its structures. When a street is opened for all kinds of public uses by compensating the land owners for the property thus taken, the government that represents the ownership of the acquired domain may authorize the erection thereupon of anything that will not interfere with such use. But even if the government can do this, the right has been most wantonly exercised, with no proper limitations and with a recklessness with regard to both public and private interests that is simply astounding. The people of New York will bitterly repent some day of this gross injustice, but the monopoly they have created, having seized its prey, will care nothing for their penitence."

THE AUTHORS

G. FRANKLIN MONTGOMERY ("Product Technology and the Consumer") is a senior engineering adviser in the Center for Consumer Product Technology of the National Bureau of Standards. He received his bachelor's degree in electrical engineering at Purdue University in 1941 and spent the next five years working at the Naval Research Laboratory and serving in the Army Signal Corps. Since 1946 he has been at the Bureau of Standards, where he has worked on the propagation of radio waves, electronic instrumentation and product technology.

JOHN C. FIDDES ("The Nucleotide Sequence of a Viral DNA") is a postdoctoral fellow in the department of biochemistry and biophysics at the University of California at San Francisco. He received his bachelor's degree at the University of Edinburgh in 1973 and continued his study of biochemistry at the University of Cambridge. Later he joined Frederick Sanger's group at the British Medical Research Council's Laboratory of Molecular Biology and participated in the determination of the nucleotide sequence of the bacterial virus ϕX174. He received his Ph.D. from Cambridge in the spring.

DAVID M. BOORE ("The Motion of the Ground in Earthquakes") is assistant professor of geophysics at Stanford University. He was educated at Stanford and the Massachusetts Institute of Technology, where he received his Ph.D. in geophysics in 1969. After a postdoctoral year at M.I.T. he turned to applied seismology and became a research associate of the National Research Council, working at the U.S. Geological Survey. There he was involved in evaluating seismic risk, including the "specification of ground-motion parameters for the trans-Alaskan pipeline and the Los Angeles Dam." Boore joined the Stanford faculty in 1972.

MARTIN M. KAPLAN and ROBERT G. WEBSTER ("The Epidemiology of Influenza") share an interest in molecular epidemiology: the study of the relation between infectious disease and the structure and function of viruses. Kaplan is currently director general of the Pugwash Conferences on Science and World Affairs, based in Geneva. He obtained his public-health and veterinary degrees respectively from the University of Pennsylvania's School of Medicine (1940) and School of Veterinary Medicine (1942). A member of the staff of the World Health Organization (WHO) for 26 years, he served first as head of its veterinary public-health activities and then as special science adviser and director of research promotion and development in the office of the director general. During that time he took extended leaves of absence for laboratory work on viruses at the Wistar Institute in Philadelphia. Kaplan retired from the WHO last year to assume his position at the Pugwash Conferences and to continue his research in epidemiology and immunology. Webster is a member of the staff of the St. Jude Children's Research Hospital in Memphis, Tenn., and director of one of the U.S. collaborating centers of the WHO concerned with the ecology of animal influenza viruses. A native of New Zealand, he obtained his Ph.D. in microbiology from the Australian National University in 1962. He spent the next two years as a Fulbright scholar in the laboratory of Thomas Francis at the University of Michigan's School of Public Health, where he became interested in influenza viruses. Later, together with Graeme Laver of the Australian National University, he developed a vaccine based on subunits of the influenza virus that, unlike whole-virus vaccines, is not toxic. Kaplan and Webster would like to thank Bernard C. Easterday of the University of Wisconsin for his help in the preparation of their article.

EDWIN H. LAND ("The Retinex Theory of Color Vision") is chairman of the board, director of research and chief executive officer of the Polaroid Corporation. Born in 1909, he attended Harvard College, where he developed a new type of polarizing filter in the form of an extensive synthetic sheet. In 1937 he founded Polaroid for research in the new field of applied polarization, and in 1944 he began his pioneering work in the development of "instant" photography. His one-step photographic process was first demonstrated to the Optical Society of America in February, 1947, and was made available to the public at the end of 1948. Land has received 14 honorary degrees, has held visiting academic appointments at Harvard and is currently Institute Professor (Visiting) at the Massachusetts Institute of Technology. From 1960 to 1973 he was consultant-at-large to the President's Science Advisory Committee, and in 1967 he received the National Medal of Science. This year, on the occasion of his 500th U.S. patent, he was elected to the National Inventors Hall of Fame. Land has pursued his lively interest in the mechanisms of color vision for the past 25 years.

WILLIAM F. HARRIS ("Disclinations") is reader and associate professor

THE JOS. GARNEAU CO., NEW YORK, N.Y. ©1976.

Washington's farewell at Fraunces Tavern may well have included Martell.

1783. General Washington resigns his commission and asks Generals Lafayette and Rochambeau plus his closest officers and friends to Fraunces Tavern in New York. Here he will personally say farewell and thank them for their loyalty.

And what more appropriate way for devoted officers and friends to wish this great patriot well than with a glass of Martell Cognac. For by 1783, Frenchman and American had long agreed, Martell was cognac unequalled.

It still is.

Martell. Taste history.

V.S.P.-V.S.O.P.-CORDON BLEU.
WORLD'S LARGEST-SELLING COGNACS.

This year, take off in a 1978 Thunderbird

The new Thunderbird is cleared for take-off—with the styling, features and luxurious appointments that make the Thunderbird name famous. Flight test one—you'll find its wide stance and full-coil suspension complement its legendary Thunderbird quality.

$5,808 as shown.

But the most remarkable feature of all: this 1978 Thunderbird, equipped exactly as you see it here, comes at the down-to-earth price of only $5,808—sticker price excluding title, taxes and destination charges.

So don't drive just an ordinary car—isn't

this *your* year to fly Thunderbird?

Take off for your Ford Dealer now and flight test one today.

FORD

When America needs a better idea, Ford puts it on wheels.

Base sticker price: $5,411*

Power Front Disc Brakes	No Charge
Power Steering	No Charge
Automatic Transmission	No Charge
5.0 Litre (302 CID) V-8	No Charge
Full Wheel Covers	No Charge
AM Radio	No Charge
Electric Clock	No Charge
Steel-Belted Radials	No Charge
White Sidewall Tires	$ 46
Vinyl Roof	$ 138
Color-Keyed Bodyside Moldings	$ 54
Dual Accent Paint Stripes	$ 46
Deluxe Bumper Group	$ 76
Remote Control Mirror	$ 16
Color-Keyed Belts	$ 21
TOTAL	$5,808*

*Excluding title, taxes and destination charges

THUNDERBIRD

FORD DIVISION

Ford 75TH ANNIVERSARY

Power to the People.

Harris MICROPLEX® Energy Management System

Demand for electrical energy plus escalating fuel costs have put utilities in a real bind. Harris is easing the pressure with fuel-saving computerized energy management systems that help generate and distribute power more economically.
Write: Harris Corporation,
55 Public Square, Cleveland, Ohio 44113

HARRIS

COMMUNICATIONS AND INFORMATION HANDLING

in microbiomechanics in the department of chemical engineering at the University of the Witwatersrand. A native of South Africa, he obtained his bachelor's degree in chemical engineering at Witwatersrand in 1962 and went on to study with L. Edward Scriven at the University of Minnesota. After obtaining a master's degree he returned briefly to Witwatersrand, where he studied the fracture of metals in the department of physics. In the course of that work he encountered the concept of dislocations in crystals and became interested in its possible applications to biology. He then returned to Scriven's laboratory to study the role of dislocations and related phenomena in membranes and shells; he received his Ph.D. in biophysics in 1970. Since then his research has focused on the biomechanics of the living cell. Harris spends his free time looking for dislocations and disclinations "in everything from art to zebras."

BERTHOLD K. HÖLLDOBLER and EDWARD O. WILSON ("Weaver Ants") are both professors of biology at Harvard University specializing in the social behavior of insects. Born in Germany, Hölldobler was educated at the University of Würzburg and the University of Frankfurt, where he received his doctorate in 1969. After six years on the faculty at Frankfurt he came to Harvard in 1973. He now teaches the basic course in animal behavior at Harvard and carries on a research program in the behavioral ecology of ants, both in the laboratory and in the field in Arizona and Florida. Hölldobler is often assisted in his work by his wife Turid, an illustrator of animal behavior. Wilson is Frank B. Baird, Jr., Professor of Science at Harvard and is also curator of entomology in the university's Museum of Comparative Zoology. A graduate of the University of Alabama, he received his Ph.D. in biology from Harvard in 1955 and has been a member of the faculty ever since. He is the author of *The Insect Societies* (1971) and *Sociobiology: The New Synthesis* (1975).

GEOFFREY WAINWRIGHT ("A Celtic Farmstead in Southern Britain") is inspector of ancient monuments in the British Department of the Environment. He was educated at University College, Cardiff, and the Institute of Archaeology of the University of London, where he received his Ph.D. in 1961. After three years as professor of environmental archaeology at the University of Baroda in India he joined the Department of the Environment and directed numerous rescue excavations in England and Wales. In 1975 he was appointed director of the Central Excavation Unit, a state archaeology service "designed to conduct excavations on any threatened site in England."

MAKE SOMEONE'S SHIP COME IN.

When you give a bottle of Cutty Sark Scots Whisky in this shimmering gold package, it reflects nicely on you.

But it does even more for the one who receives.

Because inside the package is Scotch with a smooth, distinctive taste that is vastly appreciated.

In fact, last year Cutty Sark was appreciated by more Americans than any other Scotch.

So if you have friends still waiting for their ships to come in, there couldn't be a more appropriate vessel.

CUTTY SARK

MATHEMATICAL GAMES

Dr. Matrix goes to California to apply punk to rock study

by Martin Gardner

"The emotion consisted wholly of glee and admiration; glee at the vividness which such an abstract idea or verbal term as 'earthquake' could put on when translated into sensible reality and verified concretely; and admiration at the way in which the frail little wooden house could hold itself together in spite of such a shaking. I felt no trace whatever of fear; it was pure delight and welcome.

"'*Go* it,' I almost cried aloud, 'and go it *stronger!*'"

—WILLIAM JAMES, at Stanford University, 5:30 A.M., April 18, 1906, reacting to the San Francisco earthquake

"In vulgar language," writes Garrett Hardin, professor of human ecology at the University of California at Santa Barbara, "we need an earthquake-predicting facility like we need a hole in the head."

This quotation is from Hardin's essay "Earthquakes: Prediction More Devastating than Events," which is reprinted in his splendid collection of papers, *Stalking the Wild Taboo* (William Kaufmann, 1973). Hardin believes that if geophysicists ever succeed in predicting earthquakes with high accuracy, the social disruptions produced will do more damage than the quakes.

By a curious synchronicity I had just finished reading Hardin's essay when a letter arrived from an old friend and magic enthusiast, Persi Diaconis, now a statistician at Stanford University. Persi wrote to tell me that a man by the name of Dr. Punk Rockwell had rented a small office building on El Camino Real, between Redwood City and Menlo Park. The building was about 25 miles south of the heart of San Francisco and only a few minutes drive from Stanford. Dr. Rockwell was head of a firm called the Punk Earthquake Prediction Corporation (P.E.P.C.), which claimed to have a new, infallible method for predicting quakes.

For the past several months, Persi went on, Rockwell and his lady assistant, Punky Anderson, had been selling their service to farmers and other residents all along the California coast, from Eureka to San Diego. It was Dr. Rockwell's contention that some of the mechanical stresses that have been building up for decades along the 600-mile San Andreas fault—in particular those resulting in the recent upward bulging of the area surrounding Palmdale—would be suddenly released late in December, 1977. To anyone willing to pay a modest fee of $1,000 Dr. Rockwell would disclose the exact time and severity of the coming quake at any designated spot within 20 miles of the fault.

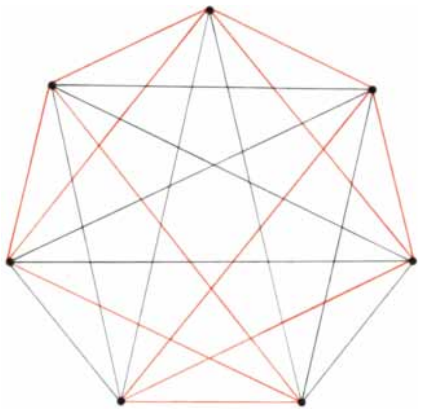
Persi's brief description of the Rockwell prediction method was so outlandish that I assumed he was putting me on. Then came the final revelation that made it all believable. "Dr. Rockwell," Persi added, "is a tall, elderly gentleman with a large hawklike nose and glittering green eyes. His assistant has unmistakably Oriental features. The pair are obviously none other than the notorious numerologist Dr. Irving Joshua Matrix and his Eurasian daughter Iva."

Whenever Dr. Matrix and Iva surface with a new swindle, I drop everything to pay them a visit. I have learned never to alert the pair to my arrival ahead of time. Although they have found me trustworthy in the past, how can they be certain I will not reveal their identity to the local police? For this reason I have found it advisable to take them by surprise.

At the San Francisco International Airport I rented a car and drove south along the Bayshore Freeway. Persi had offered to put me up for a few days at his house near Stanford. It was a fine, chilly November night, with an almost full moon rising on my left above the ominous waters of San Francisco Bay.

Persi and I stayed up late going over new variations of the Zarrow shuffle and the Ascanio spread and discussing Persi's unpublished work on the probabilities of ESP card testing. The next morning we drove to the nearby P.E.P.C. headquarters. Pressing the entrance bell

A scene in San Francisco following the earthquake of 1906



Solution to last month's Ramsey graph puzzle

produced the sound of muffled drumbeats behind the door.

Iva, as usual, looked startled when she opened the door. Then she flashed one of her seductive Oriental smiles and we exchanged kisses on the cheek before I introduced her to Persi.

I must have looked even more surprised than Iva. She was wearing the dirtiest pair of blue jeans I have ever seen. The fly was half-open, fastened with a large safety pin. A long tear in one jean leg disclosed 10 inches of bare thigh. Her shredded yellow shirt was dotted with black swastikas and the shirttails were tied in a loose knot above her bare midriff. Her hair, dyed a pale green, fell unkempt around her shoulders. A gold safety pin swung from each ear lobe. I had read about the new "punk look" that was spreading from London to California's subculture, but it was even punker than I had imagined.

"Dr. Rockwell is attending a conference at the Stanford Research Institute," Iva said, "but he should be back within the hour." Several seismologists from the Earthquake Research Center at Menlo Park and a top official of the Department of Defense were attending the conference. There was a good possibility, she hinted, that the SRI might obtain substantial Government funding for research on Dr. Rockwell's techniques.

Iva led the way to a small laboratory. Its walls were lined with hundreds of glass tanks containing what I guessed to be at least a million cockroaches. Persi and I listened with amused smiles while Iva solemnly explained how the Rockwell system operates. It was, she said, Dr. Helmut Schmidt who had made the initial breakthrough.

I nodded. I was quite familiar with Schmidt's work. He is a Ph.D. physicist who used to be the director of Dr. J. B. Rhine's parapsychology laboratory in Durham, N.C. I remembered that about seven years ago Schmidt had reported that cockroaches seemed to have the psychokinetic (PK) ability to cause a randomizer to give them electric shocks more often than chance allows. Schmidt admits, however, that he hates cockroaches. Since cockroaches presumably do not enjoy being shocked, Schmidt and his colleagues suspect that it was he who was the PK agent. The accepted theory is that cockroaches probably have some PK ability, but it is overwhelmed by the stronger, unconscious PK influence of experimenters who are not fond of cockroaches. (See Schmidt's "PK Experiments with Animals as Subjects" in *Journal of Parapsychology,* Vol. 34, 1970, and articles by Schmidt and others in later issues.)

Iva explained that Dr. Rockwell was the first to prove that in addition to their PK ability cockroaches have powers of precognition. His now classic experiment utilized a large box with doors at opposite sides, each leading to food. At one-hour intervals one of the two doors was opened at random, the choice being made by a computer using a tape of prerecorded random binary digits. Prerecording the digits, Iva added, eliminated the possibility that the cockroaches were employing PK to influence the randomizing process. The insects quickly learned to anticipate which door would open next. For several minutes before a door opened large numbers of roaches would swarm in front of it.

"How can you be sure," Persi asked, "that the roaches aren't picking up numbers on the tape by clairvoyance?"

"Good question," replied Iva, with a dim smile. Dr. Rockwell's later earthquake-prediction experiments, she explained, ruled out both clairvoyance and PK. "After all," she said, "it would be ridiculous to suppose that cockroaches have enough PK power to cause an earthquake."

"I wouldn't rule it out," I said. "Don't forget that severe quakes reduce supermarkets and kitchens to shambles and provide marvelous new food supplies. Millions of roaches, acting in telepathic union, might work up enough reinforced PK power to trigger action along a fault. Only a tiny kick is needed, you know, to snap the tension."

Iva gave me a cold stare. "You need a big kick," she said. She went on to explain Dr. Rockwell's sensational discovery of last January. Following the lead of a Russian parapsychologist who had been doing secret work on insect precognition, Dr. Rockwell found that the precognitive power of cockroaches can be greatly enhanced by two things: large doses of vitamin B-1 and smoke from burning punk.

Fortified by B-1 and drunk on punk, cockroaches become extraordinarily sensitive to future earth tremors. A careful monitoring of the animals' agitation by ingenious optical instruments provides accurate measurements of both the time and the intensity of a future quake. The length of time until the next quake is a function of the average distance a roach travels before it turns. The severity of the quake is a function of the average speed of the roach.

Iva told us that Dr. Wilhelm J. Levity, an American parapsychologist who recently emigrated to Romania, had been the most successful in replicating Dr. Rockwell's B-1-punk experiments. Working with Romanian roaches, Dr. Levity had predicted two weeks in advance the day and hour of the Romanian earthquake of last March that almost destroyed central Bucharest and killed hundreds of people.

"Wouldn't that make him a Romanian hero?" Persi asked. "How come we haven't heard of him?"

"The Romanian government has clamped tight security on Levity's work," Iva answered. "We've been trying to reach him. At the moment we don't even know if he's alive or dead."

"What do your roaches tell you about the coming California quake?" I asked.

"We know the exact day and hour," Iva said, "but we divulge this information only to our paying customers. I can tell you, though, that it will be before the end of December, and that horizontal slippage along parts of the San Andreas fault will be as much as six feet. But here comes Dr. Rockwell."

Brisk footsteps sounded outside the lab, and in strode Dr. Matrix. He was wearing tight black pants that appeared to be made of plastic garbage-bag material. A dirty T-shirt bore the smiling face of Dracula. Dr. Matrix' hair had been dyed a bright pink. A large Band-Aid on his forehead seemed to cover a fresh wound.

"So it's you again," Dr. Matrix said, his emerald eyes flashing mild hostility above a pair of half glasses. He turned toward Persi. "Dr. Diaconis, I presume."

"How did you know?" said Persi.

"I heard you lecture at Stanford last week on the probabilities of poker. And I have a friend—a dealer at Lake Tahoe—who says you do the second-best second deal west of the Mississippi."

"And who does the best?" said Persi, smiling.

"I do," said Dr. Matrix.

Iva excused herself. We followed Dr. Matrix to his office at the rear of the building, where we sat and chatted for more than an hour. A Confederate flag stood at one side of his desk. On the back wall a large portrait of President Fillmore beamed down at us.

"You may wonder," said Dr. Matrix, "what connection there is between our punk attire and earthquakes. It's more than just the use of punk to raise the psi level of cockroach consciousness. The punk movement is a refreshing protest against life's terrible unfairness. It's a punk, punk, punk, punk world. Can you think of a better symbol of nature's cru-

"The most complete and most scholarly dictionary of the English language"

—The Christian Science Monitor

As an introduction to membership in the

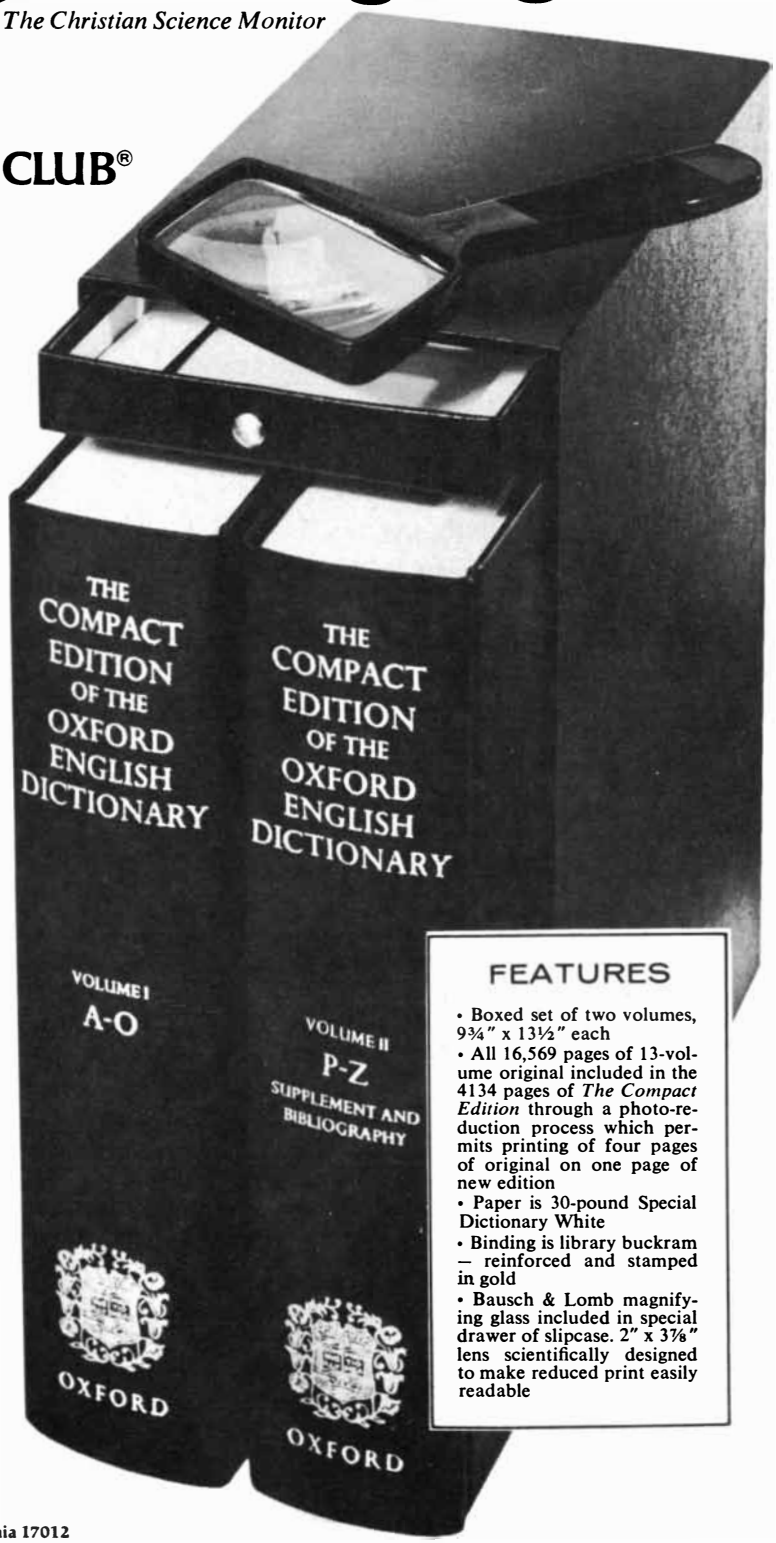
BOOK-OF-THE-MONTH CLUB®

yours for only

$17.50

[Publisher's list price $95]

THE SUGGESTED TRIAL: You simply agree to buy four Club choices within a year at substantial savings on most books you choose.



THE OXFORD ENGLISH DICTIONARY is generally regarded as the final arbiter of the English language. Until recently, it has been available only as a thirteen-volume set, currently priced at $395. Now, through the combination of an ingenious method of micrographic reproduction and a fine Bausch & Lomb optical lens, every single one of its 16,569 pages, 50 million words and close to 2 million illustrative quotations appears, in easily readable form, in *The Compact Edition.*

> Book critic Christopher Lehmann-Haupt of *The New York Times* has said of this edition: "It is something of a miracle.... *The Compact Edition* is easier to work with than the original with its thirteen separate volumes."

Even more extraordinary, as a trial member of the Club you may obtain the two-volume set for only $17.50. And as long as you remain a member, you will receive the *Book-of-the-Month Club News,*® a literary magazine announcing the coming Selection and describing other important books, known as Alternates, most of which are available at substantial discounts—up to **40%** on more expensive volumes.

If you continue after your trial membership, you will earn at least one Book-Dividend® Credit for every Selection or Alternate you buy. These Credits enable you to acquire for your library a wide variety of books, called Book-Dividends, at astonishing savings—*at least 70% of publishers' list prices.*

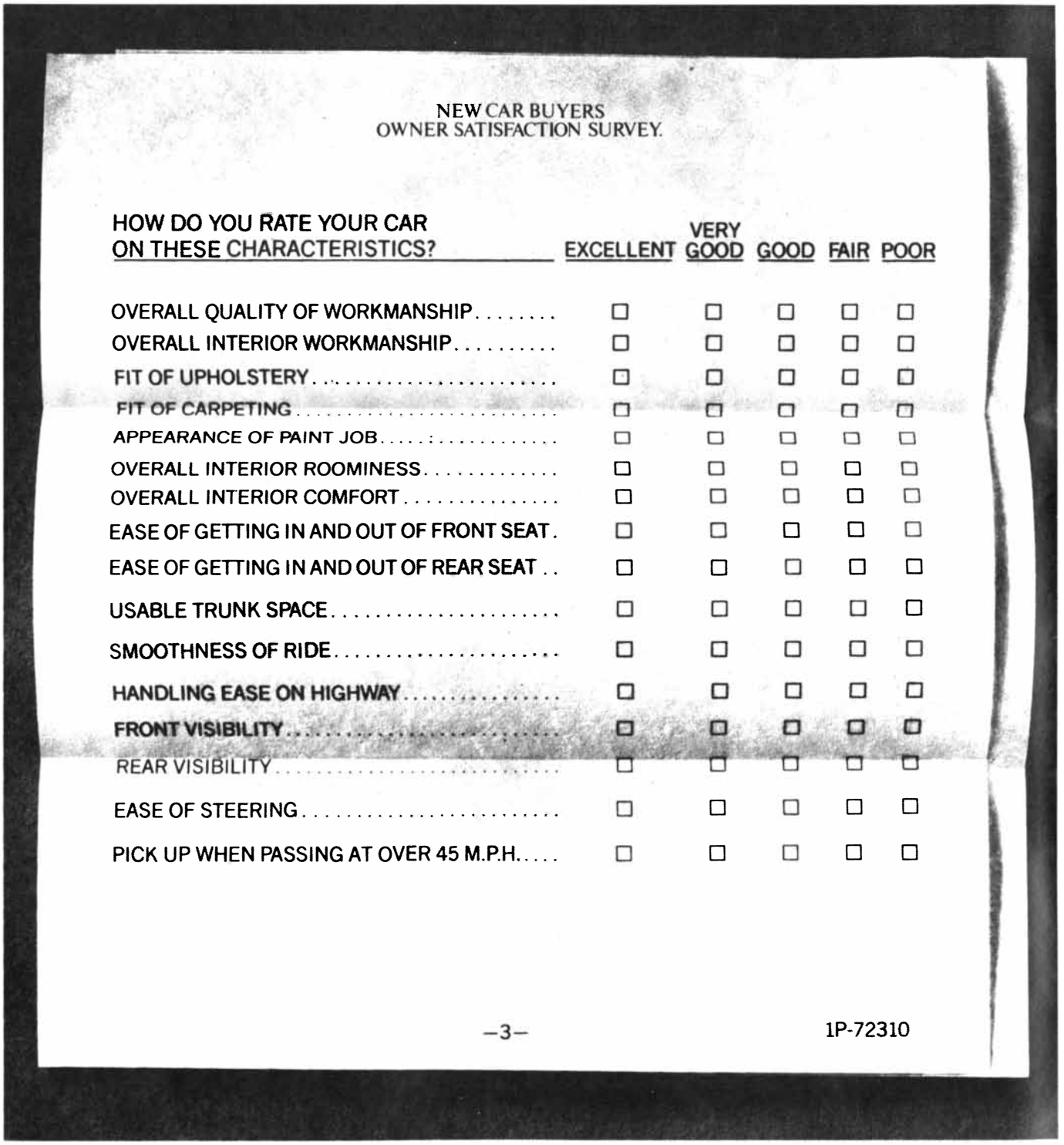
All Book-of-the-Month Club books, including Book-Dividends, are identical to the publishers' editions in content, format, size and quality.

FEATURES

- Boxed set of two volumes, 9¾" x 13½" each
- All 16,569 pages of 13-volume original included in the 4134 pages of *The Compact Edition* through a photo-reduction process which permits printing of four pages of original on one page of new edition
- Paper is 30-pound Special Dictionary White
- Binding is library buckram — reinforced and stamped in gold
- Bausch & Lomb magnifying glass included in special drawer of slipcase. 2" x 3⅞" lens scientifically designed to make reduced print easily readable

BOOK-OF-THE-MONTH CLUB, INC., Camp Hill, Pennsylvania 17012

DO YOU RATE YO
AS A VOLVO OWNE



NEW CAR BUYERS
OWNER SATISFACTION SURVEY.

HOW DO YOU RATE YOUR CAR ON THESE CHARACTERISTICS?	EXCELLENT	VERY GOOD	GOOD	FAIR	POOR
OVERALL QUALITY OF WORKMANSHIP	☐	☐	☐	☐	☐
OVERALL INTERIOR WORKMANSHIP	☐	☐	☐	☐	☐
FIT OF UPHOLSTERY	☐	☐	☐	☐	☐
FIT OF CARPETING	☐	☐	☐	☐	☐
APPEARANCE OF PAINT JOB	☐	☐	☐	☐	☐
OVERALL INTERIOR ROOMINESS	☐	☐	☐	☐	☐
OVERALL INTERIOR COMFORT	☐	☐	☐	☐	☐
EASE OF GETTING IN AND OUT OF FRONT SEAT	☐	☐	☐	☐	☐
EASE OF GETTING IN AND OUT OF REAR SEAT	☐	☐	☐	☐	☐
USABLE TRUNK SPACE	☐	☐	☐	☐	☐
SMOOTHNESS OF RIDE	☐	☐	☐	☐	☐
HANDLING EASE ON HIGHWAY	☐	☐	☐	☐	☐
FRONT VISIBILITY	☐	☐	☐	☐	☐
REAR VISIBILITY	☐	☐	☐	☐	☐
EASE OF STEERING	☐	☐	☐	☐	☐
PICK UP WHEN PASSING AT OVER 45 M.P.H.	☐	☐	☐	☐	☐

1P-72310

UR CAR AS HIGHLY
R RATES HIS?

If you put your car to the test on the left and discover it rates a lot of "fairs" and "poors," you should give some thought to owning a Volvo.

Recently, Volvo owners were sent a similar questionaire as part of a nationwide survey among new car owners. They were asked to rate their new Volvos on these, and several other points.

The results were very enlightening.

For example, when it came to quality of workmanship (something many people have stopped trying to find in a car) 86.9% of the Volvo owners rated their cars "excellent" or "very good."

Ride in a Volvo and you'll understand why. There's a definite feeling that this car has been fit together instead of slapped together. A strong, unitized body construction helps eliminate squeaks and rattles. And if you're old enough to remember when a paint job was actually a feature you could point to in a new car, you'll admire Volvo's. It's five coats deep.

Volvos also received high ratings for what their owners could fit into them. Namely, themselves.

Inside, Volvo has enough leg, hip, shoulder and headroom for five six-footers. Up front, they'll be sitting in Volvo's famous orthopedically-designed bucket seats. Which adjust in ten different directions, including from "soft" to "firm" against the small of your back. In the back of a Volvo, you'll find a trunk with room inside for, believe it or not, a trunk.

These are just a few of the reasons why Volvo owners gave higher ratings to their cars' comfort and roominess than Cadillac owners gave to Cadillacs.

Volvo owners also like the way their cars handle. Volvos have rack and pinion steering. Which, together with an advanced spring-strut suspension makes for a smooth, stable, controlled ride. Volvo also has power-assisted disc brakes on four wheels instead of just two, for quick, straight-line stopping.

Add to these Volvo's other qualities—the exceptional level of performance provided by fuel-injected overhead cam engines, Volvo's legendary strength and durability, and safety features so advanced they're being studied by the U.S. government—and you understand something else.

You understand why, in this same owner survey, Volvo owners overall were shown to be more satisfied than the owners of 48 cars from G.M., Ford, Chrysler and AMC.*

Your Volvo dealer has a full selection of sedans and station wagons to choose from. In either our 240 series or from our luxurious Volvo 260's.

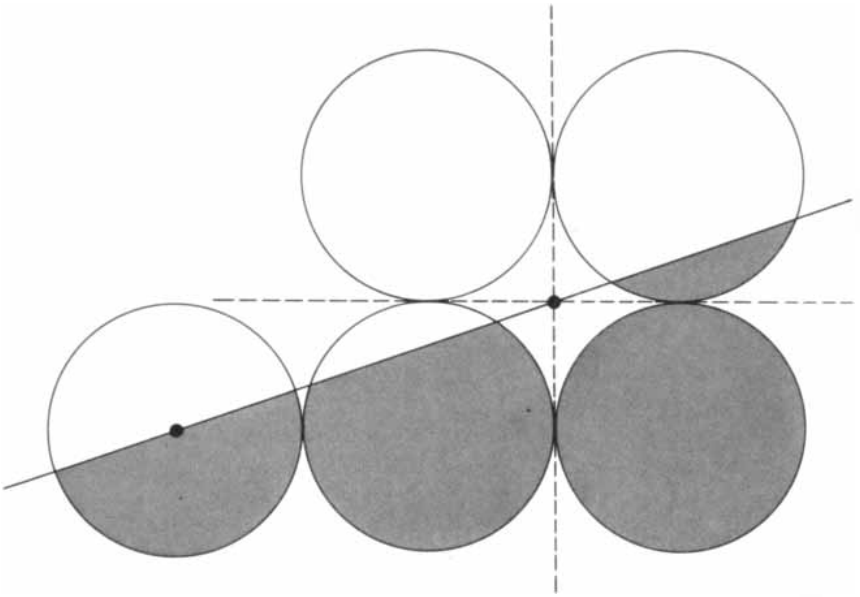
Volvos start at $6,645.†

Which to some people may seem like a lot of money to pay for a car.

But to others, it's a small price to pay for satisfaction.

*Survey conducted among owners of new cars bought in May, 1977. †Suggested retail price P.O.E. local taxes, dealer preparation, delivery charges and Lambda Sond™ units additional.

VOLVO. A CAR YOU CAN BELIEVE IN.



An improved solution to a bisection problem

elty than a monstrous earthquake that suddenly snuffs out tens of thousands of lives? Unfortunately next month's California quake will be relatively mild."

"According to *The Jupiter Effect,* by John Gribbin and Stephen Plagemann," I said, "Los Angeles will be totally destroyed by a quake in 1982."

"Sheer balderdash," snorted Dr. Matrix. "Gribbin and Plagemann think the quake will be triggered indirectly by Jupiter's gravity when it is reinforced by the gravity of all the other planets lined up on the same side of the sun in 1982. Jupiter's gravity is supposed to alter the sunspots, which in turn are supposed to alter the composition of the earth's atmosphere and produce the quake. Did you notice that Gribbin and Plagemann's book has no diagram showing how the planets will line up? That's because they won't line up. They will get to the same side of the sun—Gribbin and Plagemann are right about that—but they'll be so dispersed from a straight line that a picture would demolish his thesis. Besides, there's little evidence that the planets can affect sunspots, and no evidence that sunspots can affect the earth's atmosphere enough to disturb the San Andreas fault. Gribbin and Plagemann's book is garbage. According to our cockroaches, Los Angeles won't be destroyed until a major quake hits it in the prime year 1987."

The number 1,987, Dr. Matrix noted, is a remarkable prime. If we look for primes with positive digits in a cyclic, descending, consecutive order, 1,987 is the lowest example after 43 and the highest before 76,543, the largest-known prime of this type. If 0 is allowed, there is 109 and 10,987. Primes with digits in cyclic ascending order are commoner. Nineteen are known, starting with 23, including 23,456,789 and 1,234,567,891, and ending with the astonishing 1,234,567,891,234,567,891,234,567,891,234,567,891 which was discovered in 1972 by Raphael Finkelstein and Judy Leybourn of Bowling Green State University. Can the reader prove there are no descending consecutive-digit primes, with or without 0, that start with 9?

Dr. Matrix reached into a desk drawer, pulled out a small paper-covered booklet and tossed it across the desk. Persi and I twisted our necks to read the title: *Alphabetic Number Tables, 0–1000.* The preface begins: "It gives us great pleasure, not unmixed with profound emotion, to at last make public these alphabetical tabulations of the natural integers, ordered both in English literation and in Roman numerals." The booklet had been published at the Massachusetts Institute of Technology on April 1, 1972.

"I compiled both lists for some friends at M.I.T.," Dr. Matrix said. "You may keep the monograph if you like. It's now quite rare. I give it to you because it suggests some interesting problems."

I opened my notebook again and raised a pencil.

"The alphabetical list of the English spellings for the integers 0 through 1,000 begins eight, eight hundred, eight hundred eight, eight hundred eighteen, eight hundred eighty and so on. The last entry, of course, is zero. How many of your readers can name the 1,000th, or next to last, number on the list?"

"Delightful," I said, scribbling a note.

With the back of his hand Dr. Matrix flicked away a large cockroach that was crawling across his desk. "You can ask a similar question for the Roman numerals," he continued. "The digits themselves are now the letters. The sequence begins: C, CC, CCC, CCCI, CCCII and so on. The Romans had no zero, and so the last number is the 1,000th one. Can your readers determine what it is?"

Dr. Matrix paused until I had finished writing, and then went on: "The English spellings of numbers suggest strange problems. For example, consider the first letters of the 10 digits. What is the longest English word that can be made with those letters? You needn't use all of them, and any letter may, if you like, be used more than once. I think the longest such word is 'festoons.' Perhaps your readers can find a longer one. Another interesting question. What's the smallest number whose English name contains all five vowels plus *y?* Assume, as I did when I compiled the M.I.T. list, that 'and' is not a proper part of any number name."

"Excellent," I said. "Any others?"

"Numerological problems are as infinite as the primes," said Dr. Matrix. "But I'll give only one more. It's the invention of my friend Joe Wagner of New Wilmington, Pennsylvania."

Dr. Matrix picked up my notebook to jot down the following sequence: 10^3, 10^9, 10^{27}, 10^2, 10^0, ...

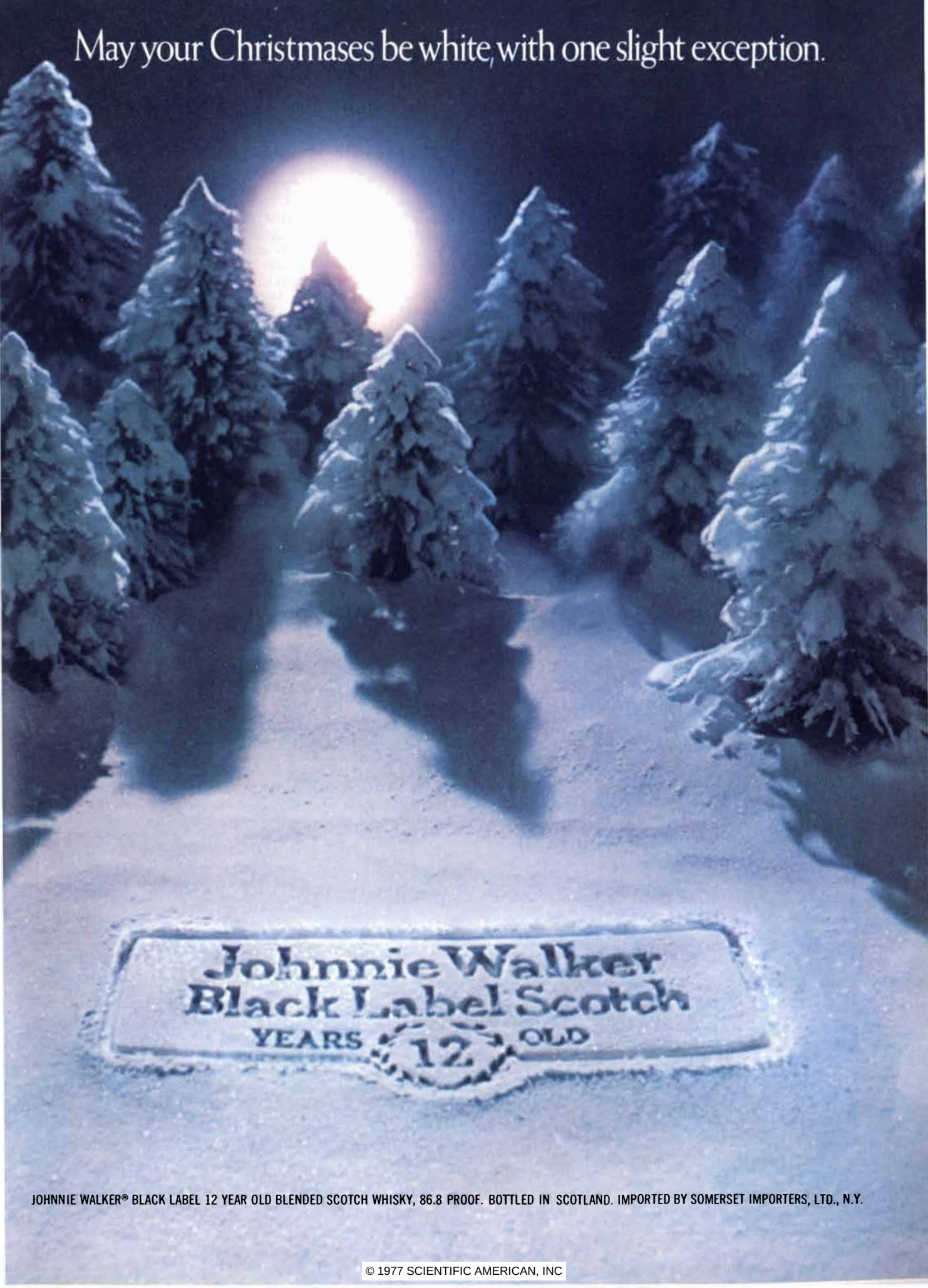
The problem, he explained, is to determine the exponent of 10 that gives the next number. The spellings of the numbers determine the pattern.

"I've been fascinated by those three cubes on your desk," Persi said, pointing to three large cubes on a tray at the top of a desk calendar. Each face of a cube bore a lowercase letter. The cubes were in a row so that the side facing Dr. Matrix spelled "nov."

"I remember," said Persi, "that in one of his columns Martin asked how to put digits on the 12 faces of two cubes so that they could be placed side by side to give any day of the month [see "Mathematical Games," SCIENTIFIC AMERICAN, April and May, 1969]. Is it really possible to turn and rearrange these three cubes to spell the first three letters of any month?"

Dr. Matrix replied by extending 10 bony fingers and rapidly manipulating the cubes so that they ran through the standard abbreviations for all 12 months. It is a pretty problem to figure out how the lowercase letters can be distributed on three cubes, one letter to a face, to make this possible. I later learned that Dr. Matrix' friend W. Bol of Geldrop in the Netherlands had solved the problem and sent the cubes to him in 1971. Next month, when I answer all the other questions, I shall show how to do it.

Dr. Matrix consulted a Mickey Mouse watch that dangled at the end of a long dog chain pinned to a shoulder of



JOHNNIE WALKER® BLACK LABEL 12 YEAR OLD BLENDED SCOTCH WHISKY, 86.8 PROOF. BOTTLED IN SCOTLAND. IMPORTED BY SOMERSET IMPORTERS, LTD., N.Y.

his T-shirt. "It's almost one-thirty," he said, pressing a button on the side of his desk. The air was instantly shattered by the voice of Johnny Rotten, the Irish punk rock celebrity, snarling out the lyrics of his latest British hit, backed by the stabbing, quaking three-chord beat of his band, the Sex Pistols. Iva materialized at the door.

"Where shall we all lunch, my dear Punky?" Dr. Matrix asked.

"There's a new place on the Camino," Iva replied, "called the Punk Elephant. The food is great, and they serve nonpunk cocktails."

"Will I have to listen to punk rock?" Persi asked apprehensively, cupping his hands over his ears.

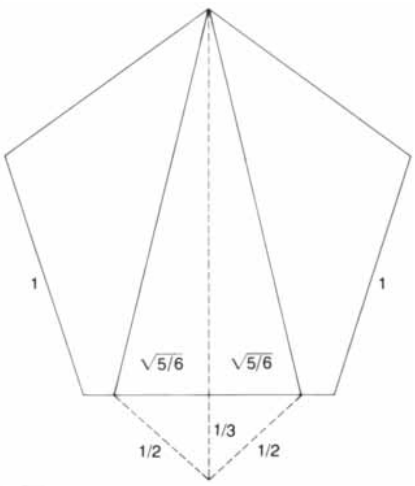
"No," said Iva. "There's a punk girl group appearing there called the Palmdale Bulges, but they don't start until nine in the evening."

Iva vanished for a moment, then returned with two long black leather overcoats. She handed one to Dr. Matrix, who had walked over to stand facing her. And then we saw an astonishing thing.

Each of them held an overcoat in front of the other, then each simultaneously let go of the coat with his (her) left hand and pushed his (her) left arm through the left sleeve of the coat opposite. Still facing each other and acting in unison, each of them then put his (her) left arm around the other's right side and with his (her) left hand grasped the top of the coat as the right hand carried it behind the other's left shoulder. Each now dropped the coat from the right hand, carried the coat around the other's back with his (her) left hand, then each pushed his (her) right arm into the right sleeve. Now they were wearing their own overcoats, and they were still facing each other. The ritual was executed so rapidly that it was over before Persi and I could observe just how they did it. Iva told me later, when she helped me to write the above instructions, that it was an old Japanese vaudeville routine.

As we walked past the entrance to the laboratory I sniffed the air. "I smell punk," I said.

"Shhh," said Iva. "If you don't talk about it, maybe no one will notice."



Trisecting the area of a pentagon

The Ramsey graph theory puzzle given last month consisted in two-coloring the complete graph on seven points so that there are no black triangles and no more than four triangles in color. A solution (note its color symmetry) is shown on page 18.

My column for July dealt with dividing figures into two or more parts. It brought a number of letters on which I want to comment briefly. The solution I gave in August to the problem of bisecting the total area of five circles required the construction of a sixth circle. Ismor Fischer, John Graves, George K. Holland, Warren Lane, Barry M. Peterson, Robert Windschitl and Louise N. Worrell were the first to send the simple construction shown in the illustration on page 22 that does the job without a sixth circle. The proof of bisection is obvious.

I stated in the August column that I knew no easy way to trisect a pentagon with lines extending from one corner. Carl F. Von Meyenfeldt of Vancouver supplied the trisection shown in the illustration on this page. Assume that the pentagon has sides of length 1.

After proving that a pair of perpendicular lines divide the area of any figure into quarters, I called attention to the difficulty of constructing such a pair of lines for the right triangle of sides 3,4,5. Von Mayenfeldt was the first to send the solution and to prove it unique. The same construction, too complicated to give here, was later sent by Windschitl and Cecil G. Phipps.

Although no reader found a general algorithm for solving all problems that involve cutting any shape into two congruent halves (when possible), many readers found procedures that apply to many such problems. For example, such problems can sometimes be solved by tracing the shape to be divided on transparent paper and then turning the tracing to intersect the original figure in various ways. Abel Bomberault, Robert W. Davis, James R. Fienup, David Fleming, Allen J. Schwenk, Daniel Sleator, Laszlo M. Vesei and Marcel Vinokur were the first to send suggestions of this kind.

In spite of the many errors in the cipher I gave in August—the one Edgar Allan Poe could not solve—many readers were able to crack it. So far correct decodings have been received from John Ambruz, James H. Andres (age 16), Alvin R. Crook, Robert Hart, A. F. Milne, James Prescott, Roger Richards, Michael Sanford and John Sladek.

We help process one-fourth of all the oil and gas the world uses every day.

And that's being accomplished by just one member of Combustion Engineering: C-E Natco.

C-E Natco helps de-salt, de-foam, de-water and do other processing needed to prepare oil and gas for the pipeline.

And, besides processing it, we also help drill for it, extract it, pump it, clean it, deliver it and refine it.

Gray Tool Company, a new member of the C-E family, pioneered the first wellhead assembly able to operate at pressures of 30,000 pounds per square inch—extending the search for oil and gas into earth layers too difficult to control before.

C-E Crest has designed systems that produce 10-million barrels of oil a day.

C-E Lummus engineered over 30% of Europe's refining capacity. And is responsible for the design and construction of deck facilities for off-shore platforms in the North Sea that can withstand 130-mph winds and 100-foot-high waves.

We do more than help process oil and gas.

For over 60 years, we've been supplying equipment, technology and research to help produce energy, conserve it, squeeze more use out of it, even convert it to new forms.

For more information about our activities around the world, write Combustion Engineering, Inc., Dept. 7008-33, 902 Long Ridge Road, Stamford, Connecticut 06902.

COMBUSTION ENGINEERING

The Energy Systems Company

BOOKS

An annual survey of children's books on science and technology

by Philip and Phylis Morrison

The Animals

WILD MOUSE, by Irene Brady. Charles Scribner's Sons ($5.95). "A wild mouse—smooth, sleek" darted across the top of the sink. The little animal was subtly caught, both brilliantly in the deliciously mouse-colored, mouse-sized watercolors of this sensitive artist-observer and in her net of adjustments and empathy. "Two weeks of constant coaxing," the woman quieter than the mouse, and he would eat sunflower seeds in her hand. ("Delightful feeling, those tickly toes on my palm!") He fitted the field-guide account to a T: the white-footed mouse, out there in an Oregon country kitchen. The easily assumed male pronoun soon did not fit the real mouse. Daily more tubby, nest-seeking, the mouse impelled a closer look: "He is a she!" Then we watch the miracle of the births, bitten umbilical cords, the placenta rolled back like a sock, three red mouselets, transparent enough to reveal the white milk in the small stomach and the dark pupils through sealed eyelids. Closely watched, lovingly and tightly drawn, the baby mice nurse, grow and change over a fortnight. Before the last page the little mice collectively outweigh their mother, and at the close they scatter, "tiny feet" in the attic. The family has grown up across two dozen pages, each one with sketches whose fitness is a delight, and with a personal, sharply seen text a few sentences in length. Our literature is rich enough in fiction about small rodents, from Beatrix Potter to E. B. White; this book of the real is as compact, engaging and instructive as any of them. Very likely it is to be a classic, equally apt in picture and in word, for both the reader and for the read-to.

THIS BROKEN ARCHIPELAGO: CAPE COD AND THE ISLANDS, AMPHIBIANS AND REPTILES, by James D. Lazell, Jr. Photographs by Martin C. Michener. Quadrangle / The New York Times Book Co. ($12.50). Dr. Lazell is a wildlife biologist for the Massachusetts Audubon Society and a man of distinct character and deep knowledge and love of New England life. He has long used "the most useful known collecting device on earth . . . called *student*." His broken archipelago, which echoes a phrase of Darwin's, is Cape Cod and its islands. In a couple of original and knowing chapters he outlines for us the origin and nature of this land and the sea around it (not forgetting double tide waves through Vineyard Sound). Then he is off, to a detailed account of the salamanders, frogs and turtles of a little, beautiful and diverse world, an account full of remarks, both fresh and salty, on his biological colleagues and their collections, their books and their foibles.

Frog distribution is no easy problem; it is traditional that "frogs are carried around by children of every race, without record." There is not much doubt that every little spadefoot is a miracle to shatter infidels, but he tells us two true tales no one could have expected. First, the largest living reptile in the world is common in these waters: the great leatherback sea turtle, and what a beast it is! There are ton-sized specimens near the Cape islands. They are found tangled in lobster-pot lines, held by the right flipper. Their prey is the big subarctic jellyfish, and they strike in error at the algae fringe on lobstermen's old buoys in a systematic fashion that can be understood. The animal that lives on such a dilute food as jellyfish has no need for a stomach; the intestine does all the work, since there is little to store but seawater. One pass down a straight gullet and a short gut lined all the way with ripping spikes does it. A second marvel is the rodent-hunting king snakes that abound on Nantucket, some acres holding 200 of them. "I have never seen so many snakes . . . outside of the Tropics."

This is a book of singular experiences by a special man. One could hope for more tolerance for the efforts at quantitative zoogeography, but let that pass. (The author does not spare learned taxonomists either.) No literate person who feels at home with the four-footed, cold-blooded classes should miss this zestful, map-rich guide to life on the Cape and living it.

ORANGUTAN: ENDANGERED APE, text and drawings by Aline Amon. Atheneum ($7.95). The chimpanzee and the gorilla have a pretty good press these days. In the deep rain forests of Borneo and Sumatra, however, live a few thousand of our other cousins, the dexterous, fruit-loving, red-furred "forest persons," with not much visibility. Slowly a new interest is growing in the orangutan too. Today the interest is in the animals in their home; in the 19th century the interest was in their skin and bones, more recently in their feats in the zoo and the laboratory (where they "explore the equipment instead of solving the problems . . . more interested in the operation of the bolt than the food inside the box").

We, the most terrible of primates, have long been the orangutans' enemies. They are gone entirely from populous Java; in Borneo remains from the Stone Age reveal cave-dwelling human beings who made orangutans a favorite meal. They are our kinfolk for all that. Most striking are the differences among individual orangutans; they may have a kind of culture. Those in Sumatra, where other big animals are found, have a "warmer family life," the big males sometimes being seen living peaceably among females and infants; in Borneo, where there are no such threats, the big males are solitary and fierce, social only for mating. They differ as individuals too; the author argues quite plausibly that the intelligence differences among the three big ape species may be less than the spread among individuals. There are clearly brighter and duller orangutans and chimpanzees.

The structure of this book is unusual. Alternate chapters present a fictional but careful account of the life of an orangutan mother and child growing up in the forest, and the chapters in between give nonfictional accounts of the history of our relations with the species, ending with the present day. The current situation is precarious in the decline of the wild forests but hopeful in the story of people who live among the apes, some trying to rehabilitate young orphan orangutans confiscated from illegal owners who made them pets. The apes are being trained for forest life; quite a few are not eager to give up those free bananas and clever friends.

DISCOVERING THE WORLD OF THE THREE-TOED SLOTH, by John Hoke. Photographs by the author. Franklin Watts ($4.33). "It looks like a little person." So said the three-year-old son of this author as the sack opened to show the first three-toed sloth (the other extant sloth, the two-toed, is quite a different species) to enter the family. They were Americans living in Surinam, where the three-toed sloth is common and is called by a name meaning "lazy man" in any one of five languages. The beast is remarkable, and this story of what kind of pets they make is bound to attract many a young reader. They do not ("indeed, they cannot") make fast

SEAGRAM DISTILLERS CO., N.Y.C. 80 PROOF. DISTILLED DRY GIN. DISTILLED FROM GRAIN.

Mercedes-Benz announces a new Coupe

The 2.8-liter gasoline-powered 280CE.

plus an extraordinary new Coupe

The 3-liter Diesel-powered 300CD.

The new Coupe
is the 280CE and it is powered by a fuel-injected, 2.8-liter Six.

The extraordinary new Coupe
is the 300CD, and it merits the term by introducing Diesel power to a class of Mercedes-Benz from which Diesels were previously excluded.

Thus the Mercedes-Benz Diesel has fully come of age. It stands today at such a peak of refinement that a limited-production Diesel sports Coupe is not sacrilege but logic. Meanwhile, savor your new luxury of choice.

No boulevard sports car

Gasoline- or Diesel-powered, this new Coupe is no faddish "personal" car but a lithe five-passenger machine you will prize as much for its quick reflexes as for its elegant lines.

You sit cradled within a wheelbase of only 106.7 inches, less than a foot longer than the close-coupled 450SL Roadster. You should find yourself negotiating even tight mountain curves with a sense of ample road to spare, slipping through city traffic like a fish. Chances are, within a few days you will stop having to envy sports car drivers forever.

But family men can rest assured. There is not only a full rear seat but a rear seat that nearly equals the 280E Sedan for headroom. And there is not only a full trunk but a trunk with the same 12.6 cubic feet of space as the 280E.

Some people don't understand how Mercedes-Benz builds such space-efficient cars. Mercedes-Benz doesn't understand how to build anything else.

Worth its weight in strength

The new Coupe's taut profile is aerodynamically "clean." Put under an X ray, it would be revealed as the skin over a steel cage welded at 4900 points to a steel platform, forming a monocoque structure.

The engineers braced that sheet-steel floor with cross members to help stiffen the midsection. And although the side windows are pillarless, the concealed "B" pillar provides additional structural integrity.

Straight-line braking stability

Engineers and not stylists are the heroes behind the Coupe. If this 1¾-ton, five-passenger machine seems to corner flatter and grip the road more tenaciously than you're accustomed to, credit them.

They in turn would credit technological nuggets like fully independent suspension, and *zero offset* steering with straight-line stability, *even* in panic stops.

Panic stops are further softened by 4-wheel disc brakes. Mercedes-Benz would never build a Coupe or any other car without them.

But *handling brilliance* is only half the tale. The other half is the ride: Over-the-road decorum worthy of an expensive motor car. *A clue:* The shock absorbers' lower chambers are pumped full of *nitrogen* gas as an extra cushion against vibration.

Power assists a purist can love

No brute force is needed to pilot your Coupe about. Power steering, power brakes, and a 4-speed automatic transmission slightly smoother than the human hand are standard features. Wives will approve. And so will driving purists—because these laborsaving devices have been designed to preserve driving "feel" at all costs.

Not a superficial ounce of chrome

Doors over four feet wide swing open to reveal a meticulously ordered cockpit. Electric windows, bi-level climate control and AM/FM stereo radio are all built in, but the mood is less opulent than crisply businesslike.

For instance, the interior has no dazzling chrome knickknacks to confuse the driver's field of vision. Glare free gauges are placed so you can check them without breaking visual contact with the road.

The anatomically contoured twin front seats are thickly padded, their springs tuned to the car's suspension. The backrests lock in place *pneumatically.*

And the walnut-root veneer that trims the instrument panel and covers the console is *real* walnut-root veneer.

A delicious dilemma

Which engine version belongs in your Coupe?

Classicists will feel most at home with the 280CE and the familiar traits of its gasoline power plant—supple and responsive in the sporting Coupe manner. And hardly a mundane choice: The 280CE marks the zenith of Mercedes-Benz six-cylinder engine development.

The 300CD allows Coupe devotees to combine their love for elegance with common-sense motoring, and Diesel devotees to combine their love for common-sense motoring with elegance. The power of its 5-cylinder engine is one surprise, the lack of noise and roughness is another. That it offers the established advantages of a Diesel should be no surprise.

But there is ultimately only one way to decide which Coupe best suits your needs. See your Mercedes-Benz dealer and experience them both.

motions. They live in forest trees alone and with no nest, simply by never attracting attention. They move without a sound, always in slow motion. They ooze upward into the trees to seek refuge in branches too small for pursuit by a big cat. They climb like telephone linesmen, their big claws on the branches. They cannot defend themselves with those claws, which can swing only in slow motion and seldom connect. In their coats live green algae, the growth of which is controlled by wingless moths that feed on them.

The family came to understand the pet sloths quite well in the circumstances of the home. The sloths always buried their wastes in the soil, very slowly to be sure, when they were let out every four or five days. They do the same in the wild, sometimes. They usually relieve themselves in the high tree canopy, the little black pellets making an audible sound on striking the ground; they do so only in the rain, when the sound is lost among the pelting raindrops. After a rare spell of rainless days the sloth may risk the descent to the ground. Apparently that is safer than making a noise that locates the animal in the forest. The famous *ai-ai* cry that gives them their name in one more language is uttered only in crisis.

The book is written particularly for early readers, perhaps too self-consciously so. It is an engaging story of a reflective experience with an extraordinary kind of pet.

THE BEE BOOK: THE HISTORY AND NATURAL HISTORY OF THE HONEYBEE, by Daphne More. Universe Books ($12.50). A NATURAL HISTORY OF TERMITES, by Frances L. Behnke. Illustrated by Turid Hölldobler. Charles Scribner's Sons ($6.95). The insect societies grow in interest as we come to learn more about how they work. These are two excellent, well-illustrated introductions, each to one of the better-studied structures. *The Bee Book* (English in origin, tone and allusion) is in no way meant for children, but its catholicity and liveliness should suit them very well. Bees frequent the green fields and feed on nectar, of course, to yield honey and wax—delicious, beautiful and useful. Termites, quite otherwise, are blind, condemned to darkness, harmful to human economic concerns, even noisome. Their value is a value to the mind alone: their societies are extraordinary in capabilities and control mechanisms. The termite volume was written for serious young readers; its text is up to date and its many line drawings are particularly instructive.

Clearly *The Bee Book* is free to begin with a review of the mythology of bees and to end with a survey of the uses of beeswax through history, including coatings for writing tablets, modeled cores for casting metal, candles and even death masks. Mme Tussaud, before her marriage a prisoner in Paris under the Terror, was set to modeling in her workshop-cell the death masks of those whose heads fell to the guillotine. Out of this gruesome experience came the wax museums that are still popular. About half of the bee book describes the bees themselves: the queen, or mother, bee, the drones and the workers, with their duties and behavior. Half of the book treats of the product of bees, the history of beekeeping, the rise of scientific understanding out of folk practice and the like.

The dark termites have no such context to set them off, but their agriculture, their architecture, their ruminantlike symbiosis with protozoa that enable them to subsist on wood and their great diversity are all strikingly set out. No reader should idealize the social insects. The queen bees seek out and break open the cells of their unborn rivals, who have announced their location in the comb by responding with a special piping noise. The workers then make sure of an undisputed succession. Termites groom one another constantly, and once in a while a rip may be made in the soft portions of their skin. Such a wound is compellingly attractive. Nestmates gather and come closer, until one termite chances to pick at the rip. Finally one starts eating; suddenly all join, until nothing is left of the injured citizen. To be sure, termite colonies are usually short of protein.

The Plants

GRAINS: AN ILLUSTRATED HISTORY WITH RECIPES, by Elizabeth Burton Brown. Prentice-Hall, Inc. ($8.95). The staff of life has been wrought with great strength and beauty in many lands. It is hard to imagine a more instructive opening for young readers to the diversity of human cultures within the unity of our species than this unusual book. From chapati to triticale, the book includes about 100 recipes based on a grain staple, each recipe associated with a chapter of interesting geographical, historical and artistic material relating the food to the people who originated it. We cover a wide gamut of space and time, of nomads and palaces, of modern elegance and antique simplicity. The economic botany and prehistory are sound and interesting, the evocation of other times and places is convincing, with many unusual illustrations (say a battle between Mongols and Russians, or a Ming festival on the river).

The good food from so many lands—not all of it easy to copy but all at least imaginable—adds a sense of verisimilitude not easy to come by through books alone. One welcome novelty is the West African section, built on the rather complicated topic of millets. The groundnut stew is good, but it would also go down well with a nongrain side dish: the yam, the great root-crop staple of the moist Tropics. Here is a profounder context for the waffles and pizza, the cornbread and tamales, the commonplace indulgences of everyday life. Any conversation at the family table that can take off from this basis is bound to be apt and civilizing yet never precious. The work is generally well informed, and it includes a welcome page on milling and a fine page or two of world production statistics. Only for white flour and polished rice is the treatment much too simplified.

POPCORN, by Millicent E. Selsam. Photographs by Jerome Wexler. William Morrow and Company ($5.95). With this book a determined grade school investigator can sample the remarkable botany of maize. Seed is easy to come by: about half of the kernels in a tin of popcorn from the grocery store will germinate. There are three steps to a useful crop, and any progress is rewarding. First, germination, which is easy to do in a drinking glass in the kitchen; the book tells just how to work the magic and to see the results. The next step is growth, which requires a dishpanful of garden soil, a place outdoors and a few summer weeks of care. Last comes the pollination step, fascinating for its direct relation to the striking structures we all know on the corn plant: tassels, silks and grains on the cob. All the changes are shown in clear photographs, some in color, some magnifying. The reader who remains only a vicarious maize farmer will still gain, since the small book also touches on the history of the crop and its economic diversity and importance.

About Living Things

POISONS: ANTIDOTES & ANECDOTES, by William Tichy. Sterling Publishing Co., Inc. ($5.95). The chemist author of this stout little volume knows both how to spin a weird, wondrous tale and how to explain in a simple way a key result of biochemistry. As it should, his discussion brings out the central strangeness of poisons: a little monkeywrench can wreck a big and powerful machine. More literally, a tiny dose of saxitoxin—the "red tide" toxin—can kill a man who weighs 500 million times as much as it does. The explanation lies in enzymes and their delicate protein "loops and twists" that are so sensitive to reactants. Most poisons strike at an enzyme, and there are so many molecules in any tiny dose that they can reach every enzyme molecule present, since enzymes themselves function in very small amounts. Certain poisons, such as botulin and curare, block some step in the molecular process by which nerves signal muscles.

Ordeal poisons, blowgun darts, the Borgias and the arsenicists are the grisly

Technics introduces cassette decks that only sound expensive.

You can build a $200-$250 front-loading deck to look expensive. Or you can build one the way Technics does—to sound expensive. Without those meaningless knobs and gadgets on the outside. But with Technics high-priced, high-performance technology on the inside.

Technology that makes Technics RS-615US and RS-630TUS sound a lot better than you'd expect a mid-priced deck to sound.

Both decks give you inaudible wow & flutter: 0.10% RMS for the 615, 0.09% RMS for the 630. The reason: The kind of electronically controlled DC motor found in our more expensive decks.

And instead of tape hiss, you'll hear music, even in soft musical passages. Because Technics low-noise circuitry in addition to Dolby* give both decks a distinctively expensive S/N ratio: −63 dB (CrO_2 tape).

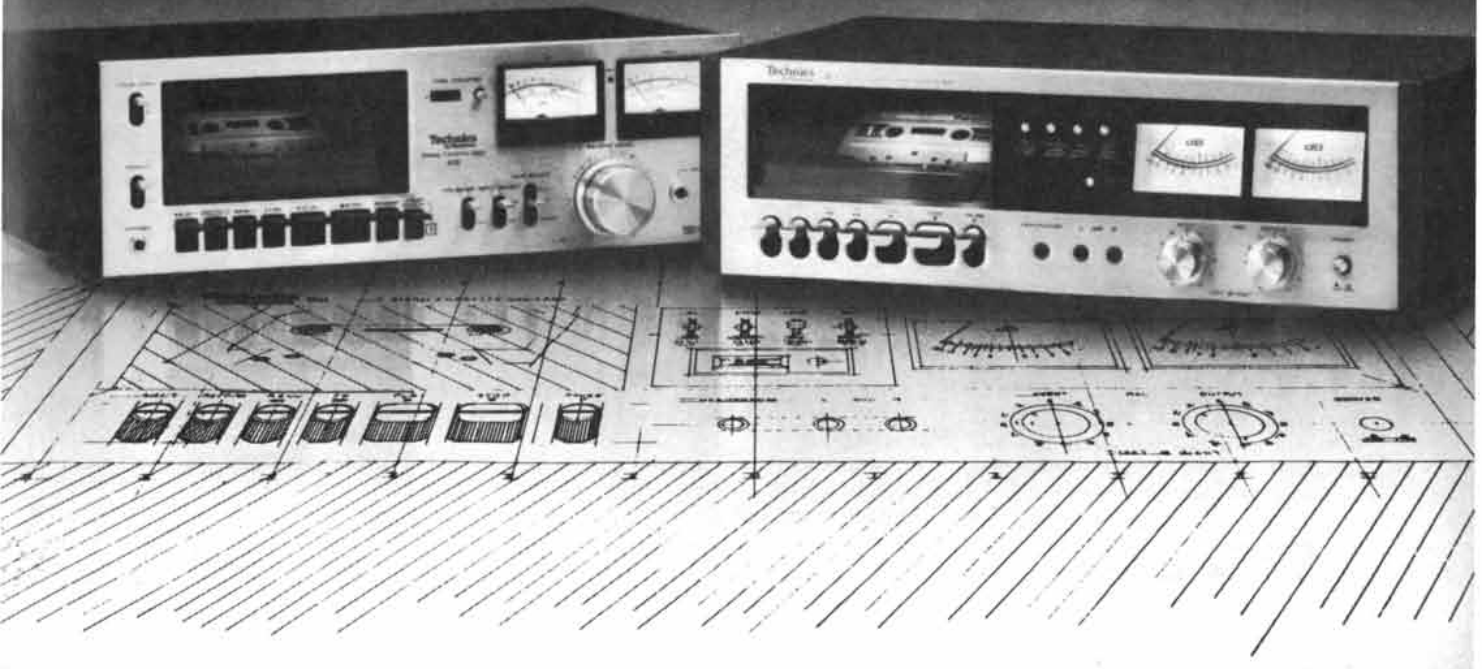
Both decks also give you long-lasting super alloy heads. Oversized VU meters. Lockable pause control. And automatic pre-timed recording or playback with a standard timer (not included).

The 630 adds a dual-output control. Separate bias and EQ switches to get the most out of normal, CrO_2 and ferrichrome tapes. And a peak check meter to help avoid overload distortion.

The RS-615US and the RS-630TUS. Everything about them sounds expensive. Except the price.

*Dolby is a trademark of Dolby Laboratories, Inc.
Simulated wood cabinet for RS-630US.

Technics
by Panasonic



topics of the long historical first chapter. Then comes a biochemical interlude, with the rationale of the only approach to a true universal antidote: adsorption on activated charcoal, either internally or in the course of the external passage of circulating blood over a filter. Next the poisonous beasts of the world are mustered, headed perhaps by the aggressive taipan of Australia, a snake whose bite has yet to be followed by human recovery. Then the poisonous plants are described, from the hemlock given Socrates to the atropine of belladonna. (The enlargement of the pupils induced by a small dose of this drug applied to the eye is the topic of a fascinating box that cites experiments showing that an enlarged pupil denotes emotional involvement. Poker players watch for that sign, and professionals, such as Chinese dealers in jade, often wear dark glasses to conceal their own reflexes of interest.)

All this dark romance and intrigue, nicely cultivated in Tichy's wry phrases, culminates in the story of poison-control centers today, their goals and methods. A sheaf of tables on antidotes and poisons and where to seek more information and help ends a book that ranges from the unreasoning and the deadly to the matter-of-factness of public health. It is a fine mix for many a stouthearted young person, from the upper grades on, both in the mode of darkness and in that of light. The work is quite reliable, although no effort at all is made to avoid the most sensational interpretation of all these good, grim but mainly accurate tales.

VITAMINS, by Alan E. Nourse. Franklin Watts ($4.47). An experienced science writer who was once a practicing physician here presents a clear sketch of the vitamins. A reasonable if distant analogy is suggested for the work of enzymes and coenzymes, and there are some chemical formulas. All the vitamins of the Recommended Daily Allowance tables are discussed: their discovery, the results of deficiency, their various sources. In a very brief account we can read of the triumphant failure of the experiments of Christiaan Eijkman with sick chickens (his controls sickened too, since all of them were being fed on polished rice), of James Lind and the able seamen who daily got two oranges and a lemon (the lime juice came later, for economy) and no scurvy, and the rest of nutrition's heroes. We see their portraits and also a few clinical before-and-after pairs.

The most surprising story is that of the blood-clotting requisite, vitamin K. It is rare in foods, but adults get quite enough from the normal intestinal flora; made by our symbionts there, it enters directly through the gut. Newborns, all of whom have sterile intestinal tracts, tend to be bleeders, so that the vitamin is normally injected into the mother's or the infant's bloodstream just before or after birth. The massive-dose enthusiasm, particularly for vitamin C, is viewed conservatively by Dr. Nourse, but not dogmatically. The list of references is brief but excellent, and it includes Linus Pauling's contrariwise book on vitamin C and the common cold. This small work is an uncommonly straightforward and compact summary for readers in the middle grades.

GROCERY STORE ZOOLOGY, by Joan Elma Rahn. Illustrated by Ginny Linville Winter. Atheneum ($6.95). Galen dissected the ape and the great Renaissance anatomists the human cadaver in order to learn the fabric of the body. The cunningly articulated skeleton and the roll call of the muscles in their Latinity still form part of our image of learning and healing. No such direct path to anatomy exists for the serious young person in, say, the upper grades, yet this knowing book has found an ingenious way around. The kitchen table of every meat-eating family is a theater for comparative vertebrate anatomy, from the chicken and the turkey through the pig, cattle and the rabbit to *sapiens,* by means of drawings and examination. That ham steak is a careful cross section of one thigh of a pig, and in it can plainly be found the named muscles gluteus primus, sartorius, biceps femoris and others, just as in the human leg. All of these are drawn in simple line diagrams and their functions are described.

Pig's feet are fine stand-ins for the bones of the human hand, and soup bones are probably preparations of the knee, ligaments and all. Kitchen knife, scissors, controlled use of saucepan and oven are the tools and techniques; supermarkets and perhaps places dealing with "foreign menus" extend the range of specimens. Everywhere the chicken is easy to get nearly whole, and the remains of a roast are the basis of the skeletal anatomy described part by part in this useful book. A raw fryer opens an easier path to the study of muscles than the Sunday dinner does. It is explained here, with all the comparisons needed, in a pleasing tone of shared interest. Patience is part of the game, since the young anatomist is not likely to plan the family dinners on Vesalian principles of order. Tables are included to help adjust science to the dinner table. Not everyone is attracted to this kind of experience; for those who are, however, the book is a pathbreaker.

THE MICROBES, OUR UNSEEN FRIENDS, by Harold W. Rossmoore. Wayne State University Press ($12.95). An entire menu of delicacies is found in this attractive book of microbiology: olives, whiskey, rum and gin, soybean cheese, preserved duck eggs, sausages and sauerkraut, cider, beer and ale, sourdough bread, rye and pumpernickel and French breads, wines and their matching cheeses, yogurt and its tart kindred, a Philippine pineapple compote, with brandy, cocoa or coffee to close. Bon appetit!

Each one of those foods owes its special nature to one or more specially cultured microorganisms, which live and usually die in the foodstuff because of the cunning manner of its preparation. Even coffee, which requires two distinct fermentations to remove the skin and the fleshy pulp from the ripe berry, and cocoa, whose big fruit is fermented away by a sequence of yeasts and bacteria, owe essential nuances of aroma to these processes.

The domestication of microbes is an old art of human beings, and it is traced here from earliest times. Nowadays the process yields not only food and drink with savor but also penicillin, silage, tanned leathers, citric acid, a list of enzymes two pages long and other industrial products. A crystalliferous bacillus produces a diamond-shaped crystal, a protein that is highly toxic to caterpillars—say the gypsy moth—and offers much promise in the specific control of a number of costly pests. And in the end, if it is biodegradable, the microbes will get it. (Activated sludge is not a mere slogan but a flow chart.) The symbionts we all have with us are described, as of course are those of the root nodules of the legumes and of the rice paddy that supply a lot of the fixed nitrogen on which all larger life depends.

This is too interesting, unusual and personal a book for such a wooden title. Professor Rossmoore deserves plenty of readers.

THE CENTER OF LIFE: A NATURAL HISTORY OF THE CELL, by L. L. Larison Cudmore. Quadrangle / The New York Times Book Co. ($8.95). "All cell biologists are condemned to suffer from an incurable secret sorrow: the size of the objects of their passion." So runs the opening sentence of this little book, a set of connected essays on the study of protists, broadly viewed, by a young member of the suffering group. The sentence suggests the style: personal, engaged, provocative, even a little overstated. The author is witty and well-read; she has studied those cells at the end of the microscope with devotion and has found some ways to put us among them there, down in the Lilliputian world, "wild and beautiful . . . half Disney, half Dali."

The book begins with the distinction between the two great stems, those cells with DNA loosely wound and those with DNA neatly packaged in a nucleus. It surveys the remarkable properties of single cells, which come equipped with most of the possibilities of life (except fire and the wheel, the author remarks), including various kinds of motion, com-

You can tell a lot about an individual by what he pours into his glass.

The "Skier" glass created for the Bushmills Collection by Henry Halem

A blend of 100% Irish Whiskies. 86 Proof. Bottled in Ireland. The Jos. Garneau Co., New York, N.Y. ©1973

Bushmills.
The world's oldest whiskey.
What individuals have poured into their glass since 1608.

munication signals, genes and sex. All of these are topics for elaboration, and besides we read accounts of the origin of life and of the repeated symbioses—of protists that long ago learned to live together—that now form the powerful organelles of our proud eukaryotic cells. More fateful topics still, energy and death itself, provide the last two essays, a virtuoso display of prose weaving in the newest themes of biology. There are some stunning photographs as chapter heads that show remarkable protist forms, particularly one photograph of a cell that lays itself a mosaic pavement of hexagonal silica plates. (The author is not quite fair to neutrinos, ball bearings or silicon in a number of her asides.) No biology student of high school age or past it who has a taste for the written word will want to overlook these lively reflections.

THE PREHISTORIC WORLD, by Alan Bartram, Basil Booth, Michael Chinery, Euan N. K. Clarkson, Barry Cox, Diane Edwards, Christopher Maynard and W. D. Ian Rolfe; edited by Angela Sheehan. Franklin Watts ($6.90). An up-to-date exemplar of a fine old tradition, this beautiful, intelligently written encyclopedic treatment of paleontology for readers from the early grades to senior high is a bargain book of reference and pleasure. Of course there is a thrilling bright painting, across two pages, showing a big toothy *Allosaurus* chasing a mortally frightened little coelurosaur, and on another spread a *Tyrannosaurus* holding a hadrosaur by the throat. The paintings are obligatory, and rather freshly done here, but there is much more among the 500 illustrations. The frontispiece is a wonderful fish fossil in real color, from the Eocene.

Along the way one finds photographs of the oldest rocks known, the gneiss of eastern Greenland; a stromatolite from Shark's Bay; an account and a model of a lobe-fin walking; an old painting of Mary Anning, the discoverer of the first complete ichthyosaur; world globes with continents adrift period by period, and a plausible painting of a hungry little early mammal feasting on dinosaur eggs. The genre should be clear; the text material is accurate, simply written and logical in development. The La Brea tar pits and woolly mammoths come near the end of the book. There are eight pages devoted to the rise of mankind, perhaps too sketchy in themselves but a good foreshadowing of some other book. Richard Leakey and an old site in Kenya can be seen. The volume is part of a well-made series first published in Britain; the authors and contributors carry excellent credentials in the British educational and scientific world.

A WALK IN THE FOREST: THE WOODLANDS OF NORTH AMERICA, by Albert List, Jr., and Ilka List. Illustrated with photographs and drawings by the authors. Thomas Y. Crowell Company ($9.95). PYRAMIDS OF LIFE: ILLUMINATIONS OF NATURE'S FEARFUL SYMMETRY, by John Reader and Harvey Croze. J. B. Lippincott Company ($12.95). Ecology has topped out as an explicit topic for the didactic style of instruction. These two books convey the sense of the web of life less by example than by precept, by the analysis at some level of an ecosystem in being. The first of them is aimed directly at young readers, particularly those who might now and again be fortunate enough to take a long walk in the American woodlands. It begins with a general overview of what a forest is and how and where forests are found, and of the complex turnover within them. One device is used throughout: the text is a series of very informal questions and answers written at grade school level. The style is often aptly chosen: "Chlorophyll is the stuff that makes leaves look green, and the easiest way to describe its complicated work is to say that chlorophyll serves as a kind of antenna tuned to receive the wave energy of light." Occasionally it is mannered rather than helpful: "*What happens when lightning strikes a tree trunk?* Sometimes nothing happens, but occasionally... zap! millions of volts of electricity zip down the juicy cambium layer." The book follows carefully the life of the high canopy, of leaves, katydids and the flash of the tanager, the middle layers of woodpeckers and berries, the ground cover of moss, eft and deer, down to the soil of the earthworm and the star-nosed mole. A series of workable projects for study and a helpful list of the actual sizes of the forms pictured close a quiet and excellent book.

Pyramids of Life, the second book, is in the same vein but is played in a portentous tone to an adult audience, as the subtitle makes clear. Its authors are a well-known English photojournalist and an American ecologist with years of experience and responsibilities in Kenya. Their volume is large, rather imposing and dramatically illustrated, although it is "neither a scientific monograph nor a 'safari spectacular.'" Its theater presents three scenes: the great grasslands of the Serengeti; the rivers and lakes, sweet or soda, of the Rift, and the forests of that complex land.

Each scene is treated in a long series of striking photographs, every image made to accompany an explicative and unifying brief text. The scheme is to move around the cycle from the primary produeers that convert the sun's energy, through the grazers and their predators and back to the decomposers, humble but inevitable, that close all the chains. A few of the scores of vignettes will suffice as samples. The wildebeest herd in migration is first. Seeking the green grass as the rains vary, the animals respond to signals we do not yet know. Then there is the termite queen, a grotesque mother-sausage, constantly attended by king and servants, a source of chemical signals also still largely unknown. Here are the lesser flamingos, specialized filter feeders on the slimy, thick blue-green algae of alkaline Lake Nakuru. They crop the top inch of the well-stirred lake; when the blue-green algae fail (as they did in the dry year of 1974), the flamingos go too. The well-camouflaged chameleon is sharply shown, live-bearing to save the eggs from predators, the tiny young able to flick that projectile tongue at a flying insect even before they are quite dry. The golden-rumped elephant shrew appears in a superb shot, a six-inch killer eating its weight in insects and snails daily and loping long-legged and confident across the leafy floor. The text is manageable for any young viewer and good reader, who will gain even if the Serengeti remains a dream.

THE MATING GAME, by Robert Burton. Crown Publishers, Inc. ($12.95). The birds and the bees are proverbial analogues. This large and lavish book by an English biology writer carries out the theme at a popular level of description—this is no treatise on genetics—group by group, from the asexual animals of the reef through our species. The photographs are remarkable pieces of natural history, from the parthenogenesis of the fruitful green aphid to the coupling of snails, lions, apes and many more. Fascinating courtship displays and signals are included, many of them strikingly beautiful, but no lyrebird or impala is more elegant than the shadow close-up of a woman's face with deep and striking eye makeup. The last picture shows a couple sitting dreamily on a park bench. The book is a serious and tasteful, if explicit, resource on the entire topic.

HANGING ON: HOW ANIMALS CARRY THEIR YOUNG, by Russell Freedman. Holiday House ($5.95). The mother cat carries her kitten to bed by the nape. The photograph is endearing, but it is only the first of more than 20. Tiger, lion, raccoon and beaver mothers do the same; they are shown at it too. The kangaroo and the opossum have special gear, and the big joey peering out and the bee-sized opossums deep inside the pouch prove it. A koala bear and a striking albino brush-tailed opossum demonstrate that marsupials also carry the kids on their back when the right time comes. So do baboons and anteaters, swans and grebes. The wolf spider and the scorpion are not very cuddly, but each bears an entire passel of young ones on its back, all eyes and claws. The hairless bat baby clings to its flying mother's breast, and the sea otter, floating on her back, cradles her young one in her arms. This is a winning book of photographs and brief

Just read your blood pressure on the large easy-to-read meter of the Astropulse 10.

Home Blood Pressure Computer

A new space-age measuring system lets you easily check your blood pressure quickly and accurately without a stethoscope.

The new Astropulse 10 lets you quickly read your blood pressure without a stethoscope and without even rolling up your sleeve.

The system is based on three micro-electronic circuits and a microphone transducer. The microphone picks up pulses in your artery, and the integrated circuits measure the pulses and relay this information to a meter which visually displays your two blood pressure readings. There is no expensive stethoscope required, no guesswork or complicated steps to follow.

EASY TO USE

Taking your own blood pressure is quite simple. Just stick your hand through a self-tightening velcro cuff, slide the cuff up your arm, pull the tab and attach the tab to the velcro material. The tab will stick automatically without loosening. Then squeeze the rubber bulb to inflate the cuff, and take your blood pressure readings.

When a doctor reads your blood pressure, he uses his skill and a stethoscope to recognize your systolic and diastolic readings. Now a computer can do this in the convenience of your home and on a regular basis.

The Astropulse 10 also flashes an LED signal and an audible tone at the two blood pressure readings to assist the hard-of-hearing or those with poor eyesight.

DOCTORS ENCOURAGE USE

Knowing your correct blood pressure is very important. Statistics show that as many as 25 million Americans suffer from hypertension, yet only half know about it. Hypertension results in high blood pressure, and high blood pressure usually goes unnoticed until other symptoms of hypertension occur—often too late to correct.

The Astropulse 10 is so easy to use that it encourages regular blood pressure monitoring —exactly what doctors recommend. Even if your health has been perfect, hypertension and blood pressure can occur at anytime.

SOLIDLY BACKED

The Astropulse is powered by a readily-available 9-volt battery supplied with each unit. The Astropulse uses solid-state electronics so service should never be required. But if service is ever required, JS&A's prompt service-by-mail center is as close as your mailbox. JS&A is a substantial company selling advanced space-age products directly to consumers for over a decade—further assurance that your modest investment is well protected.

The Astropulse 10 was designed to take your own blood pressure in the privacy of your home. The cuff is easily tightened with the self-tightening bar and the velcro material. Just pull the flap and attach it to the cuff.

The entire blood pressure kit fits nicely in the carrying case supplied free with each unit. The carrying case measures 3½"x 4"x 7" and the entire system weighs only 20 ounces.

To order your Astropulse 10, simply send your check for $69.95 plus $2.50 for postage and handling (Illinois residents add 5% sales tax) to the address shown below, or credit card buyers may call our toll-free number below. By return mail, you'll receive your Astropulse 10 complete with a 90-day limited warranty, carrying case and blood pressure record book.

TEST IT YOURSELF

When you receive your monitor, see how easy it is to slip the cuff on your arm, tighten and inflate. See how easy it is to read. If for any reason you are not absolutely pleased with your unit, return it within 30 days for a full refund, including your $2.50 postage and handling. There is no risk.

Space-age technology has made it easy to know your own blood pressure. Order an Astropulse 10 at no obligation today.

JS&A® NATIONAL SALES GROUP

Dept. SA One JS&A Plaza
Northbrook, Ill. 60062 (312) 564-9000

CALL TOLL-FREE. . . . 800 323-6400
In Illinois call . . . . . (312) 498-6900

text, fine for reading and showing to the preschooler, a mixture of empathy in caring and the diversity of life.

The young author, with several similar books to his credit, is plainly a man of parts. He dedicates his book to the T'ang poet Tu Fu, and he begins it all with a noted 20-character poem, quite untranslated because it is only of personal relevance.

Folktales

NOAH'S ARK, illustrated by Peter Spier. Doubleday & Company, Inc. ($6.95). "Wide and stark, / Was the ark.... / Creatures all, / Large and small,... / Fierce and tame, / In they came, / Pair by pair, / Gross and fair." So for 30 couplets runs the beat of the 17th-century Dutch poem given us in strong rhyme by Peter Spier. Then, with not another word, this artist shows us how the voyage was, the old myth made luminous by a powerful imagination and a brush marvelously devoted to the truth. The meticulous and colorful ink and wash drawings lead us the entire way to the new world. Would you build an ark? Saws, ladders, heavy mauls and big hay rakes are needed. Jars, baskets, sacks and bales must come aboard. Spearpoints turn toward the great herds left sadly behind as the chosen pairs walk or hitchhike up the ramp. Only the bee queen and one drone can come, as Noah shoos off the swarm of workers. How the opossum clings to the elephant's tail! Even two dodos make it.

The flood covers the trees, then the city's towers and finally the mountains as we draw back for a wider view. The ark is a crowded barn indeed, with pythons curled around the roof posts, lions and tigers prudently caged and spiders and scorpions in neat terrariums. One mouse tail is lightly pinned under an elephant's toenail, eggs are retrieved from the orderly rows of laying birds and Noah with his hayfork looks in wonder and dismay at the big komodo lizard he is feeding. Then sea, storm and rain, rain, rain, which the ship sheds in torrents. Whales and dolphins and seabirds keep the ark company. The house cats live peacefully with the people, and once the candle is snuffed the staring eyes of a dozen species of owls mirror the moon.

The ark grounds roughly, but Noah is ready for it. At sunset the dove returns with the green bough. Through all the dark ship Noah runs bearing witness. Most of the beasts are elated, although the leopards seem bored. The cow eats the bough, its first fresh fodder in a period long enough for the rabbits to have increased to a dozen or more. Down goes the ramp (naturally just built for the occasion; one would not have known at the beginning how long to make it) and the animals reach dry land, rhinoceroses, raccoons and every other beast. The ark is empty, a shambles of a zoo ship, a world mess, full of husks and stems and junk, when finally the snails make their dignified way out across the threshold. On the end papers the rainbow glows as Noah plants a vineyard.

RUSSIAN FAIRY TALES, translated by Norbert Guterman from the collection of Aleksandr Afanas'ev. Illustrations by Alexander Alexeieff. Commentary by Roman Jakobson. Pantheon Books ($5.95). Here in a fluent translation are ominous little huts with chicken legs, Swan Maids and glowing Firebirds, shrewd peasants' sons, ingenious and spirited princesses. Some 200 folk stories of all types are here, as they were first set down during the decade of our Civil War, the epoch of the flowering of the Russian prose masters. The strong current of Russian oral literature runs fresh and strange: we encounter regiments, cities, debts, czars (in plenty) and a good deal else not so familiar in the well-known tales of the languages of western Europe.

The rise of the secular use of written vernacular was much delayed in Russia. Old Church Slavonic embodies a rich literary tradition, but it is thoroughly ecclesiastical; for a long time the medium for the laity, mixing saints and sinners, was oral. Not until the 17th century did written secular fiction begin in the Russian tongue, centuries later than it had in France and England. That we learn from the scholarly comment in this admirable book, with its textural, mannered, strong drawings quite similar in feeling to the text. The readers or the read-to will not care. They will be lost to the singing bun, the pike with golden wings and the clever peasant who has to divide five roast geese among six members of a noble family. (It is easy. One goes to the baron and his wife; with it they make three. One goes to the two sons; three again. One goes to the two daughters; another symmetry. "The last pair of geese he took for himself, saying: 'Now I and the geese are another three.' ") Reprinted after 30 years, this book is a gem of its kind.

Technology

THE OBELISKS OF EGYPT: SKYSCRAPERS OF THE PAST, by Labib Habachi. Edited by Charles C. Van Siclen III. Charles Scribner's Sons ($12.95). Those antique granite needles grace the busy cities of London, Paris, New York, Istanbul and particularly Rome. Thirteen obelisks are now standing in the Eternal City, compared with only eight in the rest of the world. The big monoliths are, however, imports to all those capitals. They were conceived, made and first erected in ancient Egypt: at Heliopolis, now a district at the northern edge of Cairo; in Karnak; at ancient Thebes, now Luxor, and at one other site in the Nile delta—and not elsewhere. They were sacred to the sun god, emblems of power and sovereignty in his name for 20 centuries or more.

The granite quarries of Egypt are in only one place: Aswan. There the obelisks were chopped out of the living rock with hard hammerstones, then dressed and trimmed, hauled over ramps to the Nile and sent by barge on the long voyage north to the sites they were to grace. One was erected in Heliopolis, for instance, by Ramses II in about 1300 B.C. Augustus Caesar caused it to be taken to Rome, the first of all the trophy obelisks; the very ship that carried this extraordinary load was on exhibit in imperial Rome until it was destroyed by fire. That stone was the sign of the conquest of Egypt by the legions. It stood while Rome stood, to fall and lie forgotten while Rome was dark. Now it stands restored in the Piazza del Popolo, where the popes of the 16th century placed it in their time of power. The urge for obelisks remained strong, and the ships that carried these massive souvenirs to Paris and New York became similar cynosures in the age of steam.

The author is a veteran Egyptian archaeologist and former chief inspector of antiquities there. He offers a clear story of what obelisks meant, how they were made and where they went, as far as we now know it. No other nontechnical book tells as complete a tale as this small one; indeed, most sources celebrate more the act of removal than the work of first erection. Here we follow the entire slow journey, neglecting neither the stormy Atlantic voyages nor the half-formed mass of the great unfinished obelisk still in the Aswan quarry, abandoned with a heavy heart by the engineers of the New Kingdom once their work disclosed fatal flaws in the rock. It weighs nearly 1,200 tons; it is not easy to name any heavier load moved overland in one piece since. Perhaps the Apollo rockets are the best modern candidates.

MEN AT WORK: PHOTOGRAPHIC STUDIES OF MODERN MEN AND MACHINES, by Lewis W. Hine. Dover Publications, Inc. ($3). HUTS, HOVELS & HOUSES, by Timothy Fisher. Drawings by Kathleen Kolb. Addison-Wesley ($5.95). For more than 40 years the Empire State Building remained the highest in the world. That property distinguished it at least symbolically among the skyscrapers of Manhattan. Lewis Hine was a noted photographer whose studies of Ellis Island and of child labor had played a part in social change. As the Empire State Building grew from bedrock to mooring mast he set about a task of documentation, now not in critique but in celebration. Here is a reprint of the book Hine issued in 1932, with 18 additional photographs added for this edition. We share his pictorial admira-

The new Chevy Monza.

Value, variety, economy and kicks.

Save some gas.
Have some fun.
Conflict of interests? For some cars, perhaps. But not Monza.

Because Monza, while small and efficient, is also very big on style, sport, and spirit.

This year there's a new look to Monza, front and rear. A spunky cast-iron 4-cyl. engine is now standard. And there are more Monzas than ever to choose from: sleek Hatchbacks, stylish Coupes, the sporty Spyder package. All manner of inviting variations.

As for economy, EPA estimates are 34 mpg highway, 24 mpg city, 28 combined. That's with the standard 151 Cu. In. 4-cylinder engine and manual transmission. Power train not available in California. Your mileage may vary depending on how and where you drive, your car's condition and available equipment. (Monza is equipped with GM-built engines produced by various divisions. See your dealer for details.)

Save some gas, have some fun. Make a Monza move today.

Chevrolet

SEE WHAT'S NEW TODAY IN A CHEVROLET.

tion and praise of the workingmen who nimbly walk the high girders, who turn the wrenches, guide the bucking pneumatic drills, ride the crane hooks and deftly guide the heavy swinging loads that finally articulate to shelter a village of office workers piled high into the air. It is the posture and the faces of the individual men that mark these photographs (with captions) as being the sensitive portraits they are; any child will find them fascinating, the more so if there is a chance to compare these images with similar work being done today. The most evident visual difference is obvious: on what is now the archetype of hard-hat jobs not one hard hat is to be seen. These men worked bareheaded or in a peaked cap, a felt hat or a beret. Today the cranes are bigger and the tempo is faster, no doubt, with new materials and new tools, but the poised, purposeful and taut craftsman has not changed. The turbines are bigger, the lathes and gauges more complex, but that intent look is an invariant of responsibility and skill.

In *Huts, Hovels & Houses,* the second of these books, we see another set of construction workers on the job. They are kids in the upper grades of school. They are building too, with beer cans and rolled newspapers, with builder's polyethylene sheets inflated by an attic fan, with palm leaves, snow blocks, bales of hay and milk cartons. For each scheme there are a couple of careful and knowing pages of hints and a handful of drawings, both details and overall views. (Those inflatables are too idealized here, but the rest might be drawn from the life.) The steps are given for a basic wood structure—roughly a five-foot cube—together with modifications that could turn it into an indoor electronics shack, a lemonade stand, a tree house, a floating raft house or a trailer on grocery-cart wheels. Sod, concrete, a simple water supply and solar heating are discussed very briefly. The explanations are simple and sensible and afford adequate, if terse, technical guidance. Still, the commonsense skills, and above all the capacity for sustained cooperative hard work and for flexible planning, are assumed. Rome is not to be built in a day, even with inflatables.

CASTLE, by David Macaulay. Houghton Mifflin Company ($8.95). Big, clear drawings in ink sprawl over an entire page or even two pages. You first see the limestone outcrop at the mouth of the river Wyvern in northwestern Wales. A few little one-masted vessels lie in the lee of the hill, empty fields stretch around it. Now, however, Master James of Babbington and his carpenters have arrived, with their strong infantry guard. The Castle of Aberwyvern is to be built for Kevin le Strange, created lord by Edward.I of England in the spring of the year 1283. There is a closer

view, still from high above, and we watch the moat being cut. By the end of the year we see half-finished walls of fitted and mortared stones filled with rubble. By the end of the spring the castle walls and the outlying town wall are both much grown, and in October of 1288 the structure is finished and whitewashed, "giving it the appearance of having been carved from a single enormous piece of stone." We do not see only a distant view. Quite the contrary. We read of arrow loops (the familiar slits) and garderobes (the medieval toilets), we see the tools of the workmen and the arms of the garrison, sections and details, fireplaces and gatehouses, dungeon and culprit. At a simple level, clearly and carefully, we follow the design and construction of a castle and fortified town of the period. It is no single real place; the author-artist has put together a prototype of such a structure, part of the English conquest of Wales, embodying two centuries of military engineering "throughout Europe and the Holy Land." You can see what battlements are and why the walls are formed like the top of the chesspiece.

Prince Daffyd lays siege to the castle in the spring of 1295. Catapults and battering rams, siege towers and a sapper's tunnel assail the stronghold. The design and the loyal English defense resist the brave Welsh attack. For two centuries the town grows, and Welsh and English families inhabit the place in peace side by side. Wales is no longer held by armed force but by social bonds. The castle becomes a romantic ruin there on the hill, a mere quarry for building stone; the town walls become more of a nuisance than a necessity. We have seen and read the brief biography of a castle, brilliantly evoked in text and drawings.

MAKING VICTORIAN KINETIC TOYS, by Philip and Caroline Freeman Sayer. Taplinger Publishing Company ($6.95). In 1825 P. M. Roget first described the phenomenon of the persistence of vision in the *Philosophical Transactions of the Royal Society*. In those days the magic lantern was a home standby. The phenomenon and the device together were parents of the motion pictures that came at the end of the century, but before that they dallied in a dozen ways, their issue including riffle pictures, moving slides, spinning images and wonderful peep shows. In this book, presented in authentic period style, many of these gadgets are pictured and explained, and explicit instructions are given for reproducing the simpler ones. It is not at all clear that the television screen is proof against the lure of the fine old peep show. Who would not view the battle of Hastings there, or Paris in panorama, with changing and romantic light? "It is difficult to give instructions for making a peepshow, as success depends less on technical competence than [on] artistic abilities and imagination." All networks please copy. The book is sure to be a delight for the handy and experimental.

GAMES & PUZZLES YOU CAN MAKE YOURSELF, written and illustrated by Harvey Weiss. Thomas Y. Crowell Company ($6.50). In simple, clear drawings, with directions inserted right among the drawings, making the order and nature of the work easy to follow, this experienced author offers workable recipes for a wide variety of games and puzzles. They range from simple board games, made attractive with clay or plywood boards, to games of action: pop-it-in games, sliding games based on shuffleboard, indoor versions of miniature golf, swingball bowling (the ball is tethered to the top of a doorframe), spinball (tops and table-tennis balls) and a complicated but rewarding game called soccer. Soccer uses clothespins mounted on dowel rods in such a way that they can kick a marble along the board. Several such rods fitted into the walls of the game box make up each team. The box looks good but is rather hard to build; the work is accessible and stimulating to kids in the early grades who have learned how to use a knife, glue, a file, sandpaper and maybe a drill. Some puzzles are included, such as the tough one that requires putting a stick with a loop of string on it through a buttonhole.

The Better Use of Air

HOW TO THROW A CURVEBALL, written and demonstrated by Don Sutton. Created by Jim Auker. Follett Publishing Company ($2.50). MANY HAPPY RETURNS: THE ART AND SPORT OF BOOMERANGING, by Benjamin Ruhe. Illustrated by Peter Ruhf. The Viking Press ($4.95). Sport and antisport, as these books undertake to show. The first book is a thick, hand-sized booklet of some 80 pages. The pages bear photographs that were shot in quick sequence and are arranged on the page to be viewed staccato as well. Deft riffling of the pages (the series trade name is Rifflix) yields a moving picture, as in days of old. Free access to repeated motion is here applied to instruction in the subtle art of pitching a curveball. The master is Don Sutton, a right-hander from Clio, Ala., with a record of more than 2,000 strikeouts and the ambition "to hold all major Dodger pitching records before he retires." No theory at all enters, but a clear look at Sutton's grip, a motion picture of the windup and the delivery from two directions, and a closeup sequence of arm and foot action fill up both directions of riffle. There is a lot to watch, and a sense of intimacy. A text goes with the motion pictures. The book ought to help; it does show the ball curving across the plate to the catcher's mitt, and it ends with the admonition: "Work hard, run a lot, and by all means have a

good time." This is a neat blend of earnest pedagogy, good cinema and implied aerodynamics. The spin makes the curve.

It is plainly sport, professional and proud! *Many Happy Returns,* the second book, is another curveball book, an easy, cheerful, helpful account by a Washington expert on and organizer of boomerangery. Unlike the boomerang manual last noticed in these pages, this book celebrates not records but the "abandoned and spontaneous" attributes of the complex, unpredictable devices. The author sees them as antisport, symbols for the person who wants to be an individual "in a complex, regimented, rule-filled world." The laws of aerodynamics are not slighted; the designs are clear, the history is intelligent and lightheartedly skeptical. You can read of a whiz of the modern stage, or of the aboriginal experts who used to spin one off at night bearing a live coal, a stone-age piece of fireworks as ingenious as the long list of other feats they performed. The throwing-stick was the major hunting weapon; boomerangs were an entertainment and a challenge, with fringe benefits in hunting birds. And the file-card version for which instructions are given is just right for when the teacher has left the room! This is a book that is fun to read and pleasant to act on, eccentric, informed, even a little wise.

Earth and Sky

LOOKING AT MAPS, written and illustrated by Erich Fuchs. Abelard-Schuman ($6.95). EARTHWATCH: SPACE-TIME INVESTIGATIONS WITH A GLOBE, by Julius Schwartz. Illustrated by Radu Vero. McGraw-Hill Book Company ($6.84). ORIENTEERING FOR SPORT AND PLEASURE, by Hans Bengtsson and George Atkinson. The Stephen Greene Press, Brattleboro, Vt. ($5.95). Three steps to the Parnassus of maps are here themselves well mapped. *Looking at Maps* is for readers in the early grades, the kind of youngsters who might like to come to write, as the protagonist does, Ben's Room, Garden Street, Peacedom—out to the Milky Way Galaxy and the Universe. The cheerful, whimsical drawings of this Stuttgart author-artist show Ben's room, and all the rest, up to a fine painting of the Milky Way, with the sun's place pretty plain. Then we follow mock maps scale by scale right down to Ben's room. A few pages more show real and stand-in maps, their symbols and their nature, including road maps, a moon map, aircraft and ship charts, all as drawn by Fuchs (although some are of real places). There is a real, if playful, world map with physical signs, and there are a few maps of North America made to show the produce of field and factory. An artist's light touch makes palatable a very straightforward pedagogical approach to set the young eye firmly on a lifelong pleasure: looking at maps.

Earthwatch is just that, a close look at the earth as a whole. Well, not the real earth, but the next best thing, which an interested schoolchild, rather beyond the third grade, or a class or club of schoolchildren, can easily set up. It requires only thought, care, a place in the sun and an inexpensive free globe of the earth (which could well be a large hand-painted ball). This wonderful 18th-century activity—the "rectified globe"—is made newly rich by a fairly complete description of its wonders as the sunlight falls on a well-placed model earth. The terminator, the time zones, the date, the season, temperature and insolation differences, the moon and its orbit, even eclipses are here, and more. It offers the best of sundials and a clear, varied introduction to the main phenomena of relations between the earth and the sun. Classrooms everywhere need it, and many a backyard. The rectified globe is more suited than the silver globe to many a well-appointed garden, and the book should be a beacon to public centers concerned with the education of children and the public. Exactly why you are sitting "on top of the world" wherever you live might have been made even clearer. This study supplies a retort to those folk who have been put down by being "down under." You can safely bet that Sydney too is the top of this round world (along with Lagos and Tokyo and Coffeyville, Kan.).

Orienteering is two things. First, it is the art of navigating one's way through new territory with a map and a compass; nowadays it is also a competitive outdoor sport based on that skill. In Sweden the sport rivals soccer for popularity among participants. The ideal orienteering ground is a square mile or less of complex wooded terrain with a few old roads and fences. The great national meet in Sweden has drawn more than 10,000 participants, and in 1976 in Quebec we transatlantic tyros mustered some 1,500. Canada is ahead of the U.S., but there are now "O clubs" in nearly half of the states. The sport's origins were military, since it is obviously close to the task of the infantry patrol; the military still take an interest in the sport. Now it is seen as a public-health measure, good mental and physical exercise for the young and the old. The concept is simple but the details and variants are as rich as terrain itself. One follows a mapped and prepared course from one fixed control point to another, traversing the chosen route (straight over the hill, or around by the path?) against the clock.

Meets are arranged with graded levels of length and difficulty, from only a couple of kilometers in easy terrain to a dozen kilometers of cliff, marsh and brush. Times run from half an hour up to an hour and a half. The map is the key; the compass and its skillful use are next in importance. The technique is spelled out carefully here from the raw beginnings (and the choice of shoes) up to the organization of clubs and a championship meet. There are many clubs to help, and of course national federations; the addresses and suppliers are all in this compact but dense guide, which also offers a sample map. The authors were the organizers of the New England club. Automobile and motorcycle map "rallies" are plainly kindred, although paving and the internal-combustion engine have few friends in real O circles. On the other hand, there are muscle-powered variants too. Consider the snow-covered case: ski orienteering is well established (cross-country style, not Alpine). Sailing? City streets? Snowshoes?

Map making is the highest level of map using. Orienteering offers a challenge to any who would improve the standard topographic maps with the richer detail of trail, stream, knoll and clearing that helps the hard-pressed runner to quickly find the flag, its verifying punch tied securely down, behind the big boulder. These races are not invariably to the swift; there is always chance, insight and the temptation of ripe berries.

THE VERSATILE SATELLITE, by Richard W. Porter. Oxford University Press ($11). Inside an unsealed satellite it is hard to transfer heat. The vacuum cannot much cool the hot spots of the electronics, and so the designer makes quite sure his transformer can drop its heat over a good conducting path to an emitting surface—a radiator—outside. Before long, as the art grows, this single problem becomes so complicated that specialists in heat flow, armed with powerful mathematical models suited for computing, are called on to study any small change in the spacecraft. This knowing but entirely descriptive level is struck throughout this excellent overall survey of satellites and their applications (with military reconnaissance strangely omitted) by an experienced and responsible engineer in the industry.

The history is quickly sketched, the physics of rockets and orbits is well explained at a nonmathematical level (a demon flings balls back out from the rocket to introduce us to ideas of thrust and jet velocity), and the applications are surveyed chapter by chapter: communications, weather watch, navigation, land and sea survey, biological monitoring, space research and gravitational studies, man in space and a modest projection into the future. All of this is well assisted by visual material, from satellite photographs in many color bands (a couple are in full color) to photographs of spacecraft and many maps and graphs. The range is large and the writing clear, accessible to junior-high children and beyond that to adults who

THE PLATTER IS THE MOTOR.

Fisher introduces a major advance in audio technology: the linear motor 120 pole, direct drive turntable.

One of the problems in improving direct drive turntable performance is reducing wow & flutter due to cogging action of the motor.

With its limited number of poles (usually 12) and its relatively slow operating speed, most conventional direct drive systems also have an inherent problem of low starting torque.

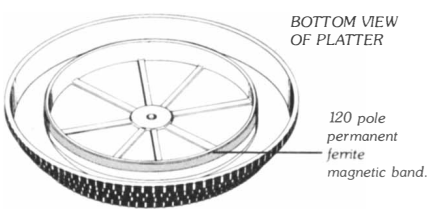
To solve these problems, Fisher has engineered the linear motor, direct drive system. The new Fisher MT6225.

In effect, the platter becomes the turntable's motor. And Fisher's 120 pole design practically eliminates cogging action, and lowers wow & flutter to a totally inaudible 0.03%.

The linear motor direct drive system further reduces turntable rumble to an extremely low −70 dB, far below hearing level.

HOW IT WORKS

A continuous band of ferrite material, containing 120 magnetic poles, is attached to the inside bottom rim of the platter. To start platter rotation, each pole opposes one of the electromagnetic drivers on the base top. Each of the 120 poles is attracted or repelled as it passes the driving electromagnets for smooth, stable operation.



An electronic sensing device monitors the platter's speed, and acts as a servo-feedback control to maintain practically perfect speed accuracy. Built-in strobe and pitch controls are provided.

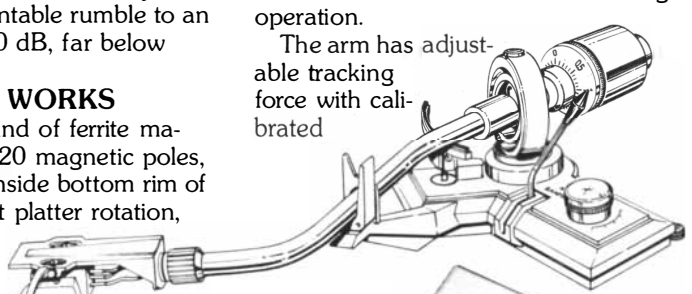
The speed accuracy of the system is independent of line voltage fluctuations.

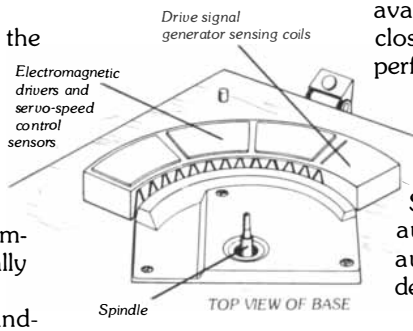
THE ARM

The Fisher MT6225 is equipped with a professional-type gimbaled tone arm for performance to equal its technically advanced motor design.

The arm accepts all standard cartridges. The auto-return system, which returns the arm to rest at end of a record, is designed with no restraints on its lateral or vertical motion during operation.

The arm has adjustable tracking force with calibrated settings from 0.6 to 3.5 grams. There's also precise variable anti-skate, and viscous damped cueing. The MT6225 also has a heavy cast-aluminum platter, and a massive integrated base to absorb vibration.





In all, there is no other turntable available that comes closer to the perfect performance . . . or the ultimate in reliability. Fisher MT6225, for about $250.† See it now at fine audio stores and the audio section of department stores.

Motor	Linear Direct Drive
Wow & Flutter	0.03% WRMS
Tracking Force Range	0.6 – 3.5 grams
Rumble	−70 dB (DIN B)
Maximum Tracking Error	± 1.5°
Auto Stop	Yes
Auto Reject	Yes
Cueing	Viscous Damped
Anti-Skate Control	Adjustable
Platter Weight	2.2 lbs.
Speed Selector	33/45 rpm

FISHER

The first name in high fidelity

Fisher Corporation,
21314 Lassen St.,
Chatsworth, Calif. 91311

†Manufacturer's suggested retail price. Actual selling price determined by individual dealer.

want a compact survey of this remarkable 20-year-old industry, already quite mature.

MESOZOIC AND CENOZOIC PALEOCONTINENTAL MAPS, by A. G. Smith and J. C. Briden. Cambridge University Press ($3.95). Behind the jargon of this formidable title lies an atlas of dreams, which came clear to some children before the geologists awakened. Here are 52 carefully plotted page-sized outline maps of the world, in four different projections. Each map presents the continents as they were at some point in the past, the snapshots having been taken every 10 or 20 million years back to the early Triassic. If you would watch India sail north to smash into Asia, crinkling up the Himalayas, why, here is the place to do it. With a little practice you can flip the pages and see it as a motion picture.

The maps are computer-generated by rotation around the varying north paleogeomagnetic poles, the amount of rotation controlled by sea-floor magnetic measurements. Expanding motions, and not the concealing compressive ones, must be relied on. Only the longitude reference of the entire land mass remains arbitrary. Anyone who likes maps and the world would find these maps a resource, particularly if there is some incentive to compare the climate, the landscape or the fauna and flora through geologic time in this shifting theater of sea and land.

Numbers and Shapes

I CAN COUNT THE PETALS OF A FLOWER, by John and Stacey Wahl. National Council of Teachers of Mathematics, Reston, Va. ($3.50). NUMBER MACHINES: AN INTRODUCTION TO THE ABACUS, SLIDE RULE, AND POCKET CALCULATOR, by Forrest M. Mims III. David McKay Company, Inc. ($5.95). A calla lily has a single large white petal, but in the golden "flower of the Incas" 10 red-tawny petals circle the thread-gold center. There is a flower for every integer in between; another count takes us along a different garden path to 16. This strikingly beautiful book offers a petal count to reward the young eye with color close-ups of exquisite flower after exquisite flower. The second time around five different photographs grace the page numbered 12: a dozen-petaled black-eyed Susan, four threefold purple spiderworts, three fourfold dogwoods. For 13 a prime cineraria stands alone. Symmetry, factors, primes and composites, even and odd—all enter gently the eye of the mind through the flower heads. The artist-teachers who prepared this treat, as tasteful in depth as at the surface, include a final page of wonder: some of the flowers, clearly familiar from earlier pages, repeat, but now as botanical oddities with an extra petal or two. The world of life is not to be caught uniformly in the static net of the natural numbers; we learn that even a count can open onto novelty and surprise.

Number Machines, the second book, is a neat and simple story of tangible counting, not for the delight of the flower but for practical reckoning. It makes plain how the Chinese abacus and the slightly trickier Japanese one work, with photographs and sketches. The slide rule too is presented, not mainly by cookbook rules for manipulating the lettered scales—although they are included—but by an explanation of how to make your own prototype slide rule on paper both for addition and, with the help of logarithms (very briefly explained), for the usual multiplication and its inverse. One might be a little disappointed by the absence of any account of what fractions can mean as exponents, but the logs of integers from 1 to 10 are given to two places, without explanation. The pocket calculator is explained at a block-diagram level, rather sketchily; some nice tricks are given for those who own one. A couple of pages describe the prowess of programmable pocket calculators but not their nature or use. This is a useful book for the right middle-grade readers.

CURIOSITIES OF THE CUBE, by Ernest R. Ranucci and Wilma E. Rollins. Illustrated by Henry Roth. Thomas Y. Crowell Company ($6.95). In the natural world spheres are many but cubes are few. This well-illustrated little book of recreational mathematics is built on cubes. They are in the main abstract cubes, although we do learn how to fold cubes out of flat paper and how to build them out of drinking straws and string or out of a variety of paper pyramids. We learn about their various axes of symmetry, how to dissect cubes into cubes, how to cut them to reveal polygons of all kinds (although never a regular pentagon).

Cubes are also projected in many ways, with their edges described in all the relations of lines intersecting or skew, and a cube is passed through a square channel bored through another cube. The one that passes down the channel can have an edge 6 percent larger than its host, although that takes a well-designed channel. Even the hypercube, or tesseract, is examined. There are cube games but no dice. A good student from the upper grades on who is interested in geometry will enjoy this book and is apt to fulfill the hope of the authors, experienced teachers of mathematics, that readers "will never look at a cube again and see only a cube."

ARRL CODE KIT, by Charles J. Harris. (A brochure and two tape cassettes.) The American Radio Relay League, Newington, Conn. ($8). Citizen's-band radio communications flared into a craze, placing no demands on its millions of practitioners beyond $100 or so and maybe a car; you could become a good buddy on crowded voice channels to the rolling world. Amateur radio, with its technical open-endedness, its worldwide flavor, its long traditions and its inherent diversity, is quite another thing. There is above all a license barrier, the most difficult part of which is surely the skill at sending and receiving Morse code (actually International Morse, or Continental). This pair of tape cassettes represents the newest aid to learning that still indispensable skill, which is legally required and still functional for narrow-band low-cost communications, even though it is fading in this computers' world.

The tapes offer a meticulous set of graded dit-dah practice sessions, half an hour each at 5, 7.5, 10 and 13 words per minute. They provide for a gradual increase in speed matched to the levels of the several U.S. license examinations (although not to the most advanced). Encoding random letters makes the tapes hard to memorize and therefore offers enough practice for proficiency. The lower speeds are sent at a high character rate but with long periods in between. Any standard cassette player can be used. Twenty minutes a day for as long as it takes is what is recommended. There is a detailed account of how to go about it, and what you can expect, in the brochure. Code is not the only requirement for the would-be ham, but perhaps it is the toughest one. Good luck; you are on your way.

STAINED GLASS TESSELLATION POSTERS, designed by Sheryl Cotleur. Creative Publications Inc., Box 10328, Palo Alto, Calif. ($6.95). The foot-square heavy-paper posters of this set glow with the softly modulated colors of stained glass. They present photographs of leaded stained-glass windows in geometric designs made particularly for this form of dissemination, aimed at elementary students of geometry. (If one aims at proofs, more mature students will find plenty to interest them.) There are 11 designs, each design a piece of an infinite plane tiled with a plane-filling pattern of polygons, the same at any point. There can be no more; they of course include square tiling, such as the checkerboard, the hexagons and equilateral triangles. Eight more are semiregular, so named; they mix a variety of regular but unequal polygons, including octagons and dodecagons. What the polygons are, the notation, the generation of the entire set, the meaning of angles, the sum of the angles in any polygon and other questions touching on the grand problems of space and its symmetries form the challenge of the handsome posters. A three-page brochure with references and suggestions is an adequate first step to what can be drawn from this agreeable visualization.

Courvoisier

V.O.C. "The Brandy of Napoleon," beautifully held in handsome Baccarat Crystal.

V.O.C. decanter not available in all stores. For shopping assistance, call 800-327-5702, toll free.

THE 1978 SCIROCCO.

Feast your eyes, while you can, on the 1978 VW Scirocco.

We got one to stand still long enough to get its picture taken.

But standing still hasn't been the Scirocco's favorite thing.

Its favorite thing has been winning.

We designed a 170 horsepower, 150-mile-an-hour souped-up racing

version of Scirocco that blew the doors off Jaguars, Corvettes, etc. at Nelson Ledges and went on to win the '76 Trans Am Championship for Touring Cars under 2 liters.

For you, we build the '78 Scirocco with a more modest 1.5-liter fuel-injected overhead cam engine. With stripes and spoiler, a la the Limited Edition above. Or in dazzling street clothes, a la the best-looking Volkswagen that you've ever seen.

Either way, Scirocco goes cheerfully from 0 to 50 mph in 8.2 saucer-eyed seconds. It also holds 4 breathless people and 16 cubic feet of trophies.

Front-wheel drive. Rack-and-pinion steering. Mind-warping handling. Giugiaro styling.

And you're just sitting there?

VOLKSWAGEN DOES IT AGAIN

Jameson Irish

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

If he likes Scotch, he'll love light Jameson Irish.

Let him try a glass of Jameson Irish Whiskey the way he would his favorite Scotch.

He'll notice how much it tastes like fine Scotch—only lighter and more delicate. Not smoky tasting like Scotch.

Chances are, as a dedicated Scotch drinker, he'll instantly appreciate this flavor difference.

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

Jameson. World's largest-selling Irish Whiskey.

80 & 86 PROOF • CALVERT DIST. CO., N.Y.C.

Established 1845 SCIENTIFIC AMERICAN December 1977 Volume 237 Number 6

Product Technology and the Consumer

A buyer who wants to compare products in terms of safety and energy consumption is now largely in the dark. Work at the National Bureau of Standards will put more data on some products at the point of sale

by G. Franklin Montgomery

In 1975 the people of the U.S. bought nearly 127 million major household appliances, which had a retail value of more than $14.5 billion. Few of the buyers had much information enabling them to compare one brand with another wisely, particularly if they were concerned about the durability of the appliance and its lifetime cost (a sum of what the buyer would spend to buy it, to maintain it and to pay for the energy it would require). At the end of 1977 the situation is much the same.

Help is on the way. Partly as a result of the "consumer movement" of recent years and partly because of rising concern over the efficient use of energy, Congress has passed legislation calling on various Federal agencies to establish certain standards for the performance of major appliances. One of the agencies involved is the National Bureau of Standards, which is concerned primarily with such measurements as length, mass and time and with the properties of materials but has also been active in consumer problems periodically. Work in progress at the bureau's Center for Consumer Product Technology, with which I am associated, promises to make available to the consumer a substantial amount of information on the durability and efficiency of appliances.

What one can expect to see as a result of this work is a labeling system for certain appliances that will set forth, at the point of sale, a measure of the energy the appliance requires in normal use and the efficiency with which the energy is used. Thus the buyer can ascertain by comparing labels that the appliance with the lowest retail price may not be the cheapest in the long run because of its large energy requirements or relative inefficiency. The difference between this labeling system and the information now available to the consumer in product brochures and consumer publications is that the label information will be standardized, based on accepted engineering tests of product performance and available at the point of sale.

The energy required to run a household appliance can be considered in two ways: the primary energy content of the fuel (oil, coal or natural gas) and the energy content of the refined fuel (or electricity) delivered to the household. An energy administrator may make his accounting based on the primary measure; the householder pays his bill on the basis of the end-use measure. For a given appliance the two figures may differ considerably. Many appliances run on electricity, which is generated by burning fuel at a central power station. From the primary-energy viewpoint running an appliance electrically is less efficient than burning fuel directly. From the end-use viewpoint, however, this inefficiency is partly concealed, because electricity is billed on the basis of the kilowatt-hours supplied and not the primary energy required to generate it.

About 20 percent of all the energy used in the U.S. is accounted for by apartments and single-family dwellings. Most of it provides space heating and central air conditioning. About a third of the energy (approximately 6 percent of the national total) operates major home appliances: 31 million room air conditioners, 70 million refrigerators, 23 million freezers and 55 million water heaters.

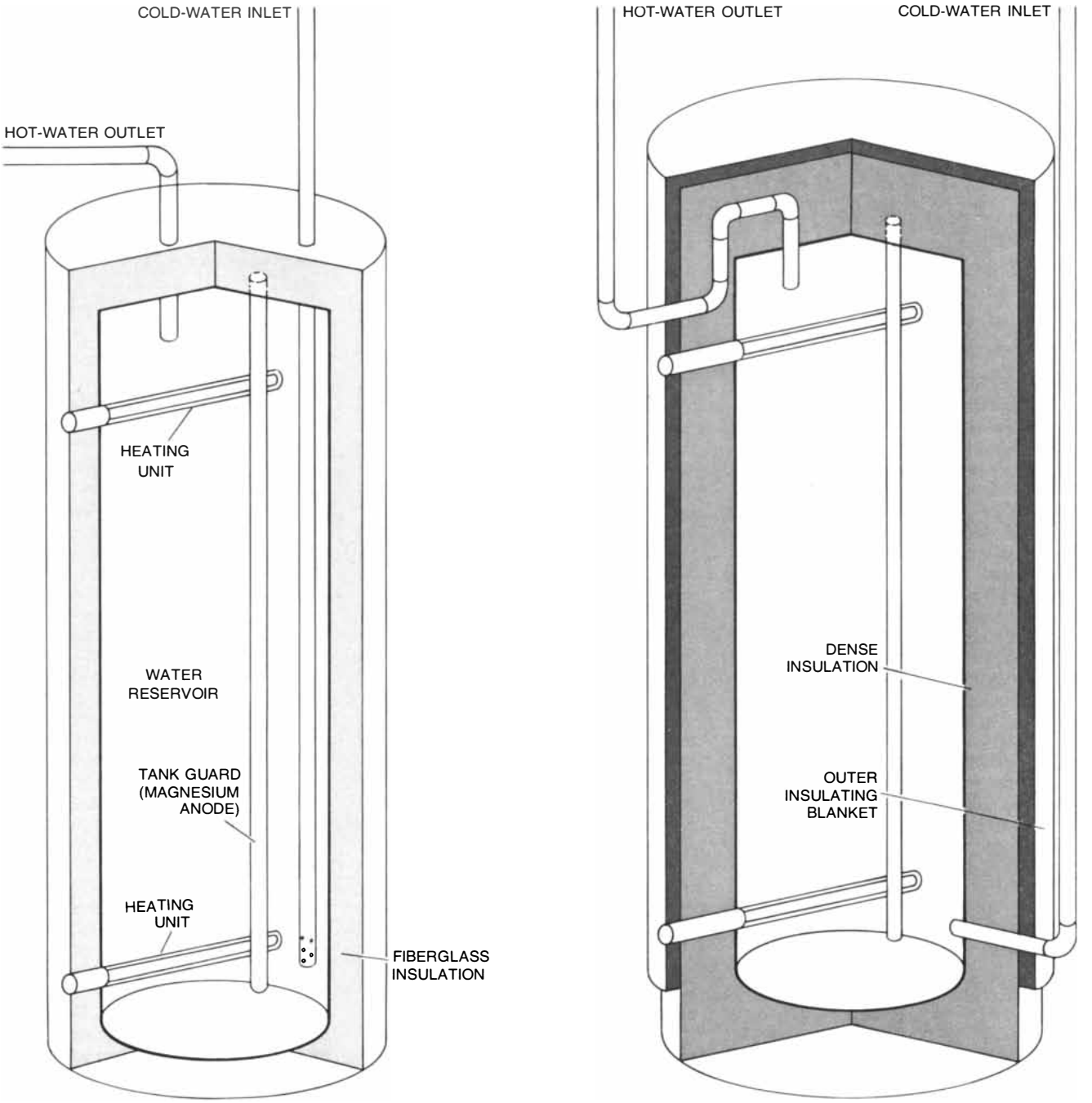
One fruit of the effort to give the buyer more information on the performance he can expect from an appliance is the labeling of room air conditioners in terms of efficiency. The program is the result of a cooperative effort by manufacturers and the Government. An air conditioner included in the program bears at the point of sale a label that specifies the efficiency of the model and the range of efficiencies to be expected of other models with the same cooling capacity. The specification, which is called the energy efficiency ratio (or EER, as it appears most prominently on the label), is the quantity of heat energy (in British thermal units) removed from the air per hour divided by the electric power (in watts) supplied to the machine.

An equivalent measure of performance is the heat energy removed from the air divided by the electric energy needed to remove it. This ratio is commonly termed the coefficient of performance; engineers employ it regularly to characterize heat pumps. For a room air conditioner the coefficient of performance is .293 multiplied by the energy efficiency ratio. A typical range for the EER is from seven to 10 B.t.u.'s per watt-hour; the corresponding coefficient of performance is from two to three joules per joule. (The fact that the coefficient of performance is higher than 1 indicates that the machine removes from the room more heat energy than it

draws from the power source. An air conditioner that is designed and operated properly should do just that.)

The labels that now appear on air conditioners are the result of a voluntary program organized by the Department of Commerce. Under a law recently passed by Congress labeling will be mandatory for a large number of products. The law requires that the annual cost of operation and at least one other measure of energy performance be displayed at the point of sale. (The annual cost of operation is itself an indirect measure of energy performance.)

Partly because of the attention that has been focused on efficiency as a result of the labeling program, manufacturers have modified the design of air conditioners to increase the energy efficiency ratio. The increase exacts a price. Greater efficiency usually requires larger surfaces for heat exchange in the condenser or the evaporator (or both), and that means more metal. Other ways of saving energy, such as installing a more efficient compressor, also increase the manufacturing cost. The result may be a difference of $40 to $50 in the purchase price between an air conditioner de-



ELECTRIC WATER HEATERS are compared. The heater at the left is a conventional model, with about $1\frac{1}{2}$ inches of loose fiberglass insulation between the water tank and the metal outer jacket. The inlet and outlet pipes go straight into and out of the top of the tank. The heater at the right is an improved model designed to conserve energy. Its inner insulation is of a high-density material; the interior wall of the tank is insulated with a ceramic material. The cold-water intake pipe enters the tank near the bottom instead of running through its length, so that the pipe does not draw heat from the tank; the hot-water outlet pipe is bent to achieve a similar effect and a blanket of insulation is added outside the jacket. The loss of heat is only about 25 percent of the amount that is lost from a conventional tank.

signed to be more efficient and one designed to attract buyers with a low price. The consumer must decide (at present on the basis of limited information) whether or not he will recover the higher purchase price of the more efficient machine through reduced outlays for electricity over the period in which he expects to use the machine.

A household appliance that most people tend to take for granted is the water heater, although the energy required to heat water is exceeded only by the energy devoted to space heating. The water heater operates automatically for years, usually receiving little or no attention unless the tank rusts through and begins to leak. By the end of the heater's lifetime the cost of the water and energy it has used is far greater than the amount spent to buy and install it. Can the design be improved so that the heater is more efficient? (I am assuming that if such an improvement resulted in a higher initial cost, as it probably would, the reduction in energy costs would more than compensate for the price increase.)

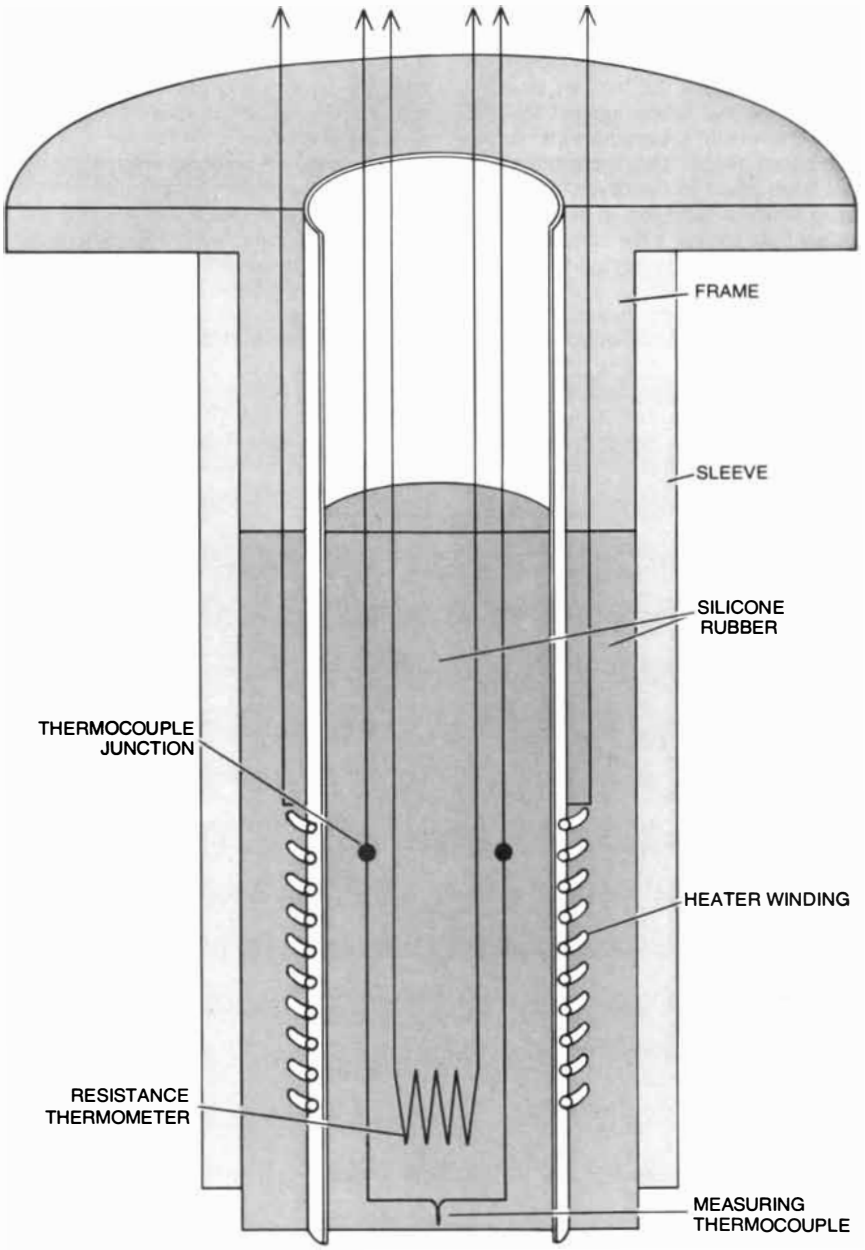
A typical gas-fired water heater consists principally of a cylindrical tank (for the water), a burner and a central flue or stack (to carry off the hot gases produced by the combustion of the fuel). In order to calculate what the cost of hot water will be, the consumer needs to know the efficiency of the heating process, that is, what fraction of the heat energy in a given quantity of fuel is delivered to the water. A typical figure for a modern gas-fired heater is 75 percent; the rest of the energy from the fuel is lost through the flue.

The efficiency could be raised to perhaps 85 percent by changes in design that are now feasible. (They would raise the cost of the heater.) Even with a more efficient device, however, the consumer could not calculate the cost of hot water accurately unless he had a measure of standby loss: the speed with which the hot water in the tank loses energy through the tank jacket to the surrounding air and through the flue wall to the chimney. A typical modern heater loses 6 percent of its stored heat per hour in this way, with an additional .5 percent per hour lost to the surrounding air through the walls of the connecting pipes.

Standby loss could be reduced by more or better insulation inside or outside the jacket and by insulating the pipes. Heat would still be lost through the wall of the flue. This loss could be reduced by installing a flue damper, which would close when the main burner was off. Then, however, there would be no way to accommodate the exhaust gas from the flame of the pilot light. With the incentive to reduce the standby loss, manufacturers are experimenting with substitutes for the gas pilot flame (electric ignition is one idea) and are providing better insulation for the tank. Presumably the increased initial cost resulting from such changes would be recovered over a period of time in reduced fuel consumption.

The household refrigerator is another candidate for improvement. In 1973, the most recent year for which complete figures are available, refrigerators accounted for 20 percent of the electricity supplied to residences. Design changes that appear to be economic when they are considered over the lifetime of the appliance could reduce this energy requirement by about half. They include denser insulation than is now installed, a more efficient motor and compressor, a return to manual defrosting and a switch for turning off the door-frame heater (for evaporating condensation) when this feature is not needed in weather of low humidity.

Drying laundry by machine is also a process that could benefit from more efficient use of energy. The task is partly



"ARTIFICIAL FINGER" was made by the Bureau of Standards to test the surface temperature of objects such as toasters. It is called a thermesthesiometer. The operator holds it in one hand and touches the bottom of it to the hot object. By means of the thermocouple embedded near the tip the device records temperature as a human finger would feel it. In this way experimenters can determine what materials and configurations produce dangerous temperatures.

mechanical and partly thermodynamic. The mechanical part is carried out in the washer when, in the final operation, the machine spins the load of laundry, pressing the wet fabric against the wall of the basket and squeezing water out by centrifugal force. The thermodynamic part takes place in the dryer, where the damp wash is tumbled in a current of hot air that converts the remaining water to vapor, which is carried away in the airstream.

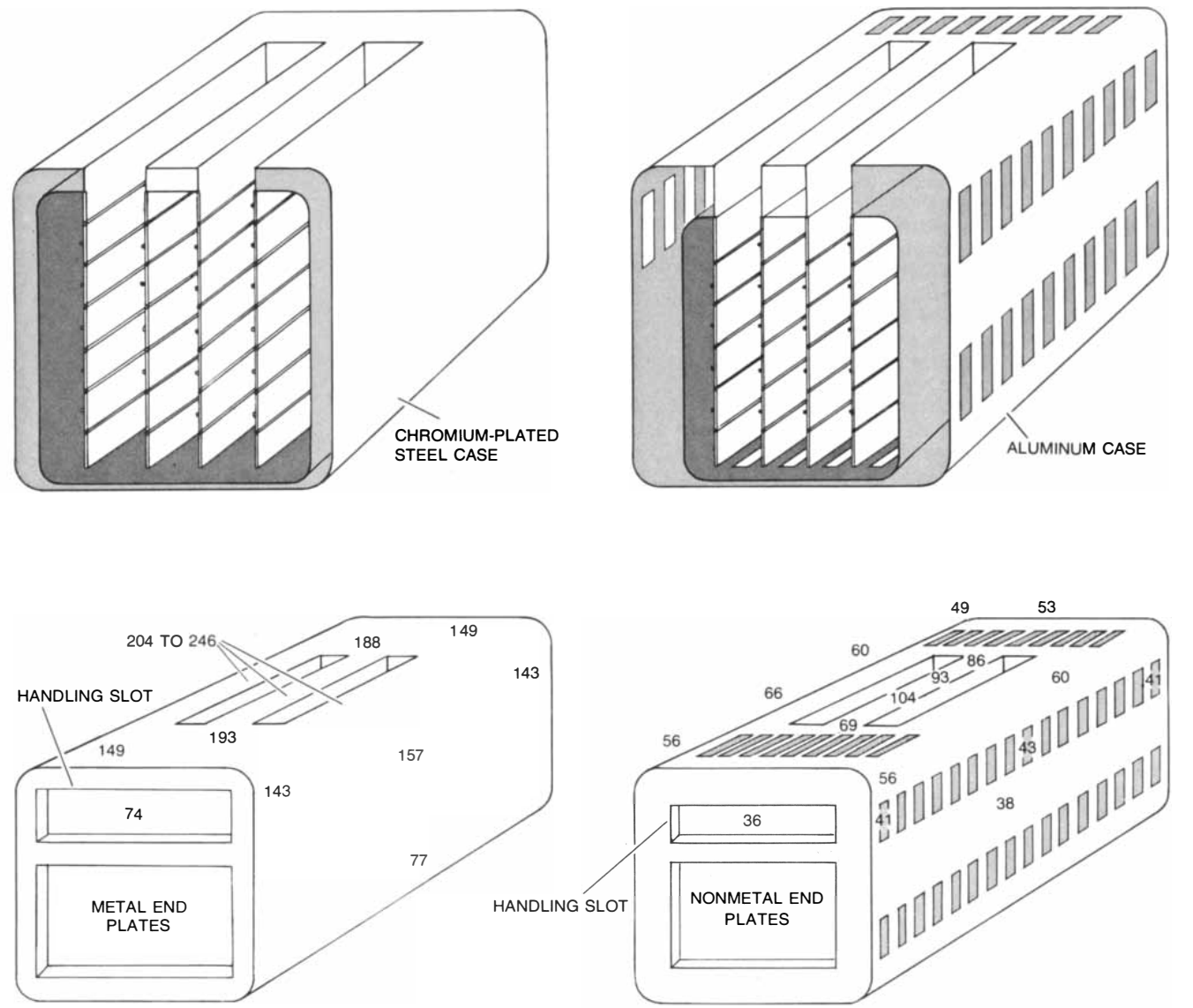
The mechanical process is by far the more efficient of the two in transforming energy. A typical drying yield for a washer during its spin cycle is 30 grams of water per kilojoule of electric energy. In an electric dryer the typical yield is .25 gram per kilojoule.

It is usual to specify the water content of cloth as a percentage of the bone-dry mass of the cloth ("bone dry" being taken to mean a water content of no more than 1 percent). Wet wash, rinsed and drained but before spinning, has a water content of 150 percent; after spinning the content is about 70 percent, and after drying it is between 2 and 3 percent. Many people judge the wash to be dry to the touch when the water content is down to 6 percent. (Evidence suggests that some automatic dryers are adjusted to dry below this level in order to assure the user that the machine is working well.)

With the figures given above one can estimate the cost of drying a three-kilogram wash to the 1 percent level, assuming that the electricity costs one cent per megajoule (3.6 cents per kilowatt-hour). The spinning washer will wring out 2.4 kilograms of water at a cost of .08 cent. The contribution of the dryer will be two kilograms of water removed for eight cents.

Plainly it would be advantageous to have the washer do more of the work. Increasing the yield of the washer, however, would require a higher rotational velocity or a basket of larger diameter. The velocity of the basket is limited by considerations of strength, vibration and perhaps safety; the diameter of the basket is limited by strength and by the width of the doors the washer must pass through to reach the laundry room.

What, then, of the dryer? Here one



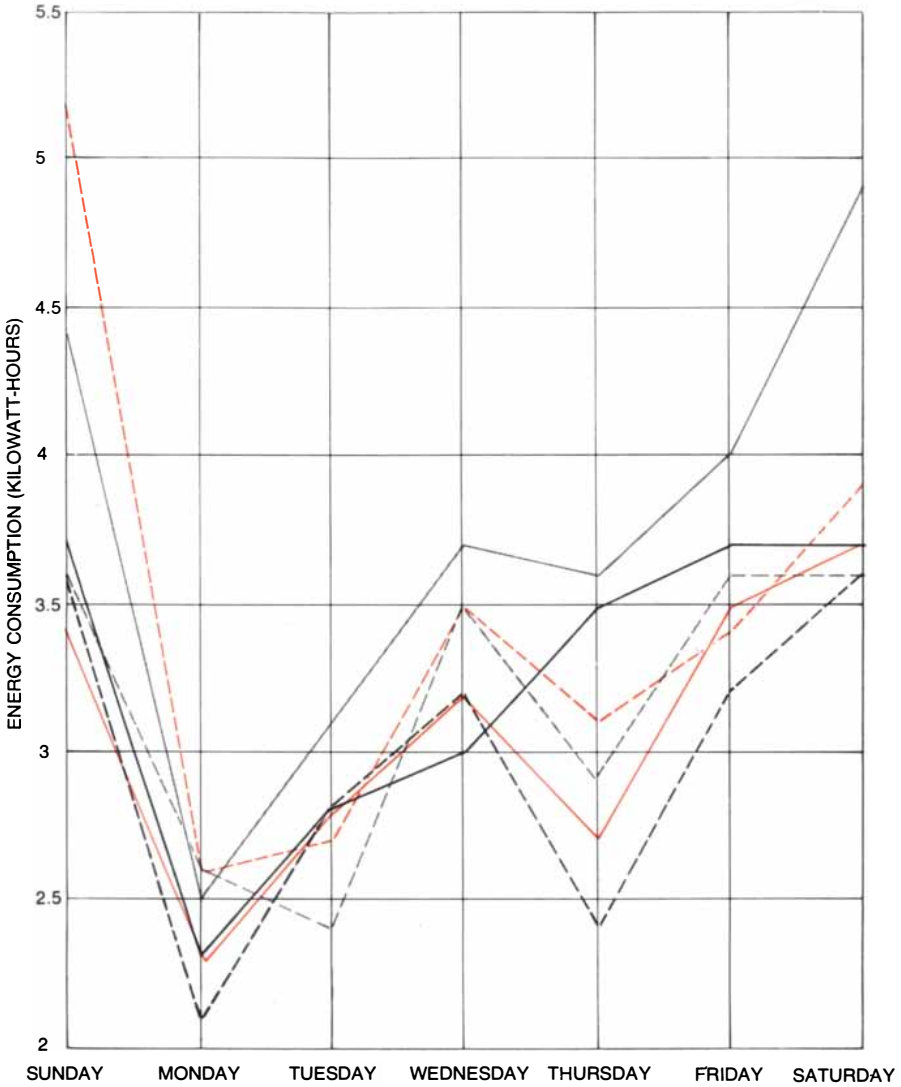
TWO TOASTERS are compared. In each pair of drawings a conventional toaster is at the left and an improved model built by the Bureau of Standards is at the right. In the standard toaster the heating wires are not insulated, so that a user who attempts to remove a piece of stuck toast with a metal instrument while the toaster is plugged in can be electrocuted. The toaster also gets hot because it has thin metal walls with no openings. In the improved toaster the heating wires are insulated from direct contact and openings are provided at the sides as well as the top and bottom to create a flow of air that reduces the temperature of the outside. In the bottom pair of relatively schematic drawings a number of external temperatures of the two toasters after five consecutive toasting cycles are compared in degrees Celsius.

encounters the even larger issue of how the heat and water vapor from the dryer are used (or not used) in the home. Often the warm, vapor-bearing air is vented to the outside. The volume of this air must then be made up by drawing outside air into the house through its windows or through the leaks that are to be found in every house.

The practice is wasteful in two ways. In cold weather warm and humid air from the dryer is wasted outside, and the cold, dry air that is drawn inside must be heated (and perhaps humidified) by the heating system. In hot weather the air from the dryer is not wanted inside, but the outside air drawn in to replace it must be cooled and dehumidified by the air-conditioning system (assuming there is one). It would be more practical to supply the dryer with outside air only, continuing to vent the exhaust to the outside. In cold weather it might be useful to supply and vent the dryer inside, using only interior air, provided that the procedure did not raise the relative humidity unduly.

These considerations are not trivial. The total volume of air passing through a dryer in the process of drying a three-kilogram wash is 130 cubic meters. The air volume of a two-story, three-bedroom house is approximately 540 cubic meters. Therefore when the dryer is vented outside, it exhausts nearly a fourth of the air in the house, air that at most times of the year already represents an investment of energy in heating or air conditioning.

A more general problem is revealed here. We usually deal with appliances one by one, buying and installing them individually and replacing them individually. If one wants to optimize the performance of the complex of household machines, however, it is necessary to consider their effects on one another. With energy costs rising, the problem of managing simultaneously the exhaust of warm, moist air from the dryer and the warm, dry air wasted from the refrigerator and the oven becomes important. A poor choice may make no more sense than running the furnace and the air conditioner at the same time.



VARIATION AMONG COOKS is revealed in this chart. Each line represents the energy required by one of six cooks who, in a test conducted by the Bureau of Standards, prepared three meals per day for seven days on identical electric ranges. The menu was different for each meal but the same for all six cooks for any given meal. The variation in the energy that the cooks used is so substantial that it outweighs variations likely to be found in the efficiency of stoves.

Growing concern over the economical use of energy means that the consumer is likely to be increasingly interested in the lifetime cost of an appliance. Here an important piece of information is missing. It is the expected useful lifetime of the appliance. To begin with, it is not easy to define "expected useful lifetime." However one defines it, numbers are not available for most products. Even if such a number were known for a product operating under rigidly specified conditions of use, the actual use would be found to vary so much from one owner to another and from one geographic area to another that a statement of the average lifetime and operating cost of an appliance might have little value for a prospective buyer and might even be misleading.

Some information is available on the average lifetime through a study made by the Department of Agriculture. Abstracting from the data, one finds that the service lifetime ranges from 10.7 years for a black-and-white television set to 20.4 years for an electric freezer. Between them, in ascending order, are a washing machine, a dishwasher, a color television set, a freestanding electric range, a gas clothes dryer, a freestanding gas range, an electric clothes dryer and a refrigerator. The authors of the study defined service lifetime as the life expectancy for a new appliance under one owner.

On the question of variations in use, which of course have a bearing on energy requirements, little is known. In an effort to develop at least one body of reliable information on this subject the Bureau of Standards recently ran a test involving six cooks using stoves and ovens. (The aim was to provide the basis for a standard method of testing the use of ranges.) Each cook prepared three standard meals per day for seven days, with a different menu for each meal. The variations in cooking habits, as reflected in the amount of energy required to operate each stove and oven, were considerable [*see illustration above*]. At times the energy used by the least conservative cook exceeded the pattern of the most conservative one by 50 percent. It seems that such variations in cooking habits may far outweigh the

variation in efficiency of ranges in terms of the effect on energy requirements. Much more information is needed about variations in the use of appliances in order to make rational estimates of the total cost of operating a particular appliance.

The safety of products is an aspect of performance that now, by law, requires the specific attention of manufacturers and Federal agencies. Several years ago, when this issue first began attracting Government action, the electric and thermal safety of the kitchen toaster was suggested as the object of a brief study. Several toasters bought at random were examined by the Bureau of Standards. Every one of them turned out to be a potential source of a serious electric shock to a user unwise enough to try to remove a stuck slice of toast with a knife or fork while the toaster was plugged in. Moreover, all the toasters became either painfully or dangerously hot to the touch after one or two toasting cycles. One model was hot enough after five toasting cycles to melt tin-lead solder.

It is fairly easy to design a satisfactory toaster that can neither electrocute nor burn the user. Such a model was built for demonstration purposes by the Bureau of Standards. The features that ensure its safe performance include insulating its heating elements electrically but not thermally and accommodating internal flows of air to reduce the temperature of the jacket. They add little to the cost of the appliance. Nevertheless, no manufacturer is now known to offer such a toaster for sale.

The work with toasters led to an interest in the general problem of specifying surfaces that could become hot without inflicting serious injury if they were touched. To this end the Bureau of Standards has developed the thermesthesiometer, an instrument that is the thermal analogue of a human finger. The cylindrical core of its probe is made of rubber with thermal properties equivalent to those of human flesh. An electric heater maintains the cylinder at finger temperature (33 degrees Celsius). A thin-ribbon thermocouple is embedded .1 millimeter below the surface of the tip of the probe (corresponding to the depth of the dermal layer of the skin) in order to measure the temperature that would be felt by a finger. When the probe is touched to a hot surface and held there, control circuits automatically yield a reading of the temperature reached at the end of a specified period of time.

The thermesthesiometer may provide the basis for a simple safe-temperature standard for products. For any surface to which it can be applied the instrument answers the safety question without regard specifically to the temperature of the surface (a wood object and a steel object at the same temperature can be respectively harmless and dangerous to the touch) or to the shape of the object, the kind of material it is made of and the thickness of the material.

Work has also been done at the Bureau of Standards on the sharpness of points and edges, mainly in order to set standards for the safety of toys. The work has included assembling data on the penetrability of skin by points of various design and by different kinds of edges. Data of this kind do not by themselves settle any questions. If a geometric criterion for a safe point or a safe edge is to be chosen, the choice must be made by whatever authority has the responsibility for regulating the hazards of products with sharp points or edges. That authority can be guided, of course, by the kind of data the Bureau of Standards has assembled. The main reason for producing such information has been to serve as a guide for agencies such as the Consumer Product Safety Commission.

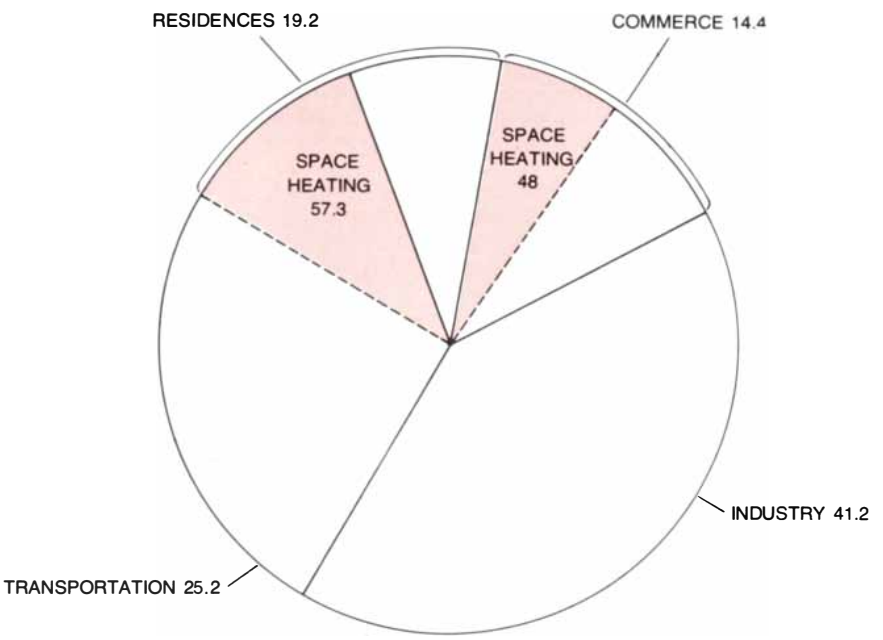
Once a geometric safety criterion has been chosen one can devise methods for testing whether or not a product meets the standard. For example, we have developed a hand-held tester for points. A point sharper than a specified "dangerous point" will penetrate the device, and the penetration is signaled by a light. We have also devised an instrument for testing sharpness. A strip of plastic tape of known physical properties is fixed around a steel mandrel. Held with a known force against the edge to be tested, the mandrel is turned through one revolution. The tape is then examined for cuts. If it has been cut through at any point, the edge is deemed to be too sharp for safety

On some questions of safety it is at present impossible to write complete specifications for products because the required information is not known. An example is provided by protective helmets. The relation between a mechanical insult (such as a blow on the head or rapid deceleration of the head) and the probability of physical injury is not adequately established.

Much of the research on the relation between insult and injury has concentrated on events that produce concussion, a symptom usually described by physicians as including temporary disorientation, perhaps with loss of visual control or speech. It is less severe than prolonged unconsciousness. Concussion thus defined is the symptom frequently observed following accidents in sports; it is widely regarded as the threshold injury to be prevented by protective headgear.

The problem is that the mechanical causes of concussion are a matter of disagreement. Some workers believe the important quantity is linear acceleration of the head; others maintain that it is not linear acceleration but rotational acceleration, or possibly a combination of the two.

This lack of fundamental understanding has not blocked suggestions for testing helmets. A drop test devised for mo-

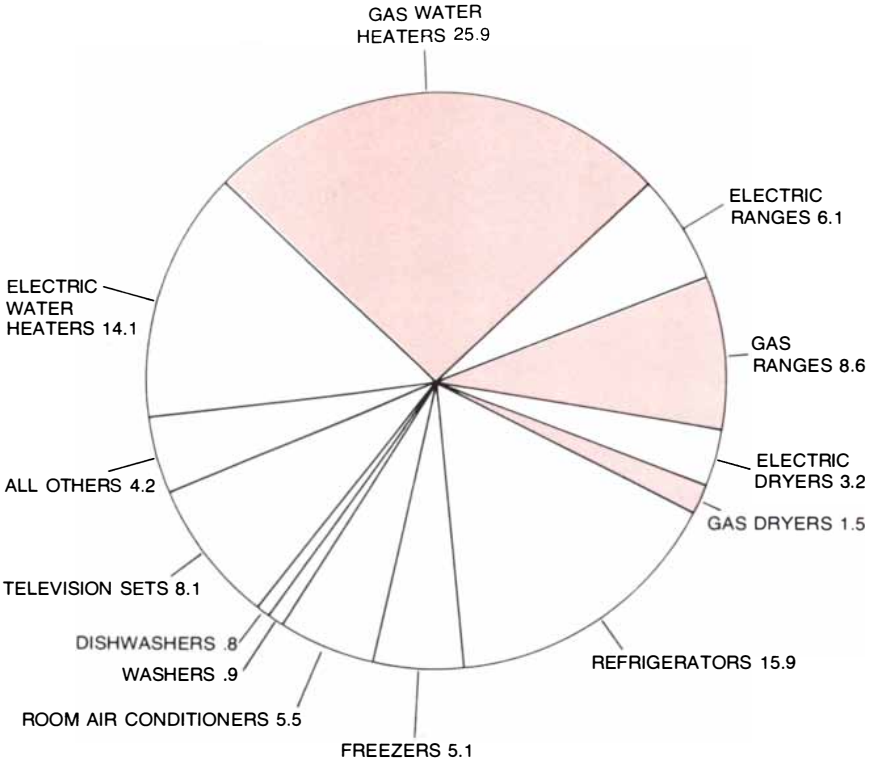


DISTRIBUTION OF PRIMARY ENERGY in the U.S. is charted in percent. Primary energy is defined as the energy originally contained in the fuel (oil, coal or gas). In the portion of the diagram relating to residences the amount of energy that is devoted to running household appliances is included in the smaller section of space representing residential energy consumption.

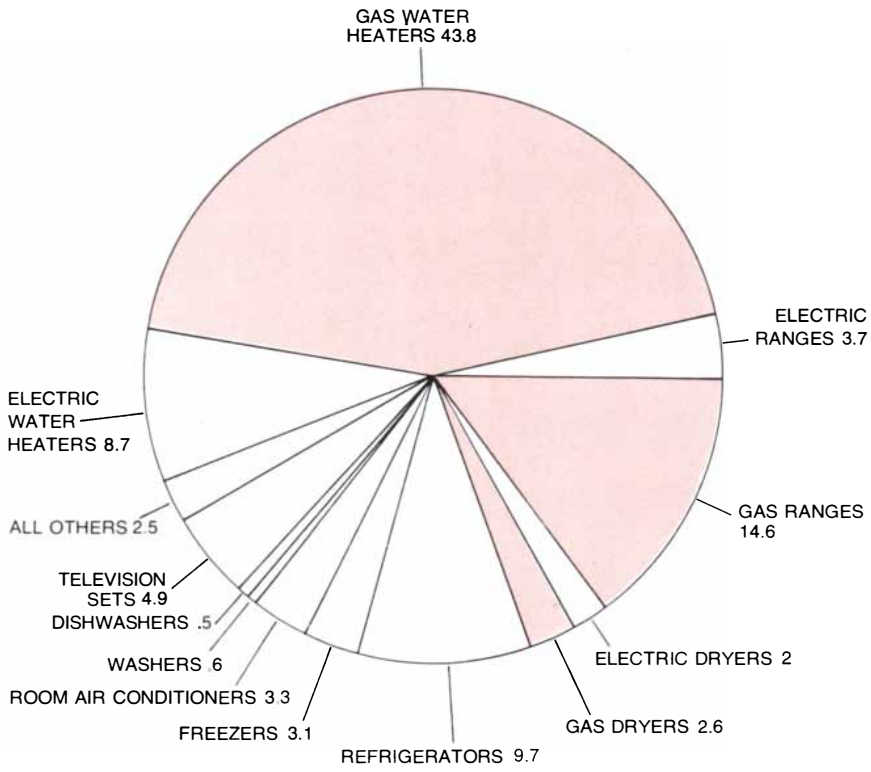
torcycle helmets by the American National Standards Institute has been employed extensively to see if it could be useful as a means of ranking football helmets. The tests have shown that football helmets do differ not only in their ability to reduce the maximum linear deceleration that would be experienced by the head in a fall but also in their ability to absorb part of the energy that would be delivered by a blow to an unprotected head. No one has yet justified a specification for these attenuations of acceleration or energy, however, and the mechanical features that would optimize the protective function of a helmet have not been explored scientifically. Adding to the problem is the possibility that certain features of helmet design, although they help to protect the wearer against some injuries, may increase the possibility of other injuries, both to the wearer and to others in the course of play.

The need for the work I have described was identified several years ago by Lewis M. Branscomb, then director of the Bureau of Standards. Concerning the lifetime cost of products, he said: "There are three possible solutions to the dilemma of initial cost versus lifetime cost. The first is greater rental of products and appliances, with the supplier assuming all responsibility for maintenance, as the telephone company does now. Such a practice can be guaranteed to bring about a swift change in engineering design that will maximize product life and minimize total costs. The second is a liberalization of warranties to cover a major part of the designed useful life, with the consumer understanding that the chances are very high that replacement will be necessary by the time the warranty expires. And the third is to provide the consumer with factual information about cost factors, to educate him to the merits of paying more for a product that will cost less to own and operate, and to help him to make an intelligent decision. A trend toward any or all of these would give the engineer a new challenge, would increase the efficiency of our economy and would reduce the rate of conversion of products to trash."

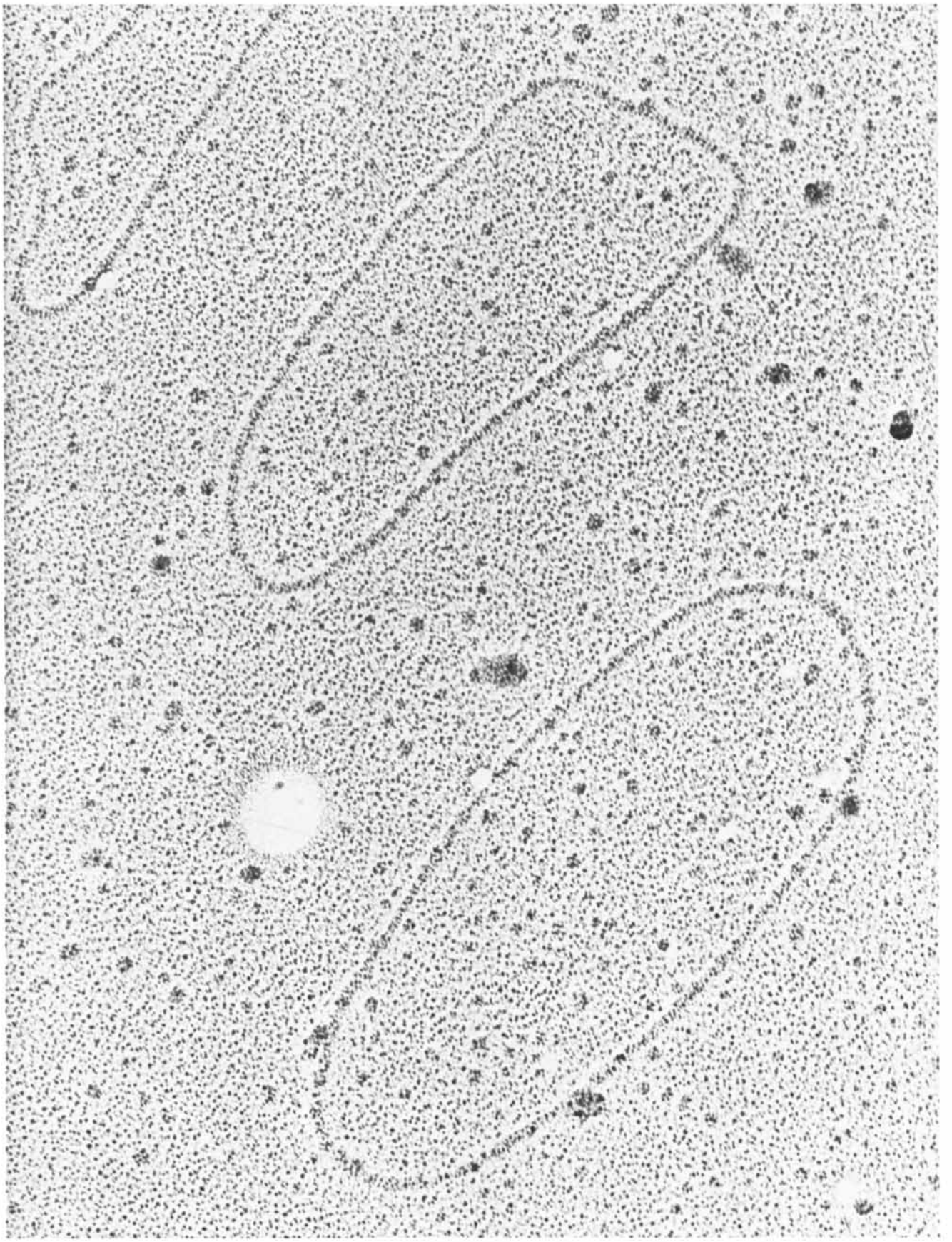
Concerning the safety of products, Branscomb said: "Technology has made an enormous contribution to American health and longevity. But we also pay a heavy price in injury and death from unsafe products.... No one expects all products to be idiot-proof. But doesn't the buyer have a right to know the nature of the hazards to which he may expose himself, his family and his neighbors? Rising public expectations of this kind require imaginative design engineering and conscientious advertising and marketing."



HOUSEHOLD APPLIANCES consume primary energy as depicted in this chart. From the viewpoint of the householder, however, the distribution is somewhat different because the primary fuel that has already been converted into electricity must be taken into account. As electricity the fuel yields less energy than it does when it is burned as a source of heat. The distribution of energy among appliances, as seen in the household, is shown at the bottom of this page.



HOUSEHOLD DISTRIBUTION of energy for appliances shows large amounts for the devices that burn primary fuel, such as gas stoves and water heaters. The differences between the end-use figures and the primary-energy distribution among appliances reflect the differences in fuel-conversion efficiency between burning fuel directly and converting it into electricity.



DNA OF THE BACTERIAL VIRUS ϕX174 is a circular molecule about 1.8 micrometers long. In this electron micrograph made by G. Nigel Godson of the Yale University School of Medicine double-strand molecules of the DNA are enlarged about 180,000 diameters. The DNA is present within a virus particle as a single strand. In an infected cell the single strand is copied to make two strands.

The Nucleotide Sequence of a Viral DNA

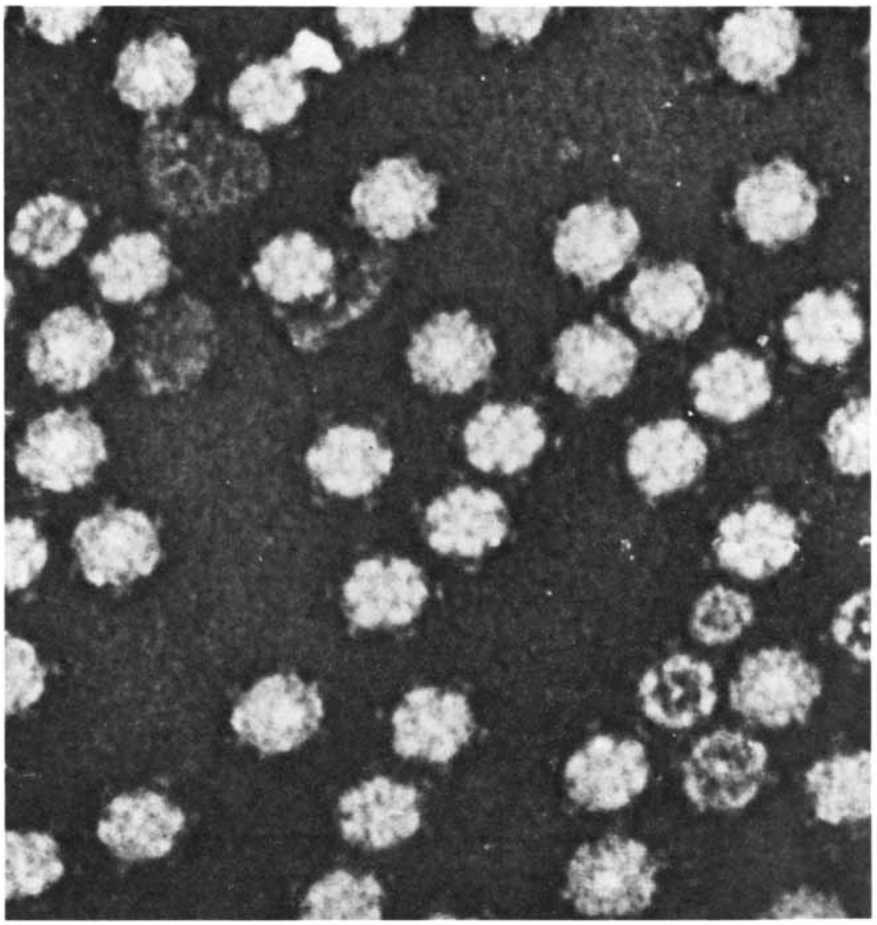
Genetic information is encoded by the order in which nucleotides are arrayed to form a strand of DNA. Now that order has been established in full for the 5,375-nucleotide DNA of the bacterial virus ϕX174

by John C. Fiddes

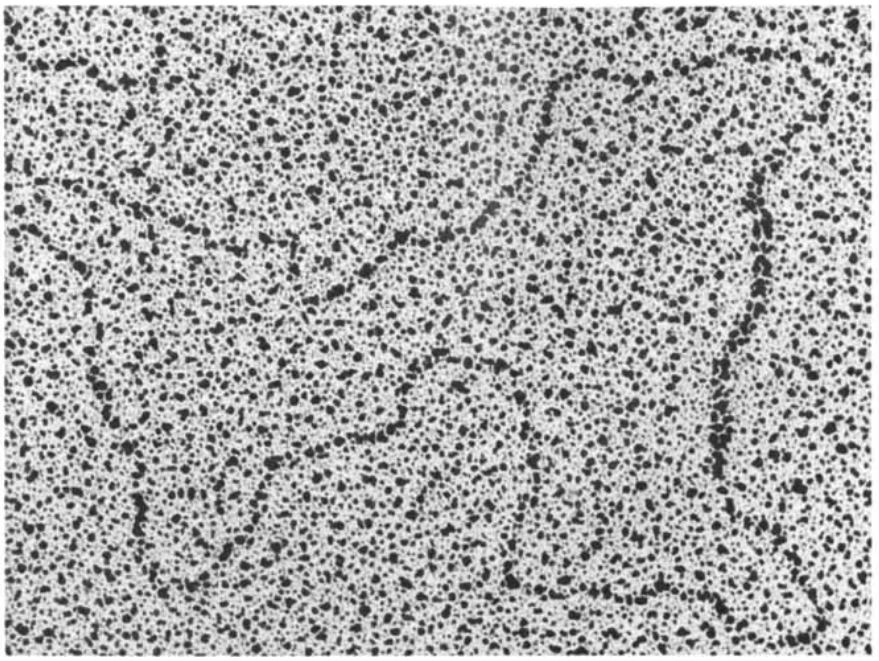
Molecular biology has come a long way since DNA was identified in the 1940's as the genetic material. In the 1950's and 1960's the structure of the DNA molecule and its mode of replication were determined, revealing the chemical basis of heredity. In the past 20 years the mechanisms and processes by which the hereditary information encoded in the sequence of DNA nucleotides is translated into proteins, and thus into the substance of all living things, have become understood in increasing detail. In all this time molecular biologists have been at a serious disadvantage: they did not know the complete sequence of nucleotides constituting the genome, or total genetic message, of any organism.

Such a message is now available: the complete nucleotide sequence of the DNA of a small bacterial virus, ϕX174, has been established. It is a short message, with only 5,375 nucleotides compared with the millions in a bacterial chromosome or the billions in the chromosomes of a mammalian cell. Because the message is complete, however, it is now possible to relate the genetic information of an organism to its proteins and functions more directly than before. In this article I shall tell how the ϕX174 sequence was established and report some of the new perceptions of the remarkable coding ability of DNA that the analysis of the viral genome has already yielded.

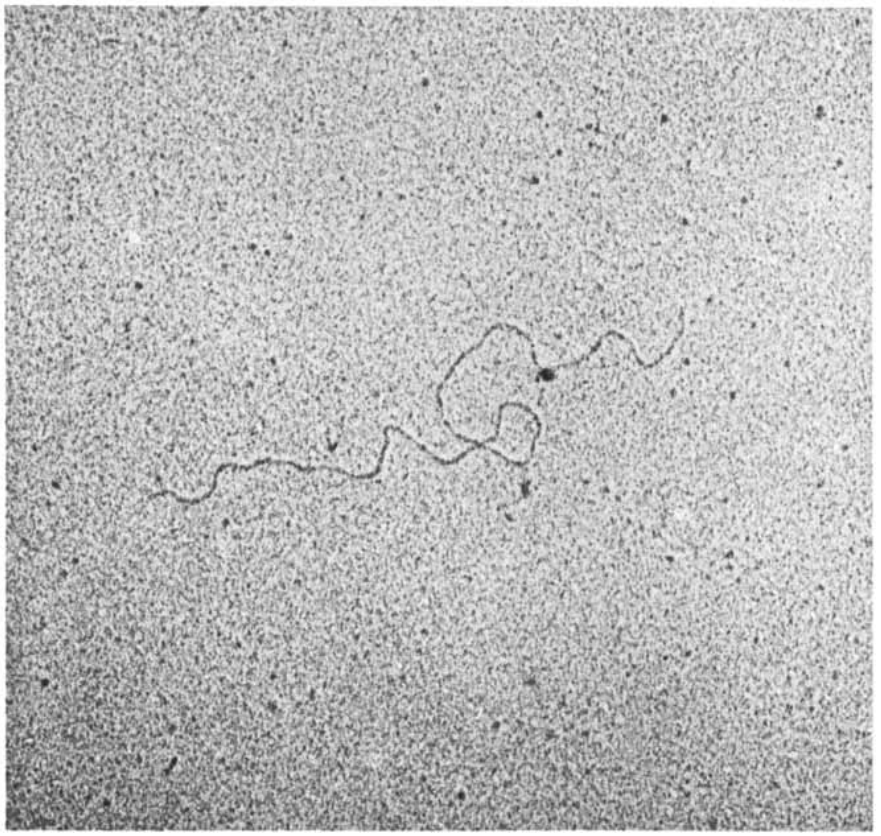
A molecule of DNA is a long strand made up of four kinds of nucleotides: deoxyadenylate (*A*), deoxyguanylate (*G*), thymidylate (*T*) and deoxycytidylate (*C*). Each nucleotide in a DNA molecule has three components. Two of them, a five-carbon sugar ring (deoxyribose) and a phosphate group, alternate to form a continuous backbone. The other component is one of four nitrogenous bases: adenine, guanine, thymine and cytosine. The bases are side chains extending from the backbone, and in the double-helix form in which DNA is present in most organisms the two strands of the helix are joined by hydrogen bonds linking the bases on each strand. (The DNA in some viruses, including ϕX174, is single-strand until it is injected into a host cell.) The strands are complementary, because the size and bonding properties of the bases ordain that *A* be always opposite *T* and that *G* be opposite *C*. The complementarity achieved by base-pairing is the basis of



ϕX174 VIRUS PARTICLES are enlarged about 400,000 diameters in this electron micrograph made by John Finch of the British Medical Research Council's Laboratory of Molecular Biology in Cambridge. The virus's capsid, or protein coat, is an icosahedron, a polyhedron with 20 faces. Each of 12 vertexes bears a spike by which virus can attach itself to a bacterium.



SINGLE-STRAND FRAGMENT cleaved from viral DNA by a restriction enzyme has been "annealed" to a single-strand viral-DNA molecule and is seen as a denser portion of the loop (*right*) in this micrograph made by Godson and W. Keegstra of the University of Utrecht. A "primer" fragment is annealed to viral DNA in this way to establish the nucleotide sequence.



VIRAL PROTEIN encoded by gene *A* of the viral DNA plays a role in replicating the double-strand form within the bacterial cell. It begins by attaching itself to a particular site on the viral strand. The location of that site, and thus of the origin of DNA replication, was identified with respect to one end of the DNA by Shlomo Eisenberg, Jack D. Griffith and Arthur Kornberg of the Stanford University School of Medicine. The electron micrograph, made by Griffith, shows the gene-*A* protein (*black spot*) about a fifth of the way from the end of the strand. Although the strand normally forms a closed loop, here it has been opened by treatment with an enzyme.

the faithful replication of DNA and hence of the transmission of genetic information from one generation to the next. It is the particular sequence in which the four kinds of nucleotides are aligned along a strand of DNA that embodies the genetic message. Some stretches of the sequence specify, by means of the genetic code, the sequence in which amino acids are to be assembled to form the organism's proteins. Other stretches do not code for proteins but include signals that regulate the turning on and off of particular genes, or coding regions, and the rate at which those genes are to be expressed.

Until recently it was a formidable task to learn the nucleotide sequence of a DNA molecule because the techniques available for DNA-sequence analysis lagged far behind those that had been applied since the mid-1960's for analyzing RNA, the very similar nucleic acid that has several roles in the translation of the DNA message into proteins. Techniques for sequencing RNA rely on the ability of several enzymes to cleave the RNA at specific sites and thus to produce short pieces of RNA whose sequence can be established readily. It was difficult to apply a similar method to DNA. The smallest DNA molecules, those of certain viruses, are perhaps 70 times longer than the 75-nucleotide transfer-RNA molecules that were the subject of early RNA sequencing. And no enzymes are known that cleave DNA with the same degree of single-base specificity as the ribonucleases that are available to cleave RNA.

In the past few years, however, two new methods have been devised for sequencing DNA, both of them very different from the classical RNA-sequencing methods and both much faster. In 1975 Frederick Sanger and Alan R. Coulson of the British Medical Research Council's Laboratory of Molecular Biology in Cambridge published a report on their "plus-and-minus" method, and early this year Allan Maxam and Walter Gilbert of Harvard University published their "chemical" method. Both of the new techniques have since been exploited intensively. Here I shall be describing the plus-and-minus method, the principal one by which Sanger's group established the ϕX174 sequence. (The group consisted of Sanger, Gillian M. Air, Barclay G. Barrell, Nigel L. Brown, Coulson, myself, Clyde A. Hutchison III, Patrick M. Slocombe and Michael Smith.)

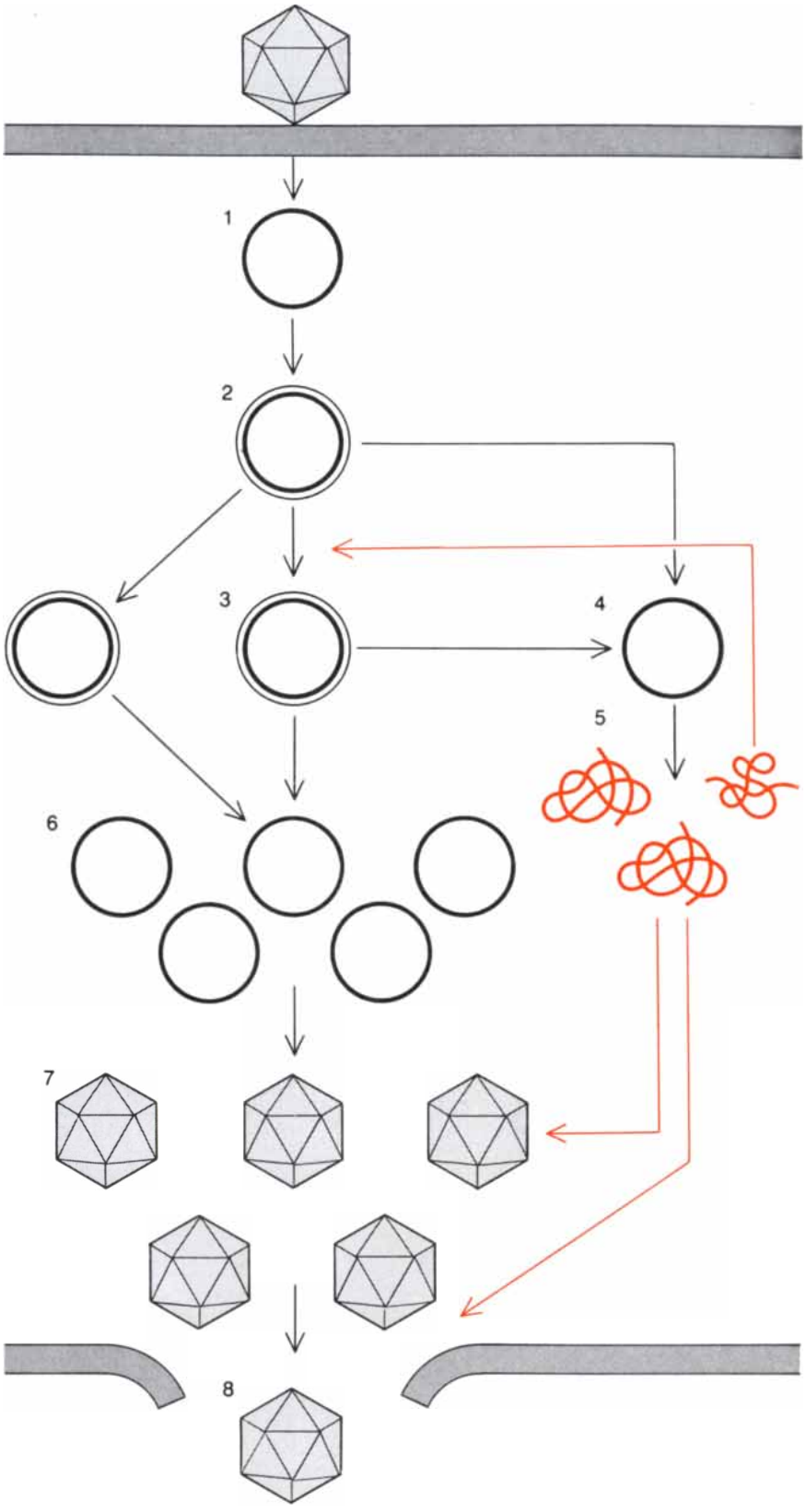
In broad outline the plus-and-minus method calls for synthesizing pieces of DNA that are complementary to various regions of the DNA being sequenced. Because of the way the synthesis is manipulated each piece ends either with or just before a particular, identifiable nucleotide (*A, G, T* or *C*). By per-

forming the synthesis to produce a large number of pieces, each one ending with a successive nucleotide in the sequence, it is possible to establish the sequence of the entire DNA molecule.

The synthesis is brought about through the agency of a DNA polymerase, an enzyme that adds successive nucleotides to a DNA strand by linking the 3′ carbon atom on the sugar ring of the nucleotide at the end of the strand to the 5′ carbon atom on the sugar ring of the next nucleotide in the chain [*see illustration on next page*]. DNA polymerases cannot start copying a template DNA at just any point. They must be directed to a specific position by means of a "primer": a DNA molecule that has a sequence complementary to the sequence of a particular region of the template and that therefore "anneals" itself to the template by base-pairing. A free hydroxyl (OH) group at the 3′ end of the primer acts as a substrate for the DNA polymerase, which catalyzes the addition of the nucleotide called for by the base-pairing rules, thereby extending the chain in the 5′-to-3′ direction.

In the first step of the plus-and-minus procedure millions of molecules of the template DNA are incubated with DNA polymerase and with the four building blocks of DNA: the highly energetic deoxynucleoside triphosphates. (Two of the phosphates are dropped as the bond is formed.) At least one of the nucleoside triphosphates is labeled by incorporating into it a radioactive atom of phosphorus. The extension of the primer molecule is done under carefully controlled conditions at a low temperature calculated to retard the enzyme's rate of progress. Because under such conditions the strands do not all grow at exactly the same rate and because the mixture is sampled frequently, every possible chain length is represented among the products of the extension process.

That this is so can be established by electrophoresis, a process that can sort molecules according to their size. The pieces of double-strand DNA that include the extension products of the primed reaction are placed on a slab of polyacrylamide gel containing a high concentration of urea, which separates the newly synthesized extensions from their template DNA. Under the influence of an electric current the extension molecules migrate through the gel. The distance they travel varies with the size of the molecules, so that all the molecules of a given size end up in a small group. The groups are visualized by autoradiography: since some nucleotides added to form the extension molecules were labeled with radioactive phosphorus, when a piece of X-ray film is placed in contact with the gel, a pattern of

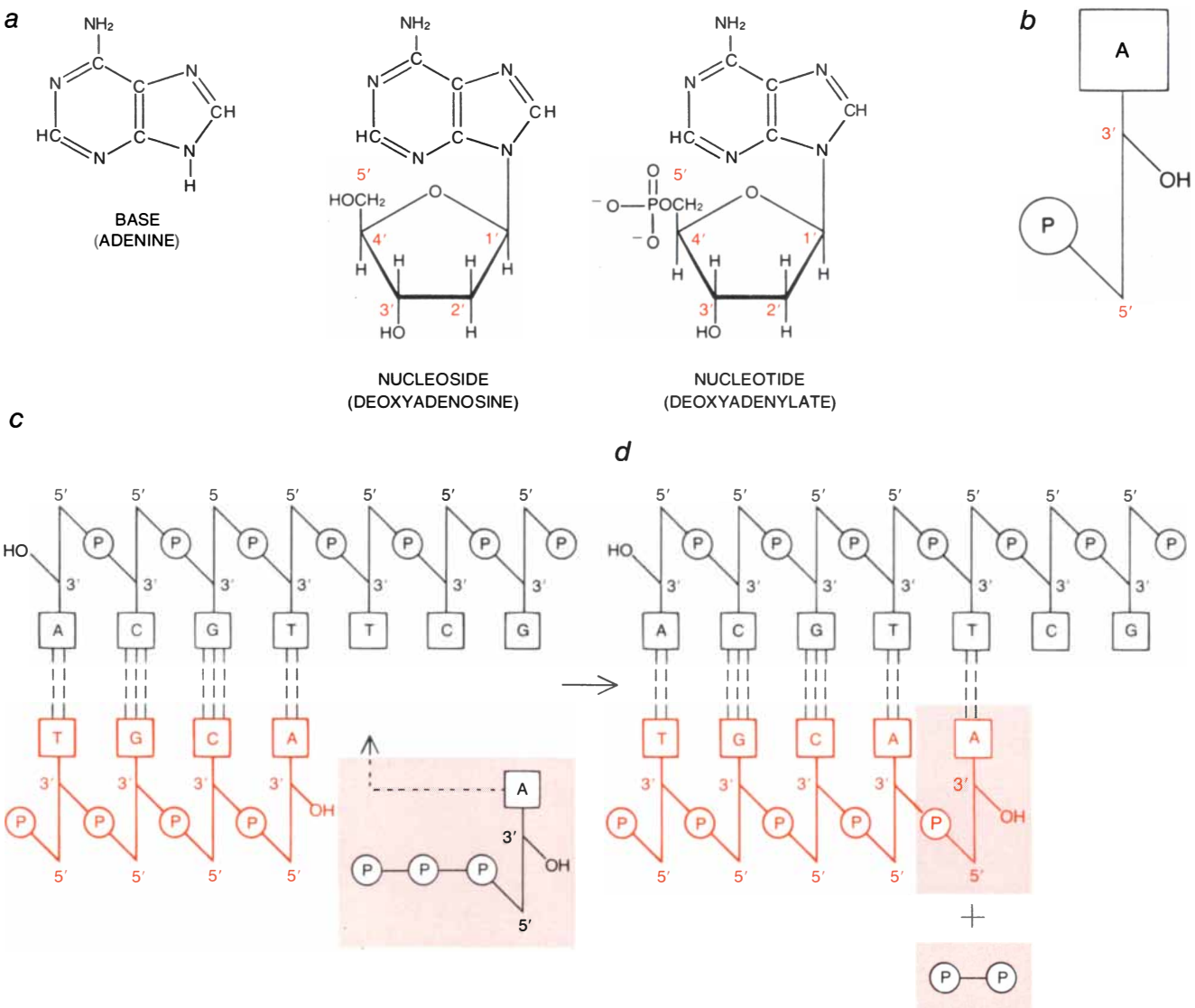


VIRAL DNA injected into a cell by an infecting virus (*1*) serves as the template for synthesis of a complementary strand (*2*). The double-strand DNA replicates to form about 20 copies (*3*). The complementary strand of the double-strand DNA serves as the template for the synthesis of messenger RNA (*4*) that is translated to make viral proteins (*5*); the proteins have enzymatic and structural roles in DNA replication, the formation of new virus particles and lysis (dissolution) of the cell wall (*colored arrows*). Some 200 copies of the single viral-DNA strand are produced (*6*), packaged in a protein capsid (*7*) and released (*8*) after lysis of the cell wall.

bands appears on the film corresponding to the positions in the gel of the groups of DNA molecules. Each successive band represents a molecule one nucleotide longer than the molecule represented by the preceding band. The purpose of this autoradiograph of the extension products is to create a standard, an array of bands representing every successive chain length. In essence the objective of the plus and the minus treatments is to create different populations of extension products: molecules that end not at every nucleotide position but specifically at or just before an *A*, a *G*, a *T* or a *C*.

First it is necessary to pass a sample of the products of the initial extension over a gel filtration column and thus to separate the relatively large extension chains attached to their templates from the small individual nucleoside triphosphate building blocks that were not incorporated into new DNA. This is important because the results of the plus or the minus treatment will depend entirely on the absence of specific building blocks in the subsequent reactions. Then the reaction mixture is split into eight subsamples, one subsample for each of the four nucleotides in a plus system and one for each in a minus system.

Each subsample to be given the minus treatment is incubated with DNA polymerase and with only three of the four nucleotide building blocks. In the minus-*A* system, for example, the extension molecules, still bonded to their templates, are incubated with the enzyme and with the *G*, *T* and *C* nucleoside triphosphates but without the *A* nucleoside triphosphate. The synthesis resumes, but now the building blocks are added to each chain only up to the position just before the one where an *A* is called for by the base-pairing requirement; the synthesis stops at that point because no *A* building blocks are available. When these minus-*A* chains are subjected to the electrophoretic size-fractionation described for the original extension products, one does not observe a radioactive band at every position on the autoradiograph. Only a limited number of bands appear, each one representing the position in the sequence just preceding an *A* nucleotide. In the three other minus systems the three oth-

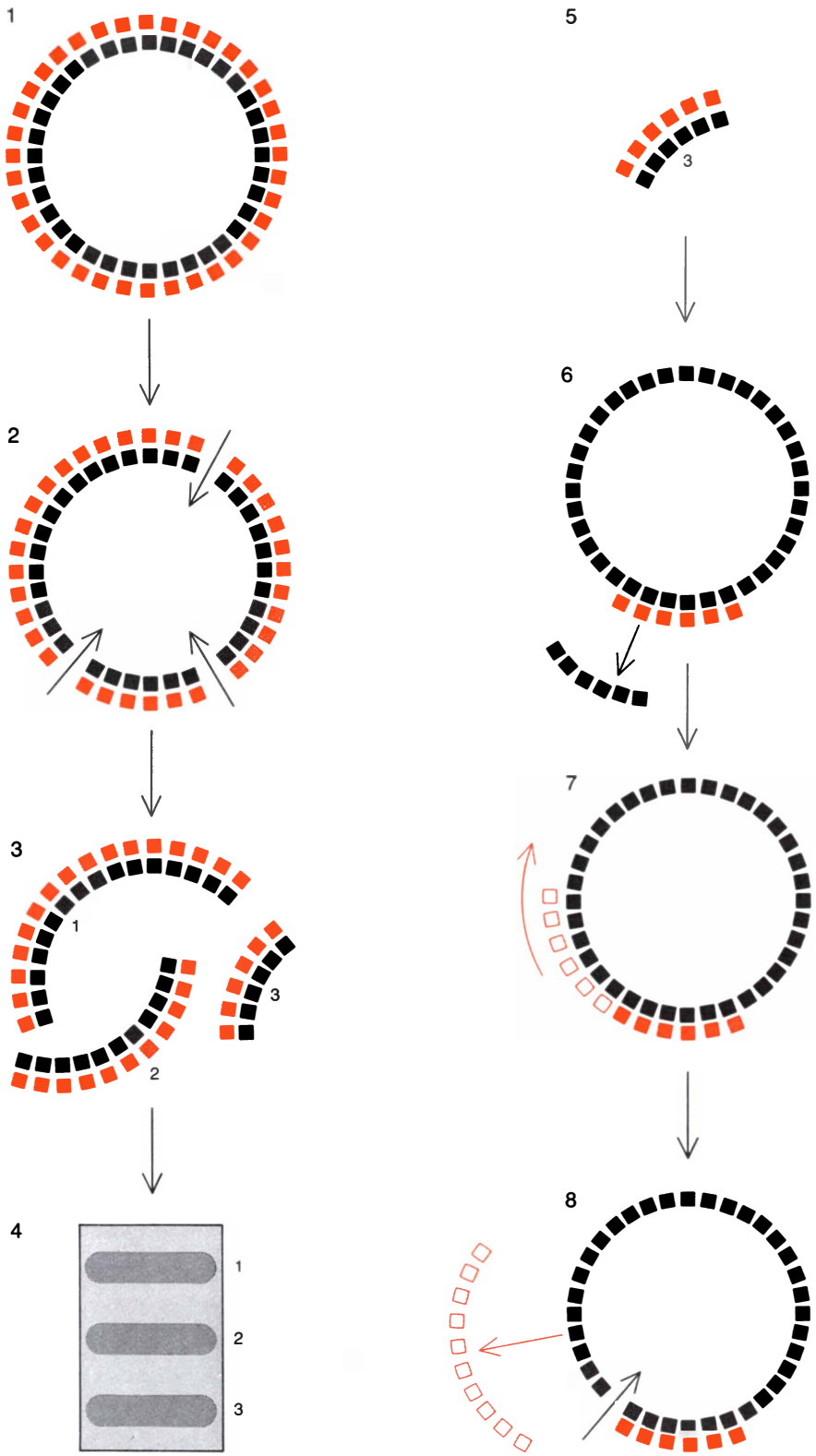


NUCLEOTIDE (*a*) consists of a base (adenine in this case), a sugar and a phosphate group. The base and the sugar constitute a nucleoside; the addition of a phosphate forms a nucleotide. A nucleotide can be represented as shown (*b*): the vertical line is the sugar, which carries the base (*A*) as a side chain and is linked by the phosphate (*P*) at its 5′ carbon to the 3′ carbon of the adjacent nucleotide's sugar. In a double-strand DNA molecule (*c*) hydrogen bonds (*broken lines*) between complementary bases link the two strands; *A* always pairs with *T* and *G* always pairs with *C*. A complementary strand (*color*) is synthesized along a template strand (*black*) by DNA polymerase; the enzyme links one phosphate of appropriate deoxynucleoside triphosphate (*colored panel*) to hydroxyl (*OH*) group at 3′ end of chain (*d*).

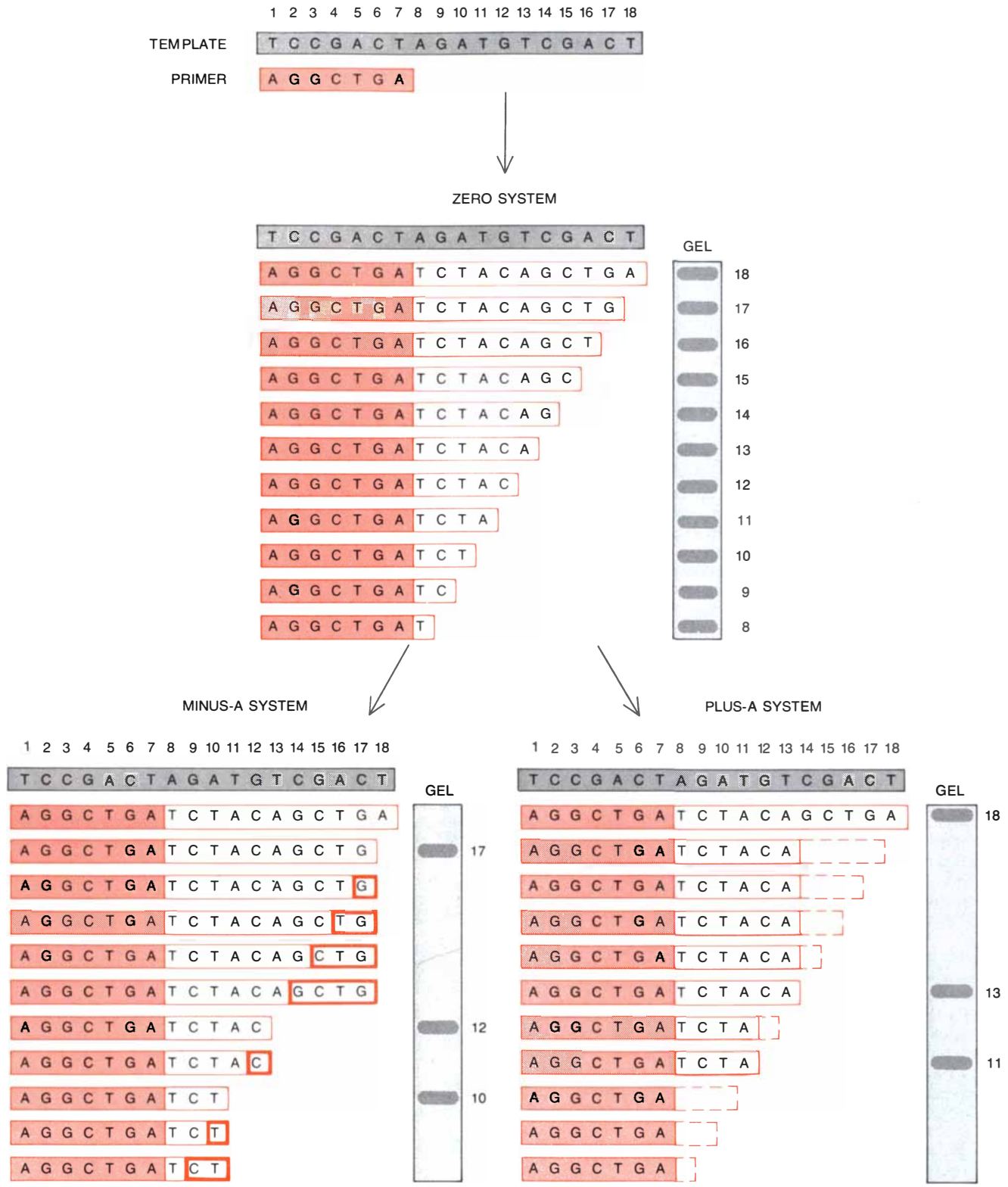
er subsamples of extension products are similarly extended in the absence of *G, T* and *C* nucleoside triphosphates; when the products of those reactions are fractionated, only the bands preceding each *G, T* and *C* are present on the autoradiograph.

In the four plus systems, also designed to yield chains ending at defined positions, extension-product subsamples are incubated with only one nucleoside triphosphate instead of with three and with a different DNA polymerase, one that is manufactured by the bacterial virus designated *T4*. This enzyme has a special property: in addition to synthesizing DNA chains in the 5′-to-3′ direction under certain circumstances, in other circumstances it degrades the DNA, sequentially removing nucleotides from the 3′ end of the chain. The *T4* polymerase degrades the original extension molecules back toward their primer. In the plus-*A* system, however, the *A* nucleoside triphosphate is present, and so the degradation of each molecule ceases at the first *A* the polymerase encounters; at these positions the easy availability of *A*'s allows the enzyme's polymerizing activity to dominate its degradative activity, so that an *A* nucleotide that is removed is immediately replaced by a new *A* nucleotide. The plus-*A* reaction hence produces populations of molecules all of which end with an *A*. Similarly, the plus-*G*, plus-*T* and plus-*C* reactions produce molecules terminating in *G, T* and *C*. Again molecules of each length are grouped by electrophoresis, and the autoradiograph bands indicate each molecule's length and thus the position of each *A, G, T* and *C*.

The extension products of the four minus systems and the four plus systems are fractionated along with the extension products of the initial "zero" system in adjacent tracks on a single gel [*see illustrations on pages 61 and 62*]. The nucleotide sequence of the extensions can be read off the gel because most of the nucleotides are represented by a band at the appropriate position in one of the plus and one of the minus systems. (In the case of a "run" of a particular nucleotide, for example a sequence such as *AAAAA*, bands are not present at every position. In the plus system only the band representing the end of the run is present; in the minus system one observes only the product representing the start of the run. The precise length of the run can be determined easily, however, by reference to the number of bands in the corresponding area of the zero-system column.) Under particularly favorable circumstances the sequence can be read on a single gel out to 200 or more nucleotides from the 3′ end of the primer. A sequence several hundred nucleotides long, which might have taken



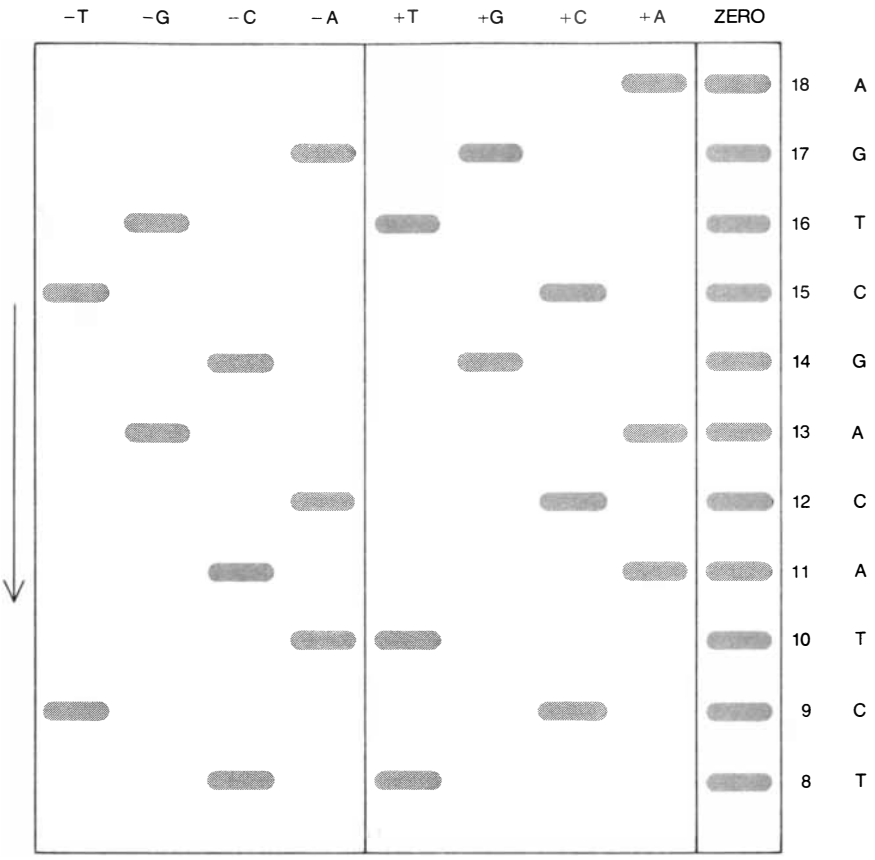
PLUS-AND-MINUS METHOD of determining a nucleotide sequence begins with double-strand DNA taken from infected cells (*1*). Cleavage by a restriction enzyme (*2*) produces fragments (*3*), which are separated according to size by gel electrophoresis (*4*). One of the fragments is selected (*5*) to serve as a primer for the sequencing procedure. Incubated with a single-strand viral-DNA template, the primer's complementary strand is annealed to the template by base-pairing (*6*). Now new DNA is synthesized: successive nucleotides (*open colored blocks*) are added (*7*) under various conditions to make extension molecules, which are separated from viral DNA by the restriction enzyme (*8*) before being analyzed for nucleotide sequence.



PRIMER IS EXTENDED under three different conditions after being annealed to the template (whose nucleotides are here arbitrarily numbered from 1 through 18). In the "zero" system the primer and template are incubated with DNA polymerase and all four nucleotide building blocks under conditions such that successive nucleotides are added, according to the base-pairing rules, to produce extension products (*open colored boxes*) of every possible chain length. A subsample of these extension molecules is separated according to size by electrophoresis on a polyacrylamide gel (*gray slab at right*). Other subsamples are subjected to "minus" or "plus" treatment. In the minus-*A* system, for example, the *A* building block is withheld from the mixture; the extension products are further extended (*heavy boxes*), but only up to the position just before an *A*. Now electrophoresis reveals only three bands, each band representing the position in the sequence one nucleotide before the next *A*. The same procedure is followed with each of the other three nucleotides withheld from the synthesis process to establish the positions one nucleotide before each appearance of a *G*, *T* or *C*. In the plus-*A* system only *A*'s are provided, and the polymerase is one that degrades DNA except in the presence of an excess of a nucleotide building block. Extension products are therefore degraded (*broken boxes*) until position of an *A* is reached. Electrophoresis again shows three bands, this time at position of *A*'s.

years to determine by the classical method of cleaving into small fragments, can now be established in a few days.

An important tool for plus-and-minus sequencing and for molecular biology in general was provided by the discovery several years ago of the enzymes called restriction endonucleases, which cleave large DNA molecules into discrete fragments. They cleave DNA not at the site of a particular single nucleotide, as the ribonucleases mentioned above do, but only at particular sites in specific sequences of four to six nucleotides, and so they generate much larger fragments than the ribonucleases do. Usually the sequence recognized by a restriction enzyme is symmetrical in double-strand DNA. For example, one endonuclease recognizes the double-strand sequence consisting of *GGCC* in one strand opposite *CCGG* in the other strand and cleaves the symmetrical tetranucleotides at their midpoint. The fragments produced by an endonuclease have an exposed hydroxyl group at the 3′ end, so that they are suitable as primers for chain extension. More than 20 restriction endonucleases are known that cleave DNA at different nucleotide sequences. Treating different samples of a DNA with different endonucleases therefore produces fragments of various sizes and with known nucleotides at each end. Biochemical experiments can establish the position of each of these fragments with respect to one another on the original DNA, so that the choice of a particular fragment to be annealed as a primer determines the portion of the total DNA whose sequence will be elucidated.



EXTENSION PRODUCTS of the four minus and the four plus systems are subjected to electrophoresis side by side along with the products of the zero system. This is a schematic representation of autoradiograph that would result from procedures illustrated on opposite page. Each group of molecules travels down gel a distance proportional to its length and shows up as a gray band on autoradiograph. Nucleotide sequence (*right*) is determined from band pattern.

The ϕX174 virus infects the bacterium *Escherichia coli*. It is a small virus and one that has been well studied in many laboratories, notably those of Robert L. Sinsheimer at the California Institute of Technology and of Irwin Tessman and Ethel S. Tessman at Purdue University. The infectious virus consists of a protective capsid, or protein shell, containing a single-strand DNA molecule that carries the genes coding for the virus's nine proteins. In electron micrographs the viral particle appears as a symmetrical icosahedron (a polygon with 20 faces) about 250 angstroms in diameter with a knobby spike at each of the 12 vertexes. The viral protein designated *F* is the major component of the capsid; the *G* and *H* proteins form the spikes.

An infecting virus particle attaches itself to a specific receptor site on the surface of the bacterium. The single strand of DNA is injected into the cell. It is copied (by host-cell enzymes) to produce a double-strand form, which then replicates (apparently with the help of viral protein *A* as well as host enzymes) to make about 20 copies, each one consisting of a viral strand like the one originally injected and a complementary strand. The complementary strands serve as templates for transcription of the DNA into messenger RNA, which therefore has the same nucleotide sequence as the viral strand; translation of the messenger RNA manufactures the viral proteins. When enough proteins have been accumulated, the DNA replication is skewed so that only viral strands are formed—about 200 of them. Under the influence of viral proteins *B, C* and *D* these viral-DNA strands are packaged in capsids to produce mature virus particles, and then the cell wall is ruptured by the *E* protein, releasing the proliferated virus particles.

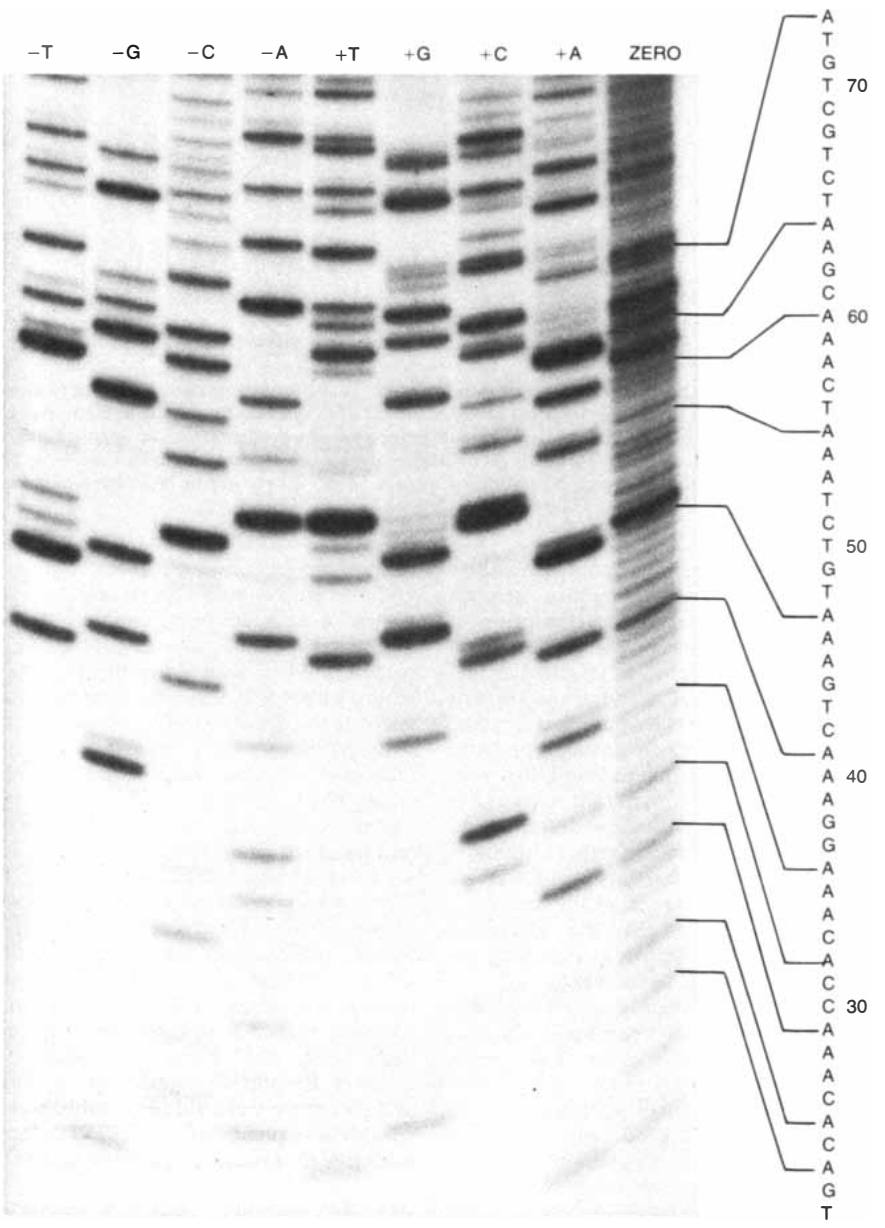
The ϕX174 virus was particularly well suited to DNA-sequence analysis because the relative order of the genes on its circular DNA molecule had been established by genetic-mapping experiments. Certain control sites had also been located, including sites where transcription into messenger RNA starts, the major site where transcription terminates and the approximate site where double-strand DNA replication begins. A physical map of the viral DNA was constructed by cleaving the strand into fragments with endonucleases of varying specificities, determining the order of those fragments and correlating them with the genetic map, so that any gene or part of a gene could be isolated as required on a small piece of DNA; such pieces could be selected as primers that would direct DNA synthesis to specific parts of the viral strand.

By priming with a large number of restriction-enzyme fragments, and with either the viral or the complementary strand or both of them in turn as the template, we determined the nucleotide sequence of many pieces of the ϕX174 DNA molecule. Advantage was taken of earlier sequencing work done by different methods in a few regions of the DNA and of the fact that the amino acid sequence of some of the viral proteins was known and could, through the genetic code, yield some nucleotide sequences. By putting together all the bits and pieces we were able to establish the complete sequence of the 5,375 nucleotides [*see illustration on pages 64 and 65*].

As the complete sequence became available we studied it for new information on the overall organization of the viral genome. The major surprise came as we located the nucleotides that constitute the "start" and "stop" signals for each of the viral genes.

DNA has usually been thought to consist of coding and noncoding regions. The coding regions, the genes, are the stretches of nucleotide sequence that specify the order in which amino acids are to be assembled to form proteins. This is accomplished by means of the genetic code, in which successive groups of three nucleotides constitute the codons, or code "words," for particular amino acids. For example, the DNA codon *GGT* specifies the amino acid glycine, *TAC* specifies tyrosine, *CAT* specifies histidine. Since the number of three-nucleotide combinations of four nucleotides is 4^3, there are 64 codons. There are only 20 amino acids, so that the code is "degenerate": most of the amino acids are specified by several codons. The coding information carried by DNA is expressed through messenger RNA, which has a sequence complementary to that of the DNA from which it is copied. (The codons are usually given in terms of RNA, but in working with ϕX174, in which the messenger RNA has the same sequence as the viral strand, we speak of the codons in terms of DNA.) Messenger RNA is translated into proteins by a complicated procedure involving the cell organelles called ribosomes.

Each coding region of the DNA starts with the sequence *ATG*, which is the codon for methionine, the initial amino acid in all protein chains. Not all methionine is at the start of a protein, however, so that clearly not all *ATG* sequences can be signals for initiating a protein; for an *ATG* to start a protein it must be preceded by another signal that tells the ribosome to begin to translate. These ribosome-recognition signals are sequences that are complementary to, and presumably recognized by, a region on one of the small RNA components of ribosomes. Since the extent of this complementarity can vary, not all ribosome-recognition signals are the same, but they are similar enough to direct the investigator to the beginning of a coding region, which is to say to the *ATG* that follows a characteristic ribosome-recognition signal. The *ATG* also establishes the "reading frame" for the coding region. For example, in a sequence including *CGATGCACT* the codons could be *CGA, TGC, ACT* or *...C, GAT, GCA, CT...* or *...CG, ATG, CAC, T....* If the *ATG* is a start signal, the correct reading frame for the coding region must be the last of the three possible frames. A coding region ends with one of three codons that do not specify a protein and therefore serve as termination codons: *TAA, TGA* or *TAG* in the correct reading frame.



ACTUAL AUTORADIOGRAPH of a gel is interpreted here. A sequence extending from 21 to 73 nucleotides from the end of the primer can be determined by reading the band pattern. Bands representing successively larger products (*farther up the gel*) are closer together because the distance of migration is inversely proportional to the logarithm of the chain length.

It had been thought that the coding regions of DNA were clearly separated by noncoding regions that function as control signals. This assumption was substantiated when Walter Fiers and his colleagues at the University of Ghent recently established the nucleotide sequence of the small RNA bacteriophage *MS2*, whose genome is a single-strand RNA molecule about 3,300 nucleotides long that codes for just three proteins. The sequence analysis showed that the three genes are separated by two noncoding regions, respectively 26 and 36 nucleotides long.

The ϕX174 virus is very different. Some of the coding regions are indeed separated by untranslated regions. In three instances, however, two coding regions merge: the termination codon of one gene overlaps the initiation codon of the next gene, leaving no space for an untranslated region. This lapped joint is formed in one of two ways. Either a *TAA* termination codon overlaps the *ATG* start signal of the next gene or an *ATG* comes just before and overlaps the *TGA* termination codon of the preceding gene; in both cases part of the sequence thus has a dual function. As a result of this kind of overlap genes *A, C, D* and *J* form a continuous coding region.

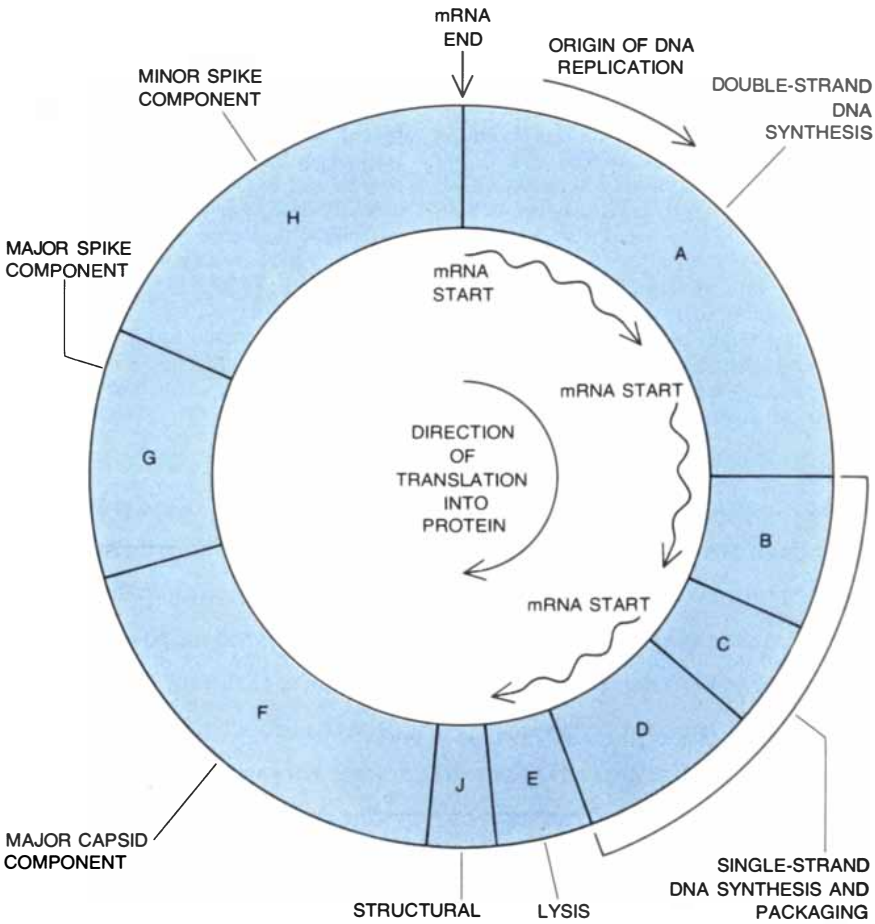
Even more surprising, in two instances a small gene is completely overlapped by a larger one, that is, the same stretch of DNA carries the information for producing two proteins with an entirely different amino acid sequence. This remarkable achievement of genetic-coding economy is accounted for by

an elegant trick: the overlapping genes are encoded in different reading frames.

The gene-within-a-gene phenomenon was discovered by looking for gene *E*. Genetic-mapping experiments had determined that in a certain portion of the ϕX174 genome the order of the genes was *D, E, J*. The complete amino acid sequence had been determined for both the *D* and *J* proteins. When the amino acid sequences were related to the newly available nucleotide sequence, it was found that the coding regions of the *D* and the *J* gene are contiguous. In fact, as was mentioned above, the termination codon of gene *D* overlaps the initiation codon of gene *J*. Apparently there was no room for *E*—which had been mapped between *D* and *J*.

In searching for gene *E* it was not possible, as it was in the case of genes *D* and *J*, to match amino acid and nucleotide sequences, because the virus produces such small amounts of the *E* protein that it has not been feasible to isolate the protein and determine its amino acid sequence. Instead a more indirect method was used: a search was carried out for a gene-*E* reading frame within the *D, J* sequence. The search was made with a ϕX174 mutant that does not make the *E* protein. The reason is not that these mutant viruses have lost the stretch of DNA that codes for the *E* protein but that a single-nucleotide mutation converts a codon for an amino acid into one of the three termination codons, and translation of the RNA transcribed from the mutated gene ends prematurely when ribosomes encounter the anomalous termination codon. There are bacterial strains, however, whose translating machinery contains a "suppressor" mutation that overcomes this premature termination codon by in effect reading the incorrect codon incorrectly, enabling the mutant virus to produce *E* protein and multiply.

The gene-*E* virus mutant was grown in one of these suppressor bacterial strains and the viral DNA was sequenced. By comparing the sequence with the normal viral sequence the nucleotide change that is responsible for the gene-*E* termination could be identified: in the mutant DNA a *G* is changed to an *A*. The change is in the region of the DNA that codes for the *D* protein. It was observed that in a particular reading frame the change would convert a codon for the amino acid tryptophan, *TGG*, into the termination codon *TAG*, thereby explaining the premature termination. The reading frame of this codon change was not the same as the clearly established gene-*D* reading frame, however, but one nucleotide to the right of it. This new reading frame accounts for the *E* gene. (It is interesting to note that in the *D*-gene reading frame the change from *G* to *A* converts the codon *CTG* into *CTA*. Both codons happen to code for leucine, so that the *E*-gene mutation does not affect the *D* gene's protein product.)



MAP OF GENOME, or complete genetic message, of ϕX174 had been determined by standard techniques before the sequence was established. Map showed relative length and position of each of the virus's nine genes (*A–J*). Function of each gene was known in a general way: they serve as enzymes for replication or packing of DNA or as structural components. Location was known of three start signals for messenger-RNA transcription and a termination signal, and of origin of DNA replication. Transcription, translation and replication go clockwise.

It was not hard to locate the beginning of the *E* gene by searching the DNA sequence for a characteristic ribosome-recognition sequence followed by an *ATG* initiation codon in the correct reading frame. The end of the *E* gene had to be the first termination codon encountered in the same reading frame. The location of the start and stop signals made it clear that gene *E* is contained entirely within gene *D*. Given the nucleotide sequence and the *E* reading frame, the amino acid sequence of the *E* protein has been predicted even though the protein has yet to be isolated.

The possibility that genes might overlap in different reading frames had been discounted on theoretical grounds because it imposes a severe evolutionary constraint. If a stretch of DNA can be read in two different frames, a favorable mutation in one frame may well be unfavorable in the other frame; for example, it may introduce a premature termination codon. The two overlapping genes have to evolve in parallel, the only permissible nucleotide changes being changes that allow both reading frames still to code for a viable protein.

In the case of the *D* and *E* genes it is possible to speculate as to the order in which the two genes evolved. The reasoning begins with the finding that the viral DNA has a large proportion of *T* nucleotides, the extra *T*'s appearing most frequently in the third position of the codons. This high incidence of *T*'s in the third position is specifically true in the *D*-gene sequence. Since the *E*-gene reading frame is displaced one position to the right of the *D* frame, the *E*-gene codons must have a high level of *T*'s in their second position. Such a codon tends to specify one of the hydrophobic, or water-repelling, amino acids. The *E* gene must therefore have an overall structure somewhat similar to that of a detergent molecule. Such a structure might account for the *E* protein's function: disrupting the host cell's outer membrane. (Bacterial cells are broken open in the laboratory by detergent so-

lutions, and some secretory proteins in higher organisms that need to move through membranes also have a detergent structure.) We therefore speculate that the *E* gene arose after the *D* gene as the result of a mutational event that turned out to be highly adaptive: it allowed the virus to capitalize on its high content of *T*'s to break down the cells it infected.

Such an explanation of the origin of the *E* gene seems to imply that gene overlap might be the result of a rather special event exploiting a particular feature of the viral DNA. There is a second pair of overlapping genes in ϕX174, however, and to them this kind of rationalization does not seem to apply.

The apparent inconsistency between the small size of the viral DNA and the fact that it codes for nine proteins had suggested that there might be more than one pair of overlapping genes, and specifically that gene *B* might be contained entirely within gene *A*. This was investigated in much the same way as the overlap of genes *D* and *E*. A ribosome-recognition sequence within gene *A* led to an *ATG* that seemed to be the initiating codon of gene *B*. This *ATG* was in a reading frame two nucleotides to the right (or one nucleotide to the left) of gene *A*'s frame. By sequencing chain-termination mutants in both genes and finding the mutations that created termination codons, the two reading frames were confirmed and the fact that gene *B* is contained entirely within gene *A* was established.

The overlap of genes *A* and *B* is surprising. The proteins encoded by these genes appear to function as enzymes, whose activity depends on their precise structure, and one would expect that evolutionary constraints on changes in their amino acid sequence would be stronger than the constraints on

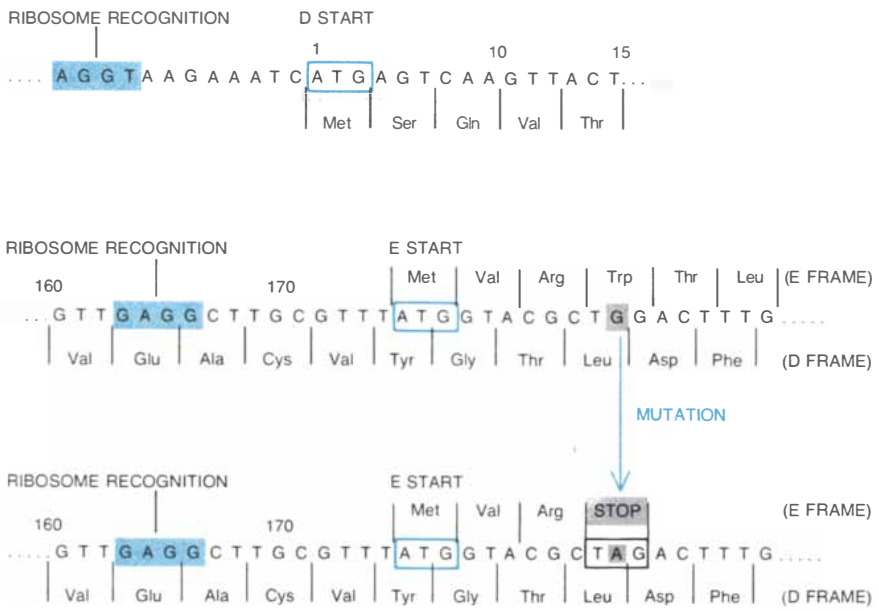
COMPLETE SEQUENCE of the 5,375 nucleotides of the ϕX174 viral DNA is listed on these two pages. Although the strand of DNA is a closed loop, here the sequence reads from left to right in each line, first down this page and then down the opposite page; the sequence

changes in a detergentlike molecule such as the *E* protein. It is possible, however, to form some idea of how these two genes evolved by studying the distribution of *T*'s in the third position of codons. In the region of gene *A* that precedes the overlap there is in the *A*-gene reading frame the usual high incidence of codons ending in *T*, but within the overlap region that high incidence becomes a feature of the *B* reading frame. Assuming that the high level of third-position *T*'s is a basic characteristic of ϕX174 DNA, this suggests that the *A* and *B* genes were once distinct, with *A* ending before the beginning of *B*, and that a mutation arose in what had been the termination codon for the *A* gene, allowing it to read on into the *B* gene.

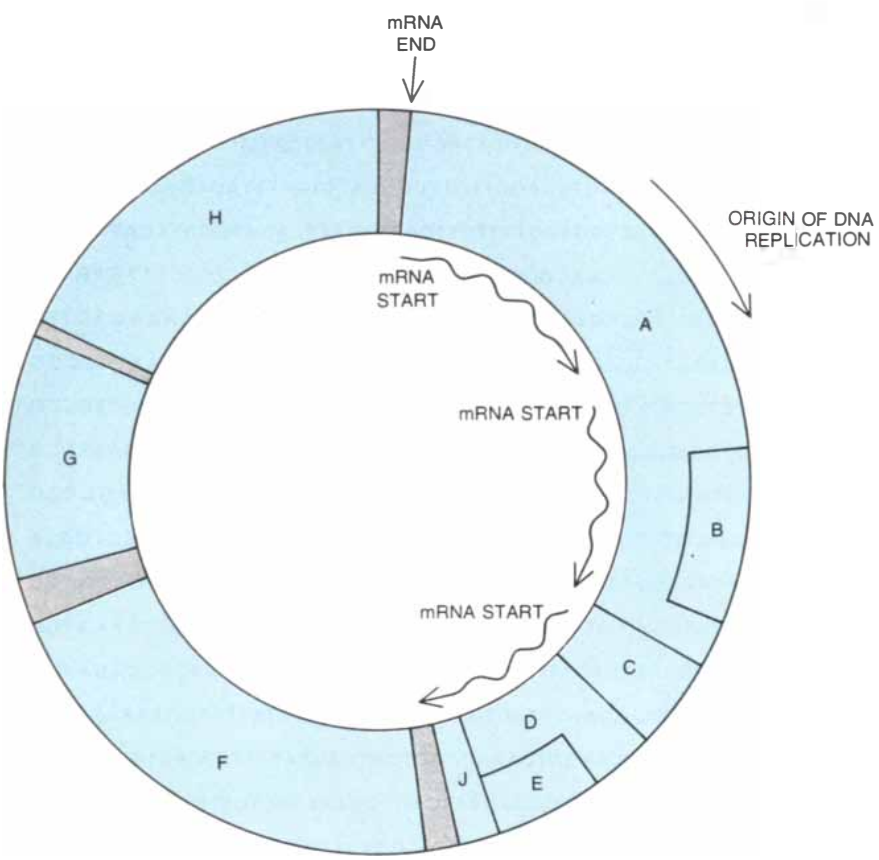
Whether overlapping will turn out to be a general phenomenon in the DNA's of various kinds of organisms is not yet clear. It can be argued that ϕX174 is a very special case, since the total amount of its DNA is severely limited by the physical size of the capsid in which the DNA is packaged. Presumably the only way such a physically constrained genome can evolve extra, advantageous functions is to develop overlapping genes. In the cells of higher organisms, including mammals, the problem seems to be quite the reverse. Instead of there being not enough DNA for the required functions there appears to be a vast excess of DNA beyond what is required for coding and control functions. This does not necessarily mean, however, that overlapping genes will not be found in higher organisms.

The combination of the two pairs of overlapping genes and the three instances in which termination and initiation codons overlap means that from the start of the *A* gene all the way to the end of the *J* gene, a region encompassing more than 40 percent of the viral genome, there is no untranslated DNA.

begins with the noncoding region before the *A* gene and ends with the termination codon of the *H* gene. The ribosome-recognition signals are shown in the shaded colored boxes, the initiation codons in open colored boxes and the termination codons in open black boxes.



OVERLAP OF *D* AND *E* GENES is illustrated. Translation of the *D* gene is triggered by the *ATG* initiation codon following a ribosome-recognition signal (*top line*). Each three-nucleotide codon is labeled with the amino acid it specifies. In the midst of the *D*-gene sequence (*middle line*) there is another ribosome-recognition signal and then an *ATG* initiation codon in a new reading frame, which begins the *E* gene. The *E*-gene frame was identified by the discovery of the mutation of a *G* to an *A* in a mutant virus (*bottom line*), changing the codon for the amino acid tryptophan (*Trp*) to a premature termination codon, *TAG*, that stops *E*-gene translation. The reading frame for *D* gene is not affected, since both *CTG* and *CTA* specify leucine (*Leu*).



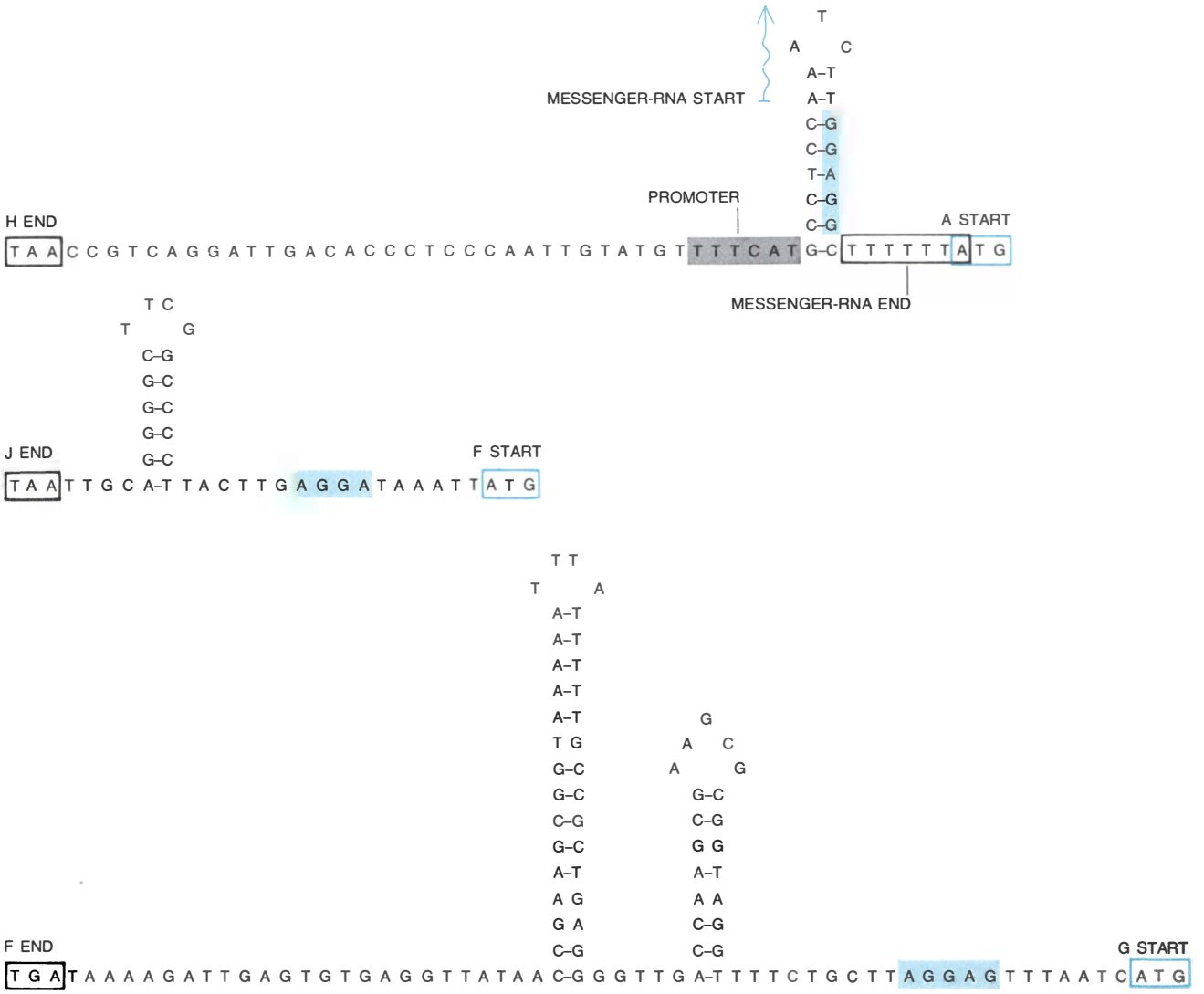
GENETIC MAP of the ϕX174 genome has now been refined on the basis of the nucleotide sequence, with the noncoding regions indicated (*gray*). From the start of the *A* gene to the end of gene *J* there is no untranslated DNA; gene *B* is within gene *A* and gene *E* is within gene *D*.

The situation is more typical, however, on the other side of the genetic map, where the structural genes are situated. These genes are separated by distinct untranslated regions: spaces respectively 39, 111, 11 and 66 nucleotides long between genes *J* and *F*, *F* and *G*, *G* and *H*, and *H* and *A*. Each of the three larger spaces contains a self-complementary sequence: a sequence such that the single strand of DNA can form a "hairpin" structure by looping out and base-pairing to itself [*see top illustration on opposite page*]. Because such hairpin loops are not nearly as common in the coding regions of ϕX174 as they are in these noncoding spaces, we think they may have some regulatory function.

Particular control functions have in fact been ascribed to several sequences in these untranslated spaces. The region preceding the start of the *A* gene contains near its hairpin loop one of three sites on the ϕX174 DNA where the enzyme RNA polymerase recognizes the DNA and initiates the process of transcription into RNA. Several such sites, called promoters, have been identified in *E. coli* and its viruses and have been found to have a characteristic sequence, to which the gene-*A* promoter conforms. The *H*-*A* space also contains the major site in ϕX174 where the process of transcription terminates, the signal for which appears to be the sequence *TTTTTTA* preceded by a hairpin loop. (How can termination follow so quickly on initiation? It has been speculated on the basis of work done in other systems that it is hairpin plus *TTTTTTA* that halts transcription; an RNA just initiated at the peak of the hairpin would therefore not be terminated.) There is a weaker transcription-termination site between the *J* and *F* genes, where the hairpin structure alone may constitute the signal. As for the longest untranslated region, between the *F* and *G* genes, no particular function has yet been ascribed to it.

Not all the control signals in ϕX174 are in untranslated regions, however. Some nucleotide sequences serve as both a coding region and a recognition signal. For example, transcription promoters preceding genes *B* and *D* are respectively situated within the coding region for genes *A* and *C;* they have sequences characteristic of RNA-polymerase recognition sites even though their stretch of nucleotides codes for amino acids.

The most striking example of economy in genetic coding is to be found in the case of the ribosome-recognition sequence for the *J* gene. The sequence, *AAGGAG*, not only serves as a recognition site for ribosomes but also codes for the *D* protein in one reading frame and for the *E* protein in another frame [*see bottom illustration on opposite page*].

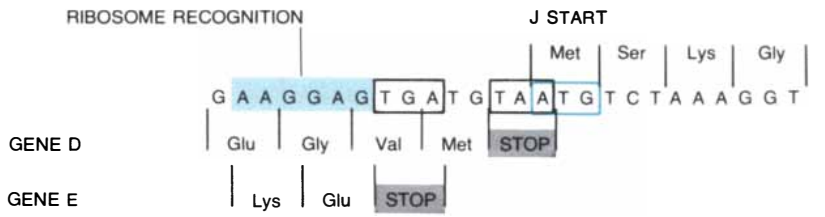


THREE UNTRANSLATED REGIONS, between genes *H* and *A*, *J* and *F*, and *F* and *G*, contain self-complementary sequences that form "hairpin" loops by base-pairing. The region before the *A* gene has a promoter: a recognition site for the enzyme RNA polymerase, which mediates transcription of the DNA into messenger RNA; the transcription begins a few nucleotides later. The hairpin loops in *H–A* and *J–F* spaces are thought to be involved in the termination of transcription. The function of the *F–G* space has not been determined.

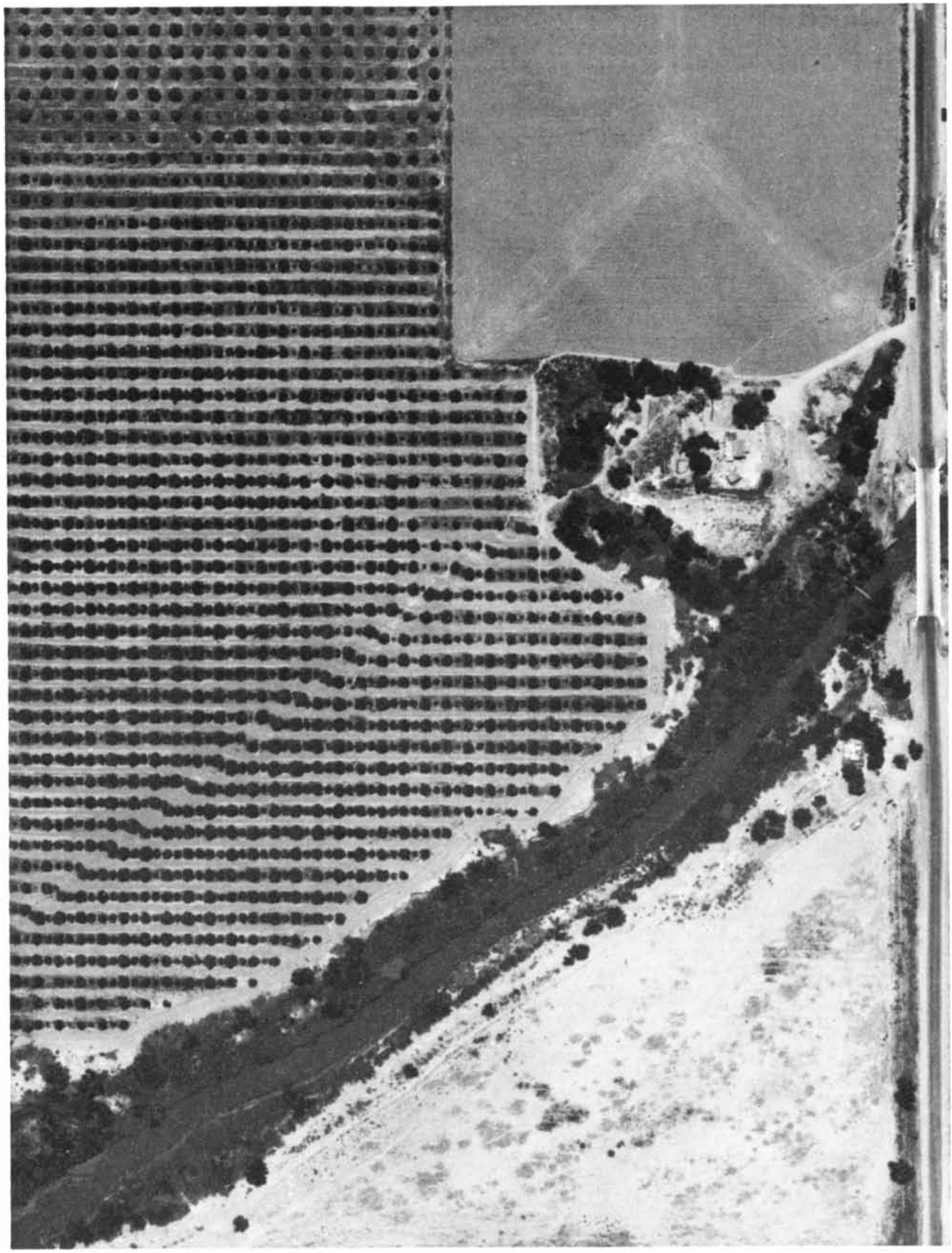
The ribosome-binding sites for genes *B*, *C*, *D* and *E* are also within coding regions, albeit regions that code for only one protein. In all these cases the DNA sequence must have had to evolve in such a way that none of its multiple functions was impaired.

The first fruits of the complete sequencing of ϕX174 have been these unexpected glimpses into the organization of the viral genome. The earlier model of DNA as having coding regions interspersed with largely regulatory spacer regions has had to be revised in view of the discovery of overlapping genes and the realization that a single stretch of DNA can have both a coding function and a control function. Moreover, such overlapping suggests that the evolutionary constraints on at least some nucleotide sequences must be much more stringent than has been believed.

Now that DNA can be sequenced readily and rapidly we can expect that in the next few years the precise composition of many DNA's will be established. In particular, new techniques for recombining the DNA of different species have already made it possible to determine the sequence of some regions of the DNA in systems other than viruses, and thus to begin to understand the complex mechanisms by which genes are expressed in higher organisms.



REMARKABLE ECONOMY of DNA coding is illustrated at the end of overlapping genes *D* and *E* and at the beginning of gene *J*. A nucleotide sequence serves three functions: the *AAGGAG* is the ribosome-recognition site for the start of gene *J*; the same nucleotides, in two different reading frames, code for amino acids in genes *D* and *E*. And the termination codon of gene *D* overlaps the initiation codon of gene *J*, which is read in a different frame.



HORIZONTAL DISPLACEMENT of the ground during an earthquake in the Imperial Valley of California disrupted the regular pattern of trees in citrus groves. In this aerial photograph of an orchard seven miles east of Calexico, made shortly after the earthquake in 1940, the path of the San Andreas fault can be clearly traced diagonally across the groves west of the Alamo River. North is to the right.

The Motion of the Ground in Earthquakes

The slippage along a fault that produces an earthquake radiates seismic waves. Exactly how these waves shake the ground bears on the design of buildings and other structures in earthquake zones

by David M. Boore

Last year half a million people were killed by an earthquake that devastated the Chinese industrial city of Tangshan. In the western U.S. over the years earthquakes have caused considerable damage, although the number of fatalities has been relatively small. The low casualty rate has been partly due to the fact that many of the major earthquakes occurred either in sparsely populated areas or fortuitously quite early in the morning, when most large office and public buildings are almost empty. Over the past few decades, however, many earthquake-prone regions of the western U.S. have become further urbanized. In them more large buildings and facilities such as dams have been constructed or are being planned. If such structures were to fail during a future earthquake, large numbers of people could be killed or injured.

Today the attention of many seismologists is being focused on ways to reduce the hazards of earthquakes by learning how to predict their consequences. To many people the term earthquake prediction probably suggests determining the time, place and magnitude of future earthquakes. Equally important is determining which of many ways the ground is likely to shake during the earthquake, how strong the shaking will be and how long it will last. Knowledge of the ground motion that can be expected during an earthquake can make it possible to design structures that do not need unnecessary and uneconomic levels of strength in order to survive being shaken.

In order to predict both the occurrence of an earthquake and the ground motion it will generate it is essential to understand the characteristics of the earthquake source. So far most of our understanding of earthquake sources has come from measurements made during actual earthquakes at seismological stations some distance from the source. Such measurements yield information about certain average properties of the earthquake source, for example the dimensions of the original disturbance and the overall movement involved in it. Average properties are useful in elucidating how seismic energy is released and how it is transmitted over large areas; they have also been invaluable in probing the structure and nature of the earth's interior and in assessing the likelihood of large earthquakes in certain regions. Such average properties, however, yield little information about the details of the ground shaking in areas immediately surrounding the earthquake source. It is this kind of specific information structural engineers require. For that reason a number of seismologists are now beginning to investigate the details of earthquake sources. This important subject, which might be called strong-motion seismology, is still in its infancy but should grow rapidly.

Historically our understanding of the cause of earthquakes is relatively new. By the middle of the 19th century it had been observed that the damage caused by many earthquakes was concentrated in a narrow zone, which suggested that earthquakes had a localized source. It was not until the San Francisco earthquake of 1906, however, that it was recognized that earthquakes were caused by slippage along a fault in the earth's crust. In a classic study conducted shortly after the earthquake Harry F. Reid of Johns Hopkins University discovered that for several hundred kilometers along the San Andreas fault fences and roads crossing the fault had been displaced by as much as six meters. Moreover, precise geodetic surveys conducted before and after the earthquake demonstrated that the rocks parallel to the fault had been strained and sheared. On the basis of such observations Reid proposed the elastic-rebound theory of earthquakes.

According to the elastic-rebound theory, rocks are elastic, and mechanical energy can be stored in them just as it is stored in a compressed spring. When the two blocks forming the opposite sides of the fault move by a small amount, the motion elastically strains the rocks near the fault. When the stress becomes larger than the frictional strength of the fault, the frictional bond fails at its weakest point. That point of initial rupture, called the hypocenter, may be near the surface or deep below it.

From the hypocenter the rupture rapidly propagates along the surface of the fault, causing the rocks on opposite sides of the fault to begin to slip past each other. A portion of the frictional stress the rocks had exerted on each other before the rupture is suddenly and violently released; the rocks along the fault rebound, or spring back, to an equilibrium position in a matter of seconds. The elastic energy stored in the rocks is released as heat generated by friction and as seismic waves. The seismic waves radiate from the hypocenter in all directions, producing the earthquake. The point on the surface of the earth above the hypocenter is the epicenter of the earthquake.

In some cases the rocks rebound not in a period of seconds but over an interval of minutes, days or even years. The seismic energy radiated at any one time is then quite small. This slow process is known as aseismic slip or creep. Why the seismic energy is released violently in some cases and not in others is not well understood.

Although the physical details of the elastic-rebound theory are still uncertain, the conceptual model of the faulting process fits well with the current hypotheses of plate tectonics. Most earthquakes are generated in zones where the huge plates of the lithosphere, which

make up the outer layer of the earth's surface, are shearing past each other.

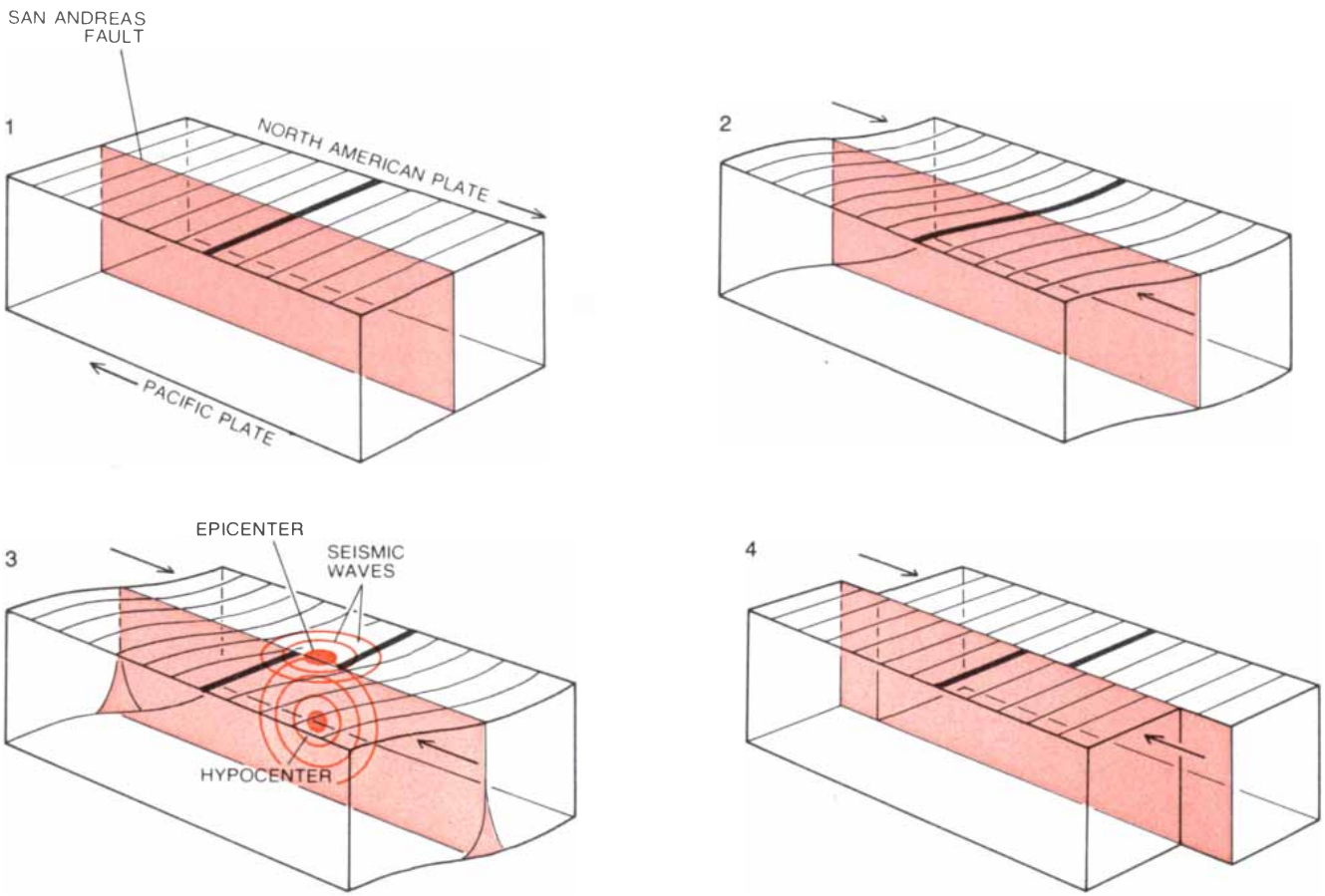
The concept of slip along a fault is at the heart of virtually all studies of earthquake sources. Indeed, the concept developed largely from investigations of earthquakes along the San Andreas fault. The San Andreas is a very long fault but not a deep one; earthquakes caused by its slippage are confined to about the upper 15 kilometers of the crust. Yet the study of this one shallow fault has led to a model that successfully explains the deformation of the ground and the radiation of seismic waves from all types of seismic sources, ranging from the shallowest slips to ruptures as deep as 700 kilometers along the advancing edge of a plate plunging below another plate.

The way the ground is deformed and the nature of the seismic waves that radiate during the earthquake provide basic information about the earthquake source: its dimensions, its shape and its orientation. The seismic waves have a wide range of period and amplitude. When a fault slips, the rupture process itself generally lasts between a fraction of a second (for a minor earthquake) and five minutes (for a major one). The waves generated by the fault's slippage can have periods ranging from essentially infinity down to less than a tenth of a second. The seismic waves with the longest period correspond to the quasi-permanent deformation of the ground around the fault. The waves with the shortest period actually fall into the low audible range. The waves with periods of about an hour have a frequency that coincides with the resonance frequency of the earth, and they cause the entire planet to ring like a giant bell.

The amplitude of the seismic waves can range from micrometers (millionths of a meter) to tens of meters. The amount by which the seismic waves deform the ground decreases with distance from the earthquake. In the great Chilean earthquake of 1960, for example, the total displacement of some points immediately adjacent to the fault ranged up to 20 meters. At Los Angeles, a quarter of the way around the world, the maximum displacement of the ground was about two millimeters.

Since seismic waves span such a broad spectrum of period and amplitude, many different kinds of instruments and experimental techniques are needed to capture all the information radiated by an earthquake source. Repeated geodetic surveys of the earth's surface can monitor deformations of the ground created by seismic waves with periods ranging from days to years. A variety of different seismographs have been designed to record seismic waves with periods ranging from an hour to a hundredth of a second. Some instruments are so sensitive that they can detect motions as minute as one micrometer, which they magnify tens of thousands of times in order to record them on paper. Other instruments are so rugged that they can withstand the jarring accelerations of the most violent earthquakes.

The record produced by the seismograph—a seismogram—holds a great deal of information; even with the aid of a computer, however, deciphering that information is neither simple nor straightforward. The waves recorded on a seismogram after passing through the earth can be thought of as violin music

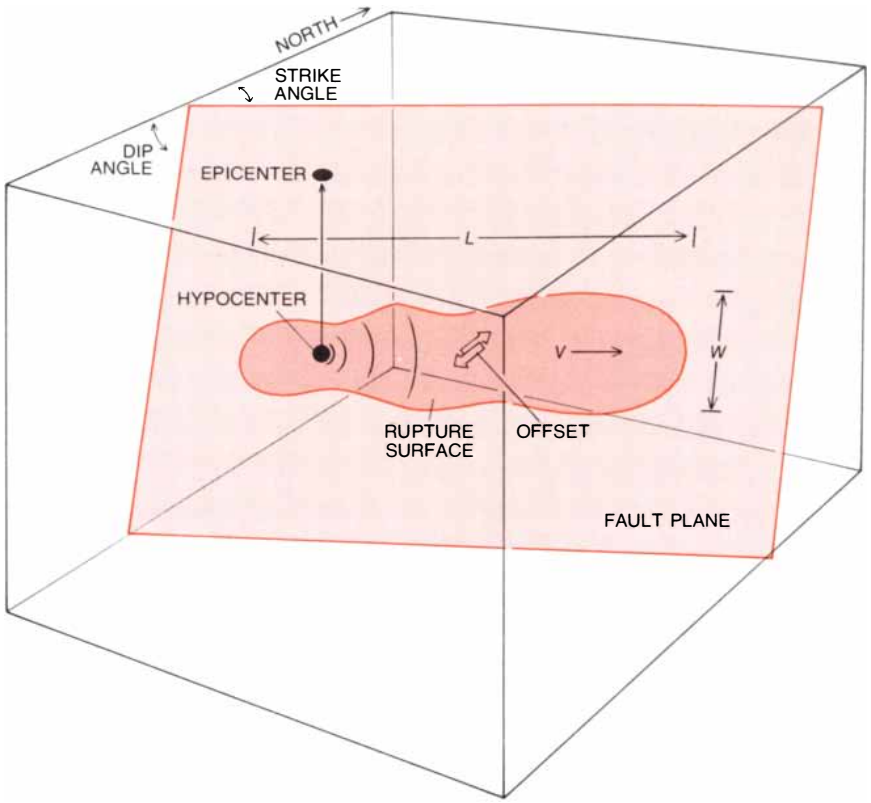


ELASTIC-REBOUND MODEL OF EARTHQUAKES assumes that two moving blocks of the earth's crust, each of which is part of a different tectonic plate in the earth's lithosphere, meet at a fault (*1*). Friction between the plates along the surface of the fault at first keeps them from slipping past each other, but the material around the fault is deformed by the stress (*2*). The deformation builds up until the frictional lock is ruptured at its weakest point, usually well below the surface (*3*). The rupture spreads out from that point, the hypocenter, radiating seismic waves as it does so. The point vertically above the hypocenter, where the seismic waves first reach the surface, is the epicenter of the earthquake. As the rupture spreads along the surface of the fault the blocks slip past each other, usually in a few seconds, coming to rest in a new equilibrium position (*4*). The stress around fault is relieved and ground rebounds to earlier state.

recorded on magnetic tape after first being transmitted over a telephone line that distorts the music. In this analogy the violin corresponds to the seismic source, the telephone line corresponds to the inhomogeneous elastic earth that distorts the signal passing through it and the tape recorder corresponds to the seismograph (which further distorts the signal as it is being recorded).

It is easy to correct for the distortion of the tape recorder. The challenge lies in trying to deduce something about the nature of the violin on the basis of the distorted sound received at the end of the telephone line. If one assumes that the telephone line is free of distortion, one might then reasonably conclude that a violin intrinsically produces a harsh sound. On the other hand, if one knows how a violin sounds when it is heard "live," one could use that knowledge to discover how the telephone line filters and distorts the music.

A similar problem faces the seismologist examining the record of an earthquake. In seismology the earth filter that distorts the seismic waves is complex because the internal structure of the earth is complex. As a result of decades of geological research, however, we now know much more about the earth's internal structure and how it distorts a seismic signal than we know about the earthquake source. Because the earthquake source is usually deep underground its seismic radiation cannot be "heard" firsthand. Seismologists must deduce the nature of the source by the indirect procedure of constructing a theoretical model of it, calculating the pattern of seismic radiation produced by the model, estimating how the seismic signal would be distorted as it propagated through the earth to the seismograph and comparing the synthetic seismogram with the actual seismogram recorded. By repeating the procedure several times with better information it is possible to refine the description of the earthquake source. Current models thus constructed attempt to describe the complex rupture process with relatively few parameters.



IDEALIZED MODEL OF EARTHQUAKE SOURCE suffices to describe most earthquakes with about a dozen variables. In the model the rupture begins at the hypocenter h kilometers below the surface, spreads across a fault plane at a velocity V and finally stops after growing into a region with an average length L and an average width W. The orientation of the fault plane is specified by its strike angle and dip angle. The slip between the two fault surfaces (*large arrows*) can have any orientation in the plane. On the average the slip requires τ seconds to reach its final offset. All these parameters are determined from recordings of the seismic waves.

At the simplest level a model specifies the location of the hypocenter and the magnitude of the earthquake. At a more complex level the model includes the orientation of the fault surface underground and the direction of slip across the surface. The model can be made even more realistic by adding the dimensions of the entire area that ruptured, the average amount of slip across that area and the average length of time required for a point on the fault surface to be offset by the maximum amount. Since friction opposes the motion of the two sides of the fault past each other, it is believed that once a fault begins to slip, its direction cannot reverse. Such models are quite successful in predicting the different types of seismic waves actually observed, particularly in predicting seismic waves with wavelengths at least as long as the dimensions of the fault.

The location of an earthquake can be determined by a procedure akin to triangulation, taking advantage of the fact that different types of seismic waves travel at different speeds. Seismic waves are of two general types: P waves and S waves. The P waves are longitudinal compression waves that travel through the deep interior of the earth, even propagating through the lower mantle and the liquid core. The S waves are transverse shear waves that travel through the solid portions of the earth.

P waves travel significantly faster than S waves. At a location close to the earthquake source the two types of waves will arrive fairly close together, but at one farther away the S wave will lag significantly behind the P wave. By observing the difference in arrival time between the two types of waves at any one station it is possible to calculate the distance of the earthquake from the station. Such a calculation from a single station does not determine the direction of the earthquake, but when observations from three or more stations are combined, the precise location of the earthquake can be determined. If there are enough data, it is also possible to locate earthquakes from the P waves alone. In fact, this is the technique used by the National Earthquake Information Service in Golden, Colo., which collates earthquake data recorded all over the world and issues information about the position of an earthquake as soon as possible after each event.

The most widely recognized measure of the strength of an earthquake is the scale of magnitudes developed in the 1930's and 1940's by Charles F. Richter and Beno Gutenberg of the California Institute of Technology. The scale is based on the notion that ideally the magnitude determined should be an absolute measure of the energy released by the earthquake itself and should not be affected by the location of the seismographic station or the particular seismograph employed. The Richter method for determining the magnitude of an earthquake is quite simple. First, the seismologist measures the amplitude of the ground motion recorded in a certain

specified part of the train of seismic waves. Second, he divides the recorded amplitude of the ground motion by the magnification of the particular seismograph to estimate the true ground motion at the seismographic station. Third, he calculates the common logarithm (the logarithm to base 10) of that ground motion. Fourth, he applies certain empirical corrections to that number to compensate both for the attenuation of the ground motion as it spreads out from the earthquake source and for the degree to which the response of the particular seismograph is influenced by local geological conditions.

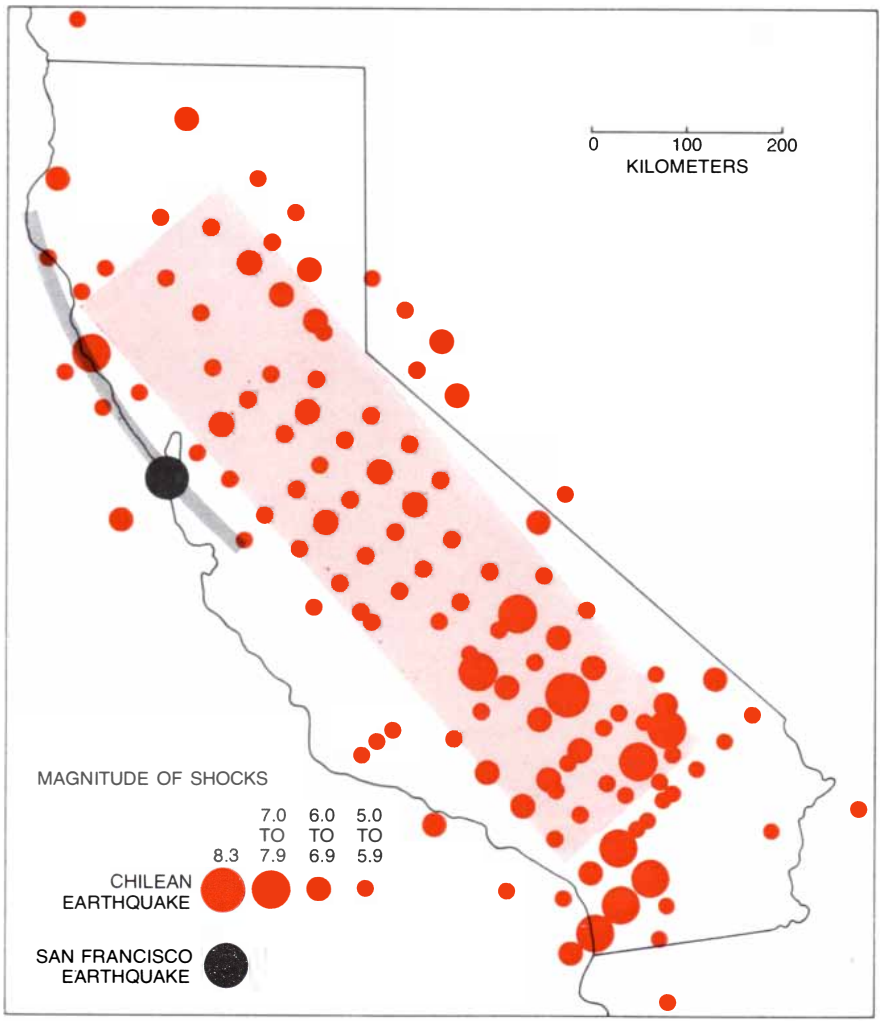
The empirical corrections are applied so that for any given earthquake the same magnitude should be determined at all seismographic stations. In practice the magnitudes differ from one station to another, and an average magnitude is calculated from all of them. On the Richter magnitude scale larger numbers correspond to larger events. Since the scale is based on the common logarithm of the corrected ground displacement, each increase of one magnitude unit implies an increase of a factor of 10 in the amplitude of the ground motion. The magnitude scale is open-ended, and negative magnitudes have been measured.

Actually there are several magnitude scales in common use, each based on a different part of the seismic wave train. One is the scale of body-wave magnitude, measured from the *P* waves that travel through the body of the earth and reach the seismograph before any other waves. By convention, *P* waves with a period near one second are used in the magnitude determination. Another scale is the scale of surface-wave magnitude, measured from the dispersed waves that travel over the surface of the earth and reach the seismograph somewhat later. The surface waves employed have periods of 20 seconds. The two magnitude scales are cross-calibrated so that on the average both will yield the same magnitude when the earthquake being recorded has a magnitude of 6.75. By measuring the two magnitudes for a particular earthquake one obtains an estimate of the overall amount of seismic energy radiated in two quite different regions of the seismic spectrum. For a large earthquake the surface-wave magnitude is generally greater than the body-wave magnitude. This fact implies that the excitation in the long-period part of the spectrum increases faster with earthquake size than the excitation in the short-period part of the spectrum.

After the location and the magnitude of an earthquake have been determined from seismograms, the kind of information that can next be most readily obtained is the geometry of the earthquake source: the orientation of the fault in the earth, the dimensions of the portion of the fault plane that has slipped and the direction of the slip in the fault plane. Just as an array of radar antennas has a defined pattern of radiation, with large amounts of energy being beamed in some directions and small amounts in other directions, so also does an earthquake source have a defined pattern in which it radiates seismic energy. The radiation pattern not only determines the amplitude of the seismic signal in different directions but also determines how the seismic waves are polarized.

The radiation pattern can be understood by means of a simple experiment with a cube of foam rubber. Slit the top of the cube and push the two sides horizontally in opposite directions parallel to the slit. You will notice that the foam is compressed in two diametrically opposed quadrants and dilated in the other two quadrants. When a fault slips, the material around it is similarly compressed and dilated. The first waves emitted from an earthquake fault display the same distribution of compressions and dilations. The distribution of those waves on the surface thus reveals the orientation of the fault plane and the relative direction of the slip.

In the experiment with the cube of foam rubber, however, the quadrants of compression and dilation are clearly separated by two orthogonal lines; one line is the fault and the other line is perpendicular to the fault. Observations of the radiation pattern from an earthquake determine the orientation of two similar orthogonal planes, either one of which may be the earthquake fault. The ambiguity can be resolved if the orientation of the true fault plane is known

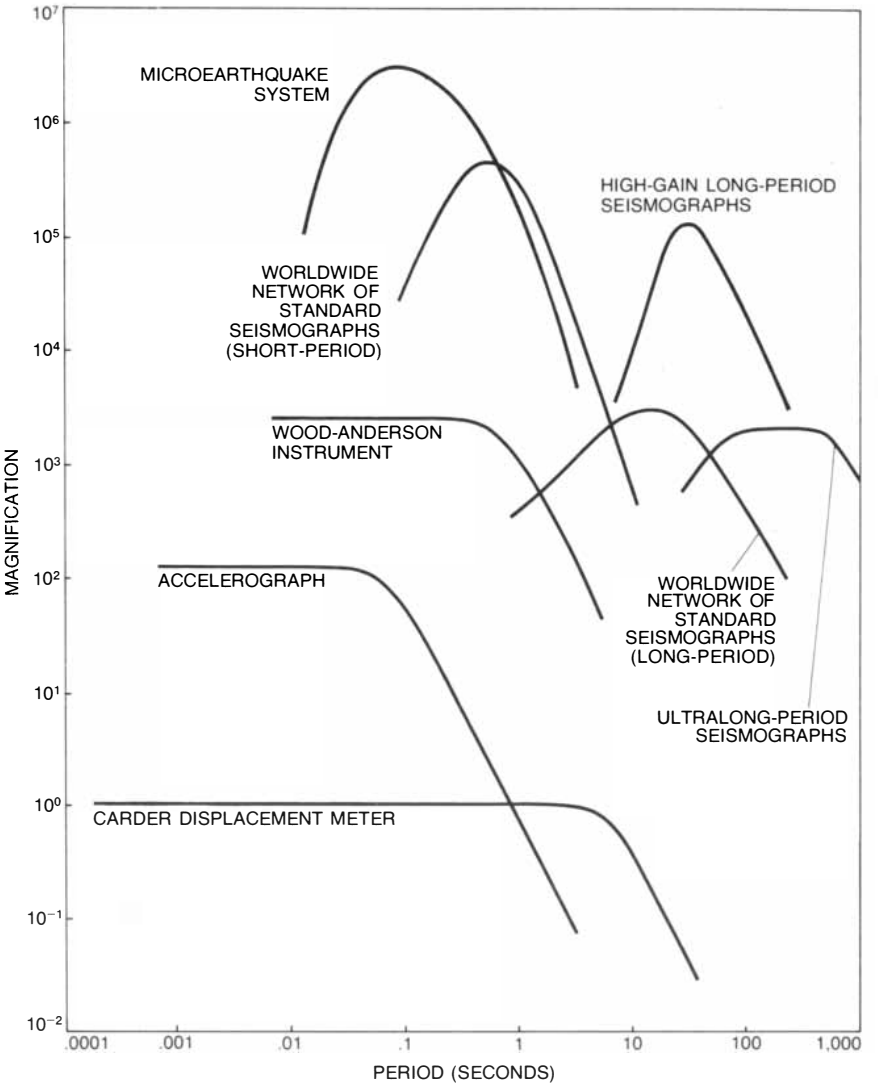


MAGNITUDE OF EARTHQUAKES is an inadequate measure of the actual size of large earthquakes. Both the San Francisco earthquake of 1906 and the Chilean earthquake of 1960 had a magnitude of 8.3. The area that ruptured in the San Francisco earthquake (*gray*), however, was approximately 15 kilometers deep and 400 kilometers long whereas the area that ruptured in the Chilean earthquake (*color*) extended to a depth equal to half the width of the state of California. The black dot represents the location of the epicenter of the San Francisco earthquake; the dots in color represent the locations of aftershocks of the Chilean earthquake with respect to its epicenter (*largest color dot*), superposed on the map of California for scale. The diameter of each dot represents the magnitude of each shock. Because earthquakes in California are caused by plates sliding past each other horizontally and not by plates subducting over each other as in Chile, no earthquake in California will be as great as earthquakes in Chile.

from the local geology. Alternatively, the orientation of the true fault plane can be determined from the pattern of aftershocks, smaller tremors that generally follow an earthquake, because the hypocenters of the aftershocks are usually scattered along the fault plane.

The information about the geometry of faults that has been amassed from earthquakes has been invaluable in developing the theory of plate tectonics. It has played a key role in identifying the faults between plates of the lithosphere and the relative motions of the plates. The seismic waves from an earthquake also yield information about the dimensions of the area that ruptured along the plane of the fault. The detail of the rupture area it is possible to resolve depends on the wavelength of the seismic radiation, just as in optics the wavelength of light limits the resolution of visual observation.

The area that ruptured in 1906, causing the magnitude-8.3 San Francisco earthquake, was 15 kilometers deep and 400 kilometers long; the area that ruptured in 1971, causing the magnitude-6.5 San Fernando earthquake in the Los Angeles area, was also 15 kilometers deep, but it was only 15 kilometers long. Seismic waves travel about four kilometers per second. The surface waves with a period of 20 seconds hence have a wavelength of some 80 kilometers. The 20-second waves might have provided a certain amount of detailed information about the source of the San Francisco earthquake, but with such waves the source of the smaller San Fernando earthquake would have appeared to be a point. By the same token, with seismic radiation having a period of several hundred seconds even the source of the San Francisco earthquake would have seemed to be a point.



RESPONSE OF SEISMOGRAPHS of different types has been tailored to monitor seismic waves over a broad spectrum of period and amplitude. The magnification of the instrument is the number of times the instrument amplifies the ground motion so that it can be recorded. The amplitude of the ground motion in centimeters is approximately equal to the inverse of the magnification. For the most sensitive instruments the magnification is limited by ambient vibrations of the ground produced by wind and surf. The microearthquake system records small earthquakes within about 100 kilometers of the instrument. The Wood-Anderson instrument records moderate earthquakes at distances of several hundred kilometers. Moderate-sized earthquakes occurring almost anywhere in the world can be recorded on the short-period and the long-period systems of the Worldwide Network of Standard Seismographs or on special instruments such as the ultralong-period seismograph or high-gain long-period seismograph. Carder displacement meter and accelerograph record strong shaking close to fault.

Clearly a fault is not a point source. As a rupture propagates over the surface of the fault the point from which the seismic radiation is being emitted moves and causes the seismic waves emitted from one portion of the fault to destructively interfere with the waves emitted from another portion. The shorter the period, the more important the destructive interference. The period at which the interference first becomes noticeable can be used to estimate the dimensions of the fault. For example, the period might be about six seconds for a fault with dimensions of 10 kilometers by 10 kilometers or 60 seconds for one with dimensions of 100 kilometers by 100 kilometers.

In the 20th century about 55 earthquakes have been observed with surface-wave magnitudes ranging between 8.0 and 8.7, and no earthquakes have been observed with a surface-wave magnitude greater than 8.7. Actually two earthquakes near the upper end of the magnitude range may have the same surface-wave or body-wave magnitude and yet radiate vastly different amounts of seismic energy. In other words, for large earthquakes the magnitude scale becomes saturated.

The reason for this saturation is easily understood. The largest earthquakes rupture faults hundreds of kilometers long. If a fault is very long, it takes more time for a wave emitted from the farther end of the fault to reach the seismograph than it does for a wave emitted from the nearer end of the fault. Since the wavelength of a surface wave can be much shorter than the length of a very long fault, the part of the wave train from which the earthquake's magnitude is measured will be emitted from only a fraction of the fault's area rather than from the entire fault. The result is that the strength of the earthquake appears to be less than it actually is, and the magnitude scale cannot accurately measure very large earthquakes.

A new measure of the strength of an earthquake, known as seismic moment, has recently come to the fore. Seismic moment is not as easy to measure as seismic magnitude, but it is a more physical measure of the size of an earthquake source. The seismic moment is determined by the Fourier analysis of seismic waves of such long period that

the details of the rupture are smoothed out and the entire fault appears to be a point source. (The periods at which the seismic moment is determined increase with the size of the fault.) If the fault is "viewed" by such long-period waves, the slip from the unruptured state to the ruptured one appears to be instantaneous. The actual pattern of the seismic radiation emitted by the instantaneous rupture is mathematically equivalent to the theoretical pattern of radiation emitted by a model consisting of two hypothetical torque couples embedded in an unruptured elastic medium.

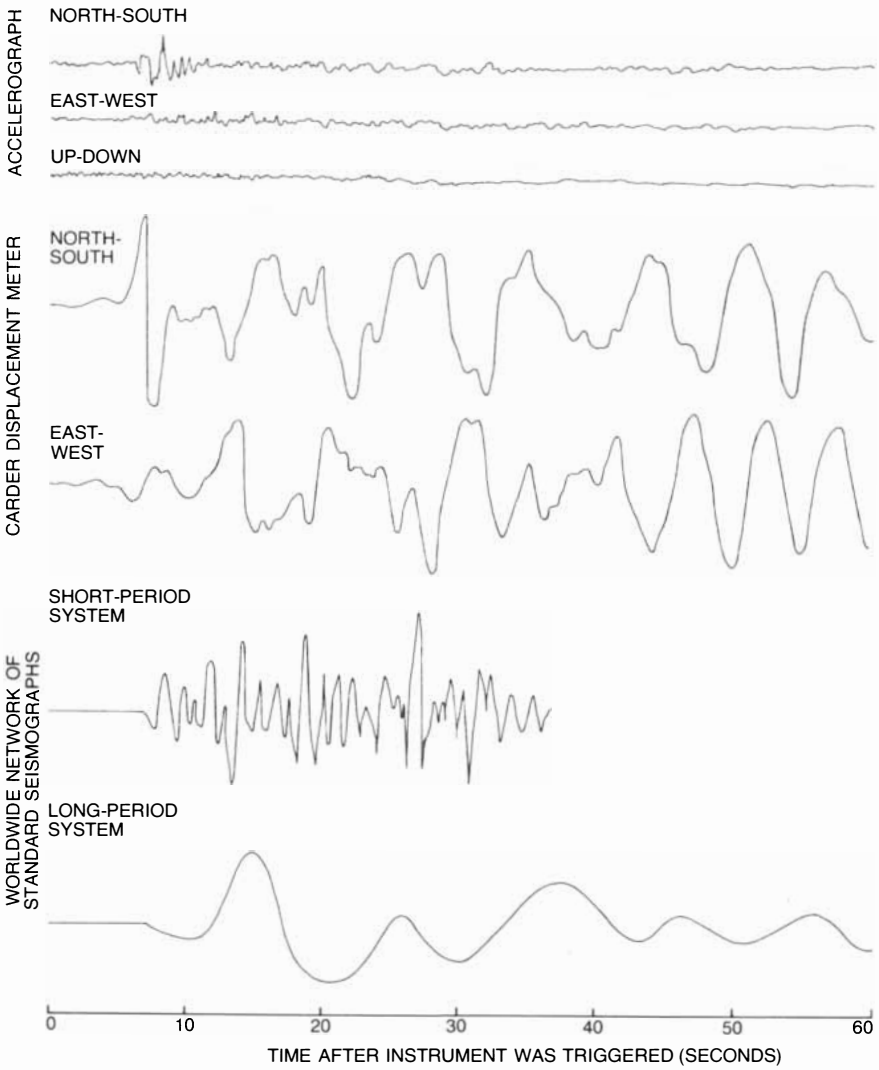
Each of the two torque couples can be visualized as a pair of small spheres, with a thin wire attached to each sphere. The wires are pulled with equal force in such a way that one pair of spheres rotates in one direction while the other pair rotates in the opposite direction. The magnitude of the rotary force—the torque—exerted by each pair of spheres on the elastic medium is the moment. Since the two torque couples rotate in opposite directions, however, no net torque is applied to the medium. The two torques nonetheless deform the medium, radiating elastic waves in a characteristic pattern: a pattern identical with the one in which an earthquake source radiates seismic waves. From this model the moment of the seismic radiation emitted by earthquakes can be calculated. The model has been named the double-couple source model.

The seismic moment measures the seismic energy emitted from the entire fault and not from just a portion of the fault, so that it is a fundamental measure of the magnitude of an earthquake. Hiroo Kanamori of Cal Tech has developed a new magnitude scale based on the seismic moment. The new scale extends the standard Richter scale so that it can accurately measure the strongest earthquakes without becoming saturated. For example, both the San Francisco earthquake of 1906 and the Alaskan earthquake of 1964 had a surface-wave magnitude of 8.3, but the seismic moment of the Alaskan earthquake was 100 times greater than that of the San Francisco one. On Kanamori's scale the magnitude of the San Francisco earthquake has been demoted to 7.9 and that of the Alaskan earthquake has been advanced to 9.2. The strongest earthquake on record is the Chilean earthquake of 1960, with a surface-wave magnitude of 8.3 and a seismic-moment magnitude of 9.5.

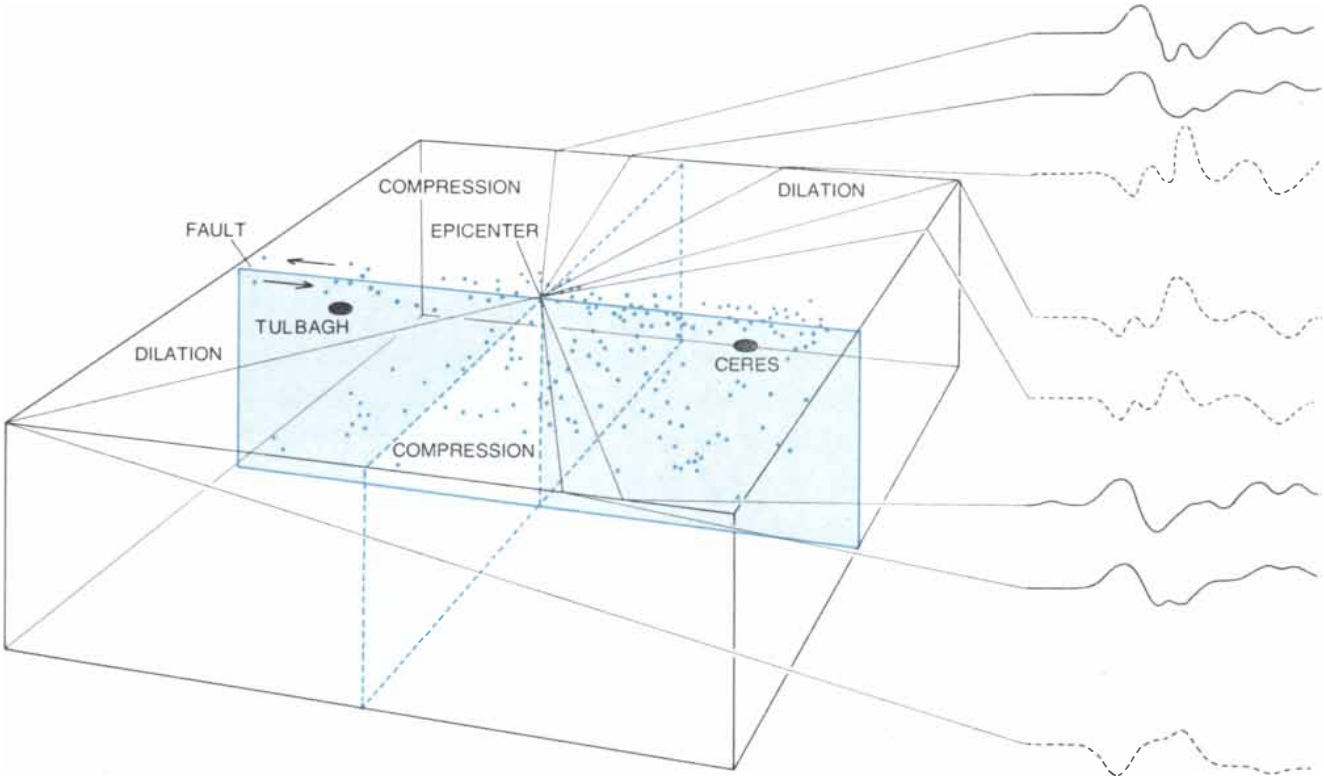
Seismic moment is more than just a convenient scale by which to rank earthquakes according to their magnitude. In 1966 Keiiti Aki of the Massachusetts Institute of Technology showed that the seismic moment is equal to the product of three factors: the average slip of the fault, the area of the rupture and the rigidity of the material that is faulted. Thus if one has independent measurements of the area of the rupture and the rigidity of the material, one can determine the average slip of the fault. The correlation between the average slip of a fault and the average strength of the resulting earthquake provides useful criteria for designing structures such as highways and pipelines that must cross active fault zones.

The total amount of slip accumulated from a number of earthquakes over a period of time also enables one to estimate the velocity at which the tectonic plates bounding the fault are moving past each other. By comparing that velocity with the velocity computed from independent geological, magnetic and geodetic evidence, it is possible to determine how much of the relative motion of the plates gives rise to earthquakes and how much gives rise to aseismic creep. It seems that in some areas, for example Chile, all the motion between plates is accomplished by earthquake slippage, and that in other regions, for example the Marianas arc in the western Pacific, the motion is accomplished by long-term steady creep.

The seismic moment and the dimensions of the fault also yield information about the amount of stress across the fault that is released during the earthquake. The drop in stress is only weakly



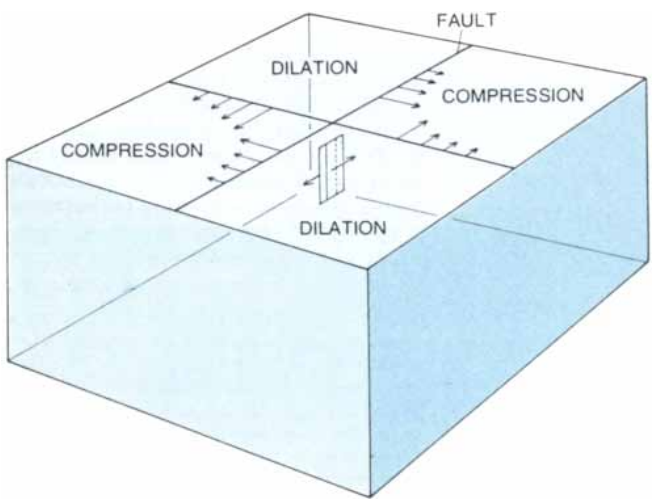
TYPICAL SEISMOGRAMS recorded by different instruments at the same site during the same earthquake can be remarkably different. The top two sets of curves are the recordings of an accelerograph and a Carder displacement meter at El Centro, Calif., from an earthquake at Borrego Mountain, some 60 kilometers away. Both instruments were triggered by the initial *P* wave, or compression wave, from the earthquake; the first strong pulse on each recording is the slower-traveling *S* wave, or shear wave, which arrived seconds later. The prominent reverberations on the recording from the Carder displacement meter are resonances of the seismic waves in the thick blanket of sediments in the Imperial Valley. The bottom pair of curves is the recording made at La Paz in Bolivia of the vertical component of the initial *P* wave from the same earthquake that was recorded by a short-period seismograph and a long-period seismograph in the Worldwide Network. By the time seismic waves had traveled to La Paz, a fifth of the way around the world, *S* waves (*not shown*) arrived approximately nine minutes later than *P* waves.

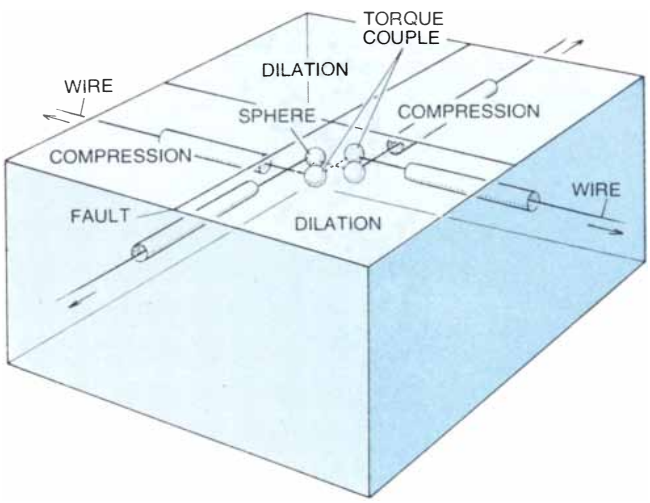


ORIENTATION OF A FAULT below the surface can be detected from the way the ground is initially compressed and dilated around the epicenter of an earthquake. This pattern of compressions and dilations is preserved in the seismic waves that are radiated by the earthquake source. In the illustration portions of seismograms (*right*) recorded during an earthquake near Ceres in South Africa show how the phase of the initial seismic waves received was shifted with azimuth between the source and the recording station. From this information alone the fault could be either of two orthogonal planes: the actual fault (*dark color*) or an imaginary plane perpendicular to it (*light color*). The path of actual fault can be determined from the location of earthquake aftershocks (*dots*) which lie along a single plane.

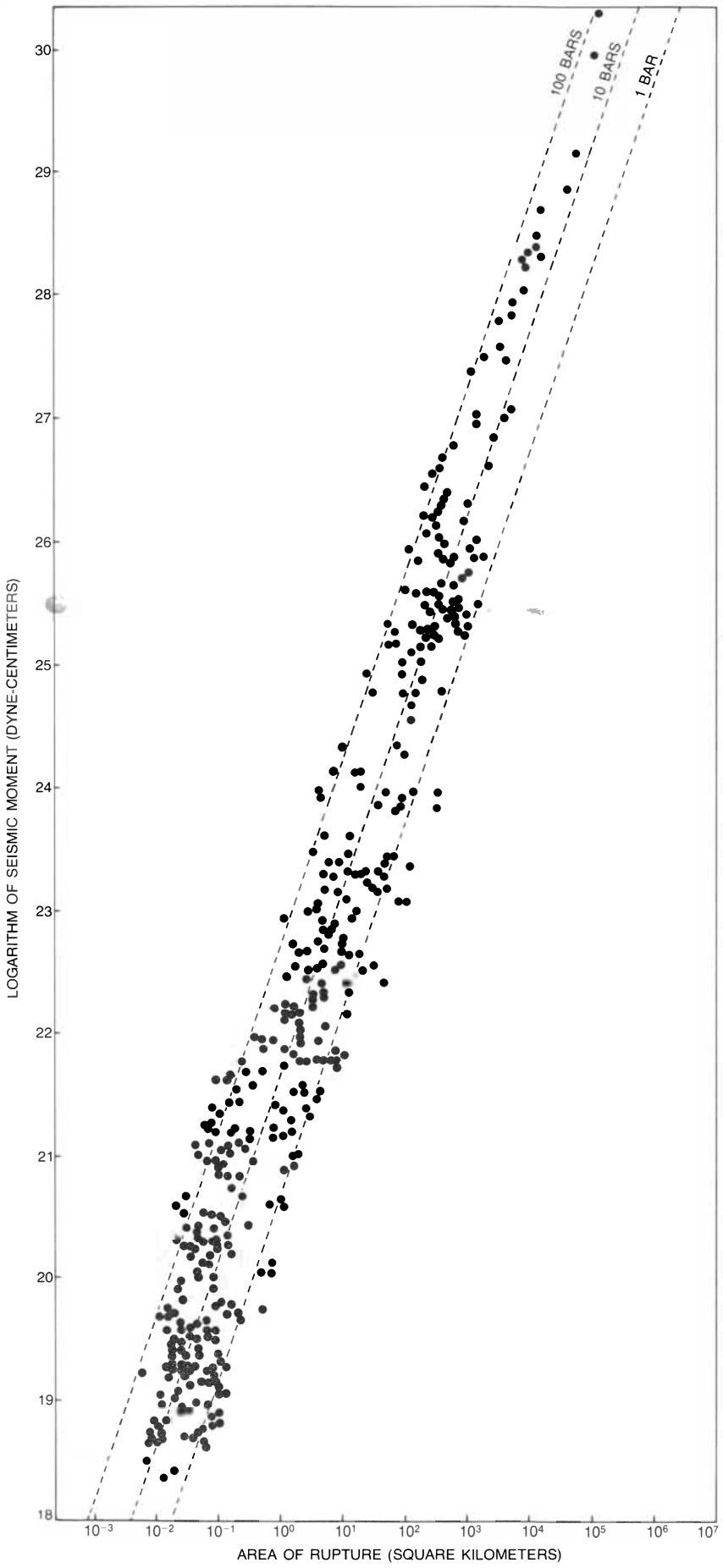
dependent on the magnitude of the earthquake. Most measurements during large earthquakes indicate that the drop in stress is between 10 and 100 bars. (A bar is 15 pounds per square inch.) The absolute, or total, stress on the faulted material could be considerably higher, but the radiated seismic waves are influenced only by the change in the stress across the fault and not by the absolute stress. Why the drop in stress should be essentially constant for earthquakes spanning such a great range of magnitude is under active debate; the explanation probably lies in the physical properties of the materials within the fault zone and in the forces driving the lithospheric plates.

The properties of the earthquake sources I have discussed so far have been deduced from seismograms made





DOUBLE-COUPLE SOURCE MODEL is mathematically equivalent to the slippage of an earthquake fault. When a small fault slips (*left*), the material closest to it slips more (*longer arrows*) than the material farther away (*shorter arrows*). Thus the material around the fault is compressed and dilated. The same deformation pattern can also be obtained if opposite torques are exerted on two torque couples embedded in an elastic medium (*right*). A torque couple can be visualized as a pair of spheres with a wire attached to each sphere running through a frictionless tube to exterior of medium. When wires are pulled with equal force, elastic medium is deformed in same way as material around a fault. Moment, or amount of torque exerted, is a good measure of strength of earthquake producing the deformation.



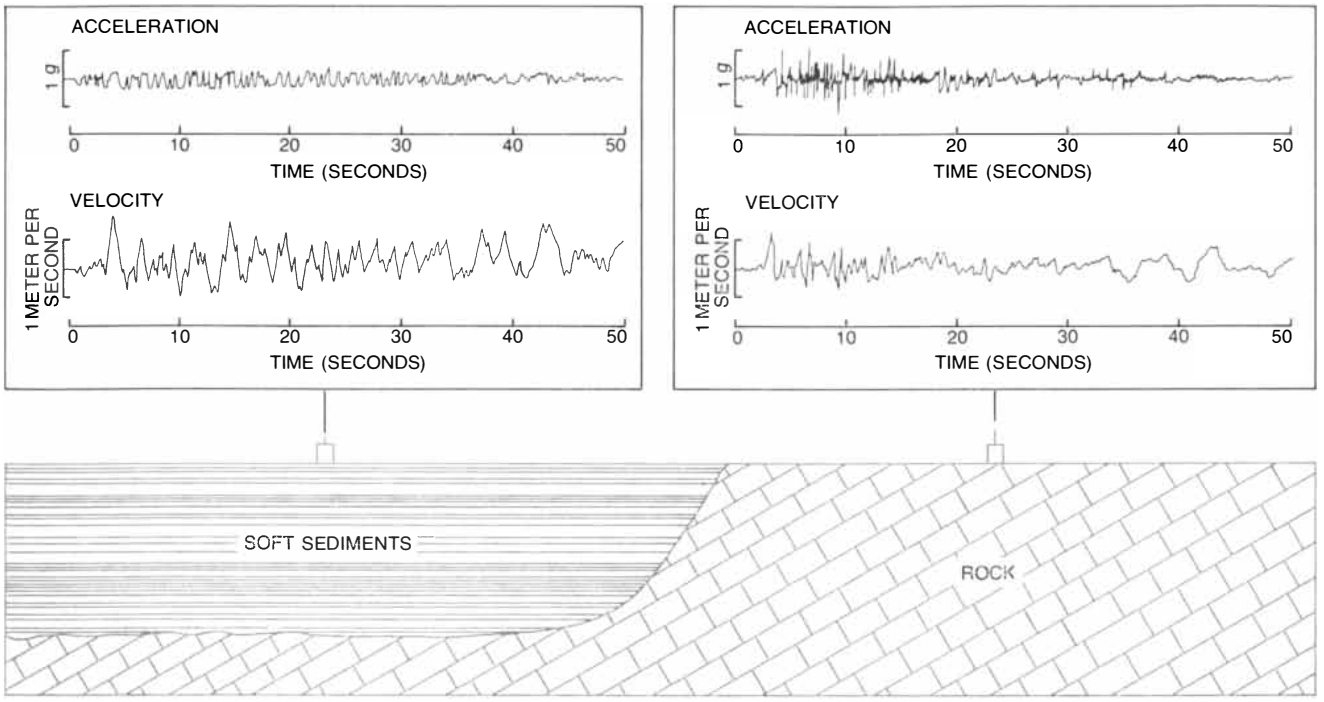
DROP IN STRESS across a fault during a large earthquake seems to be independent of the strength of the earthquake. The dots represent measurements of seismic moment obtained during many earthquakes with respect to the size of the rupture in square kilometers. Stress drop is inferred from measurements. Lines of constant stress drop are shown. Scatter in measurements for smaller earthquakes may be due in part to experimental error. A bar is a unit of pressure equal to 15 pounds per square inch; a dyne is a unit of force required to impart an acceleration of one centimeter per second per second to a mass of one gram.

at stations far from the source. Observations at a distance, however, rarely make it possible to resolve the detailed structure of the source. Recently seismologists have been devoting an increasing effort to gaining an understanding of the intricate effects involved in the propagation of the rupture along the fault and in the distortion of the seismic radiation by geological heterogeneities near the fault. This understanding is essential for the design of structures to withstand ground shaking.

Man-made structures are particularly susceptible to earthquakes because the seismic waves have frequencies that coincide with the resonant frequencies of the structures (which range from a tenth of a hertz for large structures such as the Empire State Building up to 30 hertz or even higher for small structures such as systems of pipes in an industrial plant) and because the largest ground motions are usually in the horizontal plane. All buildings are inherently capable of withstanding large vertical forces (at least 1 *g*, or the force exerted on them by the earth's gravity) but special precautions must be followed in earthquake country to ensure adequate resistance to large horizontal forces.

In general the most destructive ground motions have wavelengths smaller than the dimensions of the earthquake fault. Therefore the ground motions are strongly influenced by the details of the rupture process, such as the speed at which the rupture travels over the fault surface, the frictional strength of the fault and the drop in stress across the fault. Geological heterogeneities in the path of the seismic waves can also affect the waves' amplitude and frequency; a seismogram recorded at two stations close to each other may differ significantly. In the past seismologists have rarely been lucky enough to have a good distribution of seismographs close to the source of a major earthquake, and the few seismographs that have been close to the fault have usually been shaken so violently that the recording pen was thrown off the paper. Accordingly the short-period seismic waves are not as well understood as the long-period ones.

In recent years several types of inex-



LOCAL GEOLOGY AFFECTS GROUND MOTION near the recording site. The waves propagating from the hypocenter up to the earth's surface slow down as they encounter the deformable rocks near the surface, and in general their amplitude increases in much the same way that the amplitude of an ocean wave increases as it approaches the shore. When soft sediments are subjected to strong shaking, however, the amplitude of the motion can actually be reduced. Seismograms at the right are hypothetical recordings of acceleration and velocity of ground for an area underlain by hard rock. Seismograms for a nearby area underlain by sediments (*left*) show that the ground moves faster but amplitude of its acceleration is less. Acceleration is given in terms of *g*, acceleration of gravity at earth's surface.

pensive, rugged and reliable low-magnification instruments have been designed and installed in large numbers near many earthquake faults. The most widely used instrument is the accelerograph, which measures the acceleration of the shaking ground. There are now more than 1,200 accelerographs on station in California alone. Even with so many instruments now in operation we still do not know much about the ground motions close to a fault during a severe earthquake. So far only two useful recordings of an earthquake of magnitude 7.0 or greater have been obtained within 40 kilometers of a fault, and one of them was obtained during an earthquake in the U.S.S.R. To a large extent this lack of data is due to the fact that there have been no large earthquakes in the U.S. in the four or five years since most of the accelerographs were installed.

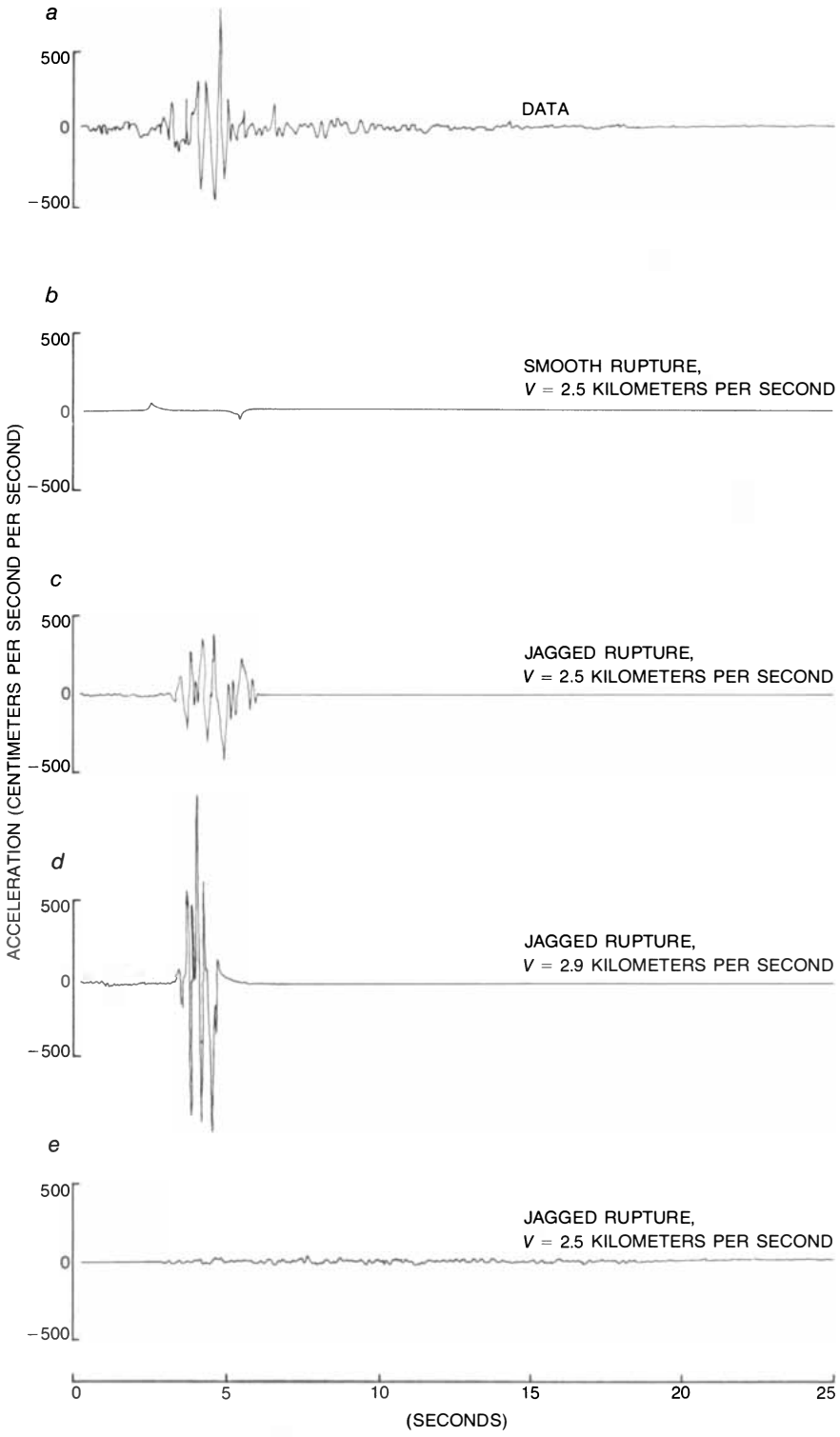
The measurements that do exist have been the main resource for estimating the strength of the ground motion. The few recordings close to faults have had a disproportionate influence on earthquake engineering design, even though these data may not be truly representative of the motions close to future earthquakes. As might be expected, the few close-in recordings have received intensive scrutiny. For example, an accelerograph on a rock abutment near the Pacoima Dam in California during the San Fernando earthquake recorded a peak acceleration of nearly 1.5 *g*, the largest acceleration yet recorded near an earthquake. The record was obtained in a region of exceptionally rugged cliffs and hills, and numerical simulations of the propagation of the seismic waves suggest that the topography may have amplified the ground acceleration by as much as 50 percent with respect to the motions that would be expected on flat ground.

At distances beyond 10 or 20 kilometers from the fault there are a fair number of recordings for earthquakes of magnitude less than 7.0. It is convenient to study the peak acceleration of the ground, expressed in terms of the acceleration of gravity at the earth's surface (*g*), which can be measured directly from the accelerograph records. The peak acceleration expected is widely used by engineers to specify the ground motion a structure should be able to withstand. The peak acceleration of the ground decreases with distance from the fault, both because the seismic waves spread out as they propagate away from the source and because their energy is attenuated by the slight inelasticity of the rocks through which they propagate. Between 20 and 200 kilometers from the fault the peak acceleration decreases approximately as the inverse square of the distance from the fault.

The data that have been obtained 20 kilometers or more from the fault imply that the peak acceleration of the ground is correlated with the earthquake's magnitude. Contrary to what one might expect from the definition of magnitude, however, earthquakes differing by one unit of magnitude do not generate peak accelerations differing by a factor of 10. Moreover, the few available data obtained some 10 kilometers from the fault indicate that very close to an earthquake, peak acceleration is hardly correlated with magnitude at all. For example, an accelerograph close to a fault near Oroville, Calif., recorded a peak acceleration of .6 *g* during an earthquake of magnitude 3.4 but another instrument near a fault in the Imperial Valley recorded a peak acceleration of only .4 *g* during an earthquake of magnitude 7.1.

The lack of correlation between peak acceleration and magnitude is easily understood. The seismic waves measured by accelerographs have a dominant frequency of about four hertz, much higher than the frequency at which the magnitude of the earthquake is measured. For all earthquakes but the smallest, seismic waves with a frequency of four hertz have a wavelength much shorter than the dimensions of the fault. Thus peak acceleration is not a good measure of the strength of large earthquakes, and for the same reason that magnitude is not. The duration of the ground motion is probably much better correlated with earthquake strength.

Strong-motion seismology is a new



SYNTHETIC ACCELEROGRAMS were constructed on a computer by the author and William B. Joyner of the U.S. Geological Survey in order to determine experimentally how an earthquake generated observed ground shaking. An actual accelerogram is shown at the top (*a*). If the earthquake were produced by a smooth rupture of the fault propagating toward the theoretical seismographic station, its accelerogram would consist of a few simple isolated peaks corresponding to the radiation emitted as the rupture started and stopped (*b*). The peaks are small because the rupture was constrained to have a gradual acceleration and deceleration at the ends of the fault. Actual data, however, generally show a more continuous shaking. To simulate this shaking random fluctuations were added to the amount by which the fault slipped. The resulting theoretical curve looked more like the actual data (*c*). Next the author let the rupture propagate toward the theoretical accelerograph at a velocity close to the velocity of seismic waves in the surrounding material. The seismic radiation from the fault then arrived in a sharp peak (*d*). When rupture propagated away from theoretical accelerograph, however, the seismic radiation was spread over a longer time interval and its amplitude was reduced (*e*).

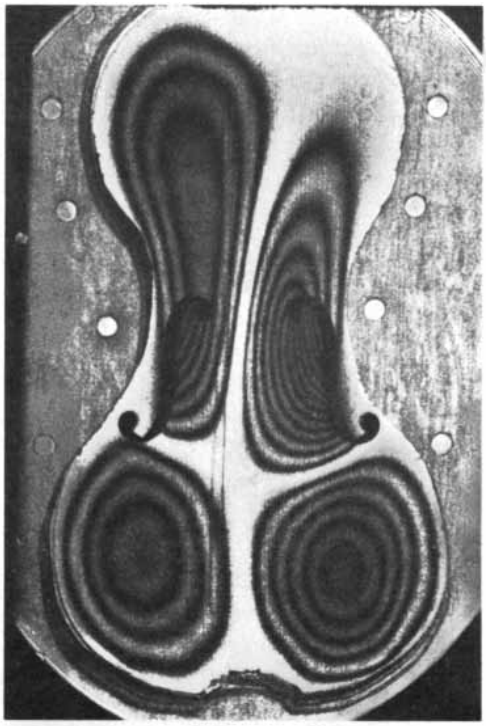
disiscipline, and there are many unknowns in it. In the future data obtained by means of accelerographs and other instruments close to earthquake faults should provide information about both the complexities of earthquake sources and the ground motions they generate. The theoretical and computational models of the seismic source should also improve. Such information will be of direct value to engineers designing major structures. Until that information is available, however, architects and engineers must continue to design buildings on the basis of the few data that do exist, some simple theoretical scaling arguments and plain educated guesses.

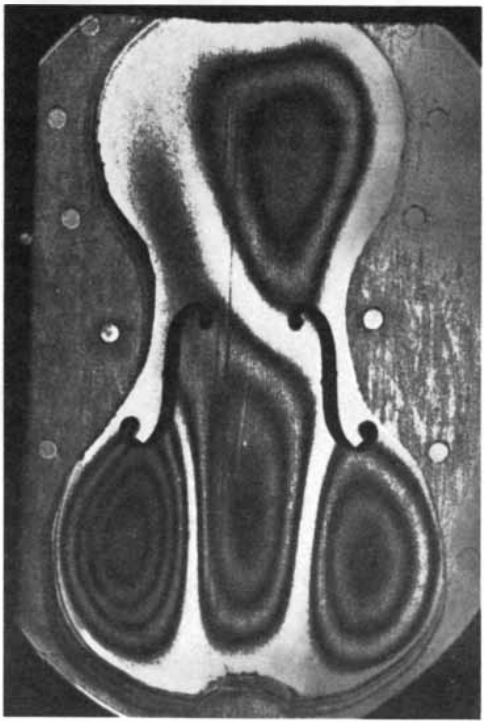
Seismologists studying models of the strong ground motions near faults are just beginning to recognize that many of the problems facing them have an essentially statistical character. For years engineers designing major structures in earthquake zones have treated accelerograms of short-period motions as recordings of random noise. On that basis they have devised many ways to generate random series of short-period motions that look much like the accelerograph recordings. The random series were generated in such a way that they matched certain constraints derived from existing data, but they paid scant attention to the physics of the earthquake source. This engineering approach is certainly a reasonable first approximation on which to base the design of a building, but it is of little value in determining what is actually happening below the ground.

Seismologists, on the other hand, have tried to predict the ground motion from earthquakes purely on the basis of deterministic models. In these models earthquake sources have been idealized as simple faults in layer-cake geological structures. Such deterministic models have been relatively successful in predicting only the long-period components of the ground motions.

Clearly the time has come to merge the engineer's statistical view with the seismologist's deterministic one. A number of seismologists are now attempting such a synthesis. Predictions of the ground motion are, however, only as good as the statistical distributions incorporated into the model and physical knowledge of the earthquake source: the properties of the fault surface and of the surrounding rocks and soil. For that information we must not only study existing strong-motion recordings but also draw on other fields such as rock and soil mechanics. I foresee an exciting future in which the skills and the learning of many disciplines, ranging from classical seismology to soil engineering, are combined to gain a better understanding of the nature of earthquakes and to reduce the hazards they create.

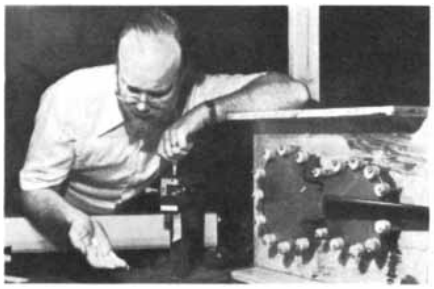
SO-115
a new universal b&w film for the sophisticated user





These are hologram reconstructions, photographs made with a camera receiving images originally recorded as time-average interference holograms. The dark fringes connect all points of common amplitude on the inside of the vibrating top of an experimental viola. It's part of an effort to improve the loudness, dynamic range, and playing ease of the viola.

Time-average interference holography is widely used by engineers to analyze vibration in lots of things besides violas. One of the two people who devised it is Karl A. Stetson (*J. Opt. Soc. Amer.* 55:1593 [1965]). These images are Dr. Stetson's work.



For both the holograms and the reconstructions he used He-Ne laser light, here seen impinging on his palm. He made the holograms on KODAK High Speed Holographic Film SO-253 and the reconstructions on KODAK Technical Pan Film (ESTAR-AH Base) SO-115.

Stetson gives two reasons for using SO-115 in this application: 1) high attainable contrast* to bring out the higher-order Bessel fringes for counting; 2) hardened emulsion permitting superproportional reduction of the negative with ammonium persulfate to accentuate the fringes by suppressing detail in the high-density areas that represent absence of vibration.

Other technical users may have quite different reasons for liking SO-115. With KODAK WRATTEN Filter No. 58 for photomicrographic contrast enhancement in phase contrast or Normarski illumination, as in chromosomal studies, it is about 2/3 stop faster than KODAK Photomicrography Monochrome Film SO-410 and KODAK Solar Flare Patrol Film SO-392, which are now discontinued. Addition of a dyed-gel backing against halation and curling has slightly reduced red sensitivity but not enough to bother those who need it to study the sun or the night skies. Red sensitivity still takes its plunge only around 690 nm. Extremely high resolving power. Extremely fine grain. Good latent-image stability. Stocked in 36-exposure magazines, 35 mm x 150 ft, 4 x 5 in. Ask Scientific and Technical Photography, Kodak, Rochester, N.Y. 14650 about other formats and about "POTA" developer for pictorial quality with this film rather than high contrast.

The standout characteristic of SO-115 is its extremely wide range of contrasts. Elaine Stetson, writer on early Americana and director/curator of the Noah Webster Foundation of West Hartford, Conn. likes the way the low-contrast end of its performance range and the extended red sensitivity bring out detail in antique furniture. The Stetsons are here shown photographing the traveling trunk used extensively by the young author of the famous American speller, who was also to become arbiter of the American language. Both pictures of the Stetsons on this page were taken for us by photographer Frances L. Funk at E. I. 25 on the very same SO-115 film. Karl processed them for 5 minutes in POTA developer, made up of 1.5 g of KODAK Balancing Developing Agent BD-84 and 30 g of sodium sulfite per liter of deionized water.

Kodak

*Like gamma 4 with 5 minutes in KODAK Developer D-19 at 20°C, for which exposure index is about 100.

ONE OF THE FEW LUXURY SEDANS IN THE WORLD THAT WOULDN'T BE LAUGHED OFF THE NÜRBURGRING.

In the heart of Germany there is a race course called the Nürburgring.

An awesome giant of a track, generally acknowledged to be the most arduous test of both cars and men.

All of the world's great high-performance cars have raced there—most have had their day.

Yet, few cars—and certainly no luxury sedans—have achieved a more impressive record on the Nürburgring than those built by the Bavarian Motor Works of Munich, Germany.

Luxury sedans? Yes. But luxury sedans built by racing engineers. German engineers who believe that extraordinary performance is the only thing that makes an expensive car worth the money

PERFORMANCE PERFECTED ON THE RACE TRACK.

While it is, of course, feasible to develop an acceptable automobile in the relative vacuum of the test track and the laboratory, it is virtually impossible to simulate the perfection demanded by motor racing.

Motor racing enables BMW engineers to develop ideas and experiment without the inhibiting constraints of economics or the cost of production—a crucial role in the development of a true high-performance automobile.

And the BMW 530i is a direct reflection of this cache of engineering intelligence.

Its suspension—independent on all four wheels—is quick and clean through the corners; its steering sharp and accurate.

Its four-speed manual transmission (automatic is available) slips precisely into each gear. And its acceleration comes up smoothly, with the turbine-like whine so characteristic of the justifiably renowned 3-liter BMW engine.

PARTICIPATION NOT ISOLATION.

The interior of the conventional luxury sedan is deliberately engineered to isolate the driver from the mechanical workings of the automobile, the world outside and the road beneath.

The interior of the BMW 530i is carefully engineered to include the driver as one of the integral, functioning parts of the car itself—the human part that completes the mechanical circuit.

Careful study has been made of the critical interrelation between seat location, visual position, steering wheel, pedals and controls.

Luxurious? Yes. Yet all functions have been assiduously planned to facilitate total, precise control at all times, under all conditions.

So successful is this integration of man and machine that when you drive the BMW 530i for the first time, you will experience an almost total oneness with the car. A unique feeling of effortless control which, if you're accustomed to conventional luxury sedans, will be completely and pleasantly new to you.

Surprising? Only to those who have never driven a BMW.

As the editors of Motor Trend magazine observed, "The reaction to a BMW is always the same. The first time driver takes the wheel and after a few minutes no other automobile will ever be the same again."

If the thought of owning such a car intrigues you, call us anytime, toll-free, at 800-243-6000 (Conn. 1-800-882-6500) and we'll arrange a thorough test drive for you at your convenience.

THE ULTIMATE DRIVING MACHINE.

Bavarian Motor Works, Munich, Germany

SCIENCE AND THE CITIZEN

The Nobel Prizes

The 1977 Nobel prizes in science were awarded to a medical physicist who developed a sensitive radioactive-isotope method for measuring trace amounts of hormones and other biological substances in the body fluids and jointly to two biologists who tracked down and isolated a new class of brain hormones; to three theoretical physicists whose contributions to elucidating the nature of the solid state of matter span half a century; to a physical chemist who has shown how systems can organize themselves into orderly structures in seeming defiance of the second law of thermodynamics, and to two economists for their theories of international trade. Each of the four prizes this year is valued at about $145,000.

Half of the prize in physiology was awarded to Rosalyn S. Yalow of the Veterans Administration Hospital in the Bronx; the other half was shared by Roger C. L. Guillemin of the Salk Institute for Biological Studies and Andrew V. Schally of the Tulane University School of Medicine and the Veterans Administration Hospital in New Orleans. Yalow, the sixth woman to receive a Nobel prize in the 77-year history of the science awards, obtained her Ph.D. in physics from the University of Illinois in 1945. In the 1950's she and Solomon A. Berson, who died in 1972, began developing methods of using radioactive isotopes to measure the level of insulin and other hormones circulating in the body. The Nobel committee described the highly sensitive radioimmunoassay (RIA) technique that ultimately evolved as "a spectacular combination of immunology, isotope research, mathematics and physics." The unknown concentration of a substance is obtained by comparing its inhibitory effect on the binding of radioactively labeled antigen to a limited amount of specific antibody with the inhibitory effect of standard known antigens. The RIA test is now used routinely to measure the physiological concentration of hundreds of hormones, enzymes, vitamins, viruses and drugs.

Guillemin and Schally, who were aided in their work by the RIA test, began more than 20 years ago to look for the hypothalamic hormones that were suspected of controlling the pituitary gland, the "master" organ whose own secretions regulate growth and structural development and the production of the thyroid, adrenocortical and sex hormones. It had been conjectured that the hypothalamus of the brain and the pituitary might act in concert not only from their physical proximity at the base of the brain but also from their development in the embryo. For a period in the 1950's and 1960's Guillemin and Schally collaborated at the Baylor University College of Medicine. In 1962 Schally moved to Tulane and in 1970 Guillemin joined the Salk Institute.

The two men were cited for discovering that the hypothalamus exercises its control over the pituitary not through nerve impulses but through low-molecular-weight polypeptides that pass through portal veins into the pituitary. From hundreds of tons of sheep-brain tissue the Guillemin and Schally groups extracted milligram amounts of several hypothalamic hormones. One, called TRF (for thyrotropin-releasing factor), controls the secretion of the pituitary hormone thyrotropin, which in turn regulates the function of the thyroid gland. A second hypothalamic hormone, LRF (for luteinizing-hormone-releasing factor), causes the pituitary hormone to release LH (the luteinizing hormone), which travels to the gonads and controls the reproductive functions of both men and women. Both TRF and LRF are now manufactured synthetically and are used in clinical medicine. In 1972 Guillemin, in collaboration with Roger Burgus, Nicholas Ling, Catherine Rivier and Wylie Vale, isolated and synthesized a third hypothalamic factor, somatostatin, which regulates the secretion of growth hormone from the pituitary gland and also of insulin and glucagon from the pancreas. Somatostatin is unusual in that it is manufactured in other parts of the body as well as in the hypothalamus.

The prize in physics was awarded jointly to Philip W. Anderson of Bell Laboratories and Princeton University, to Sir Nevill Mott of the University of Cambridge and to John H. Van Vleck of Harvard University "for their fundamental theoretical investigations of the electronic structure of magnetic and disordered systems." The eldest of the three, Van Vleck recalls being concerned that Harvard would not accept his 1922 dissertation dealing with a crossed-orbit model of helium calculated by the "old," or Bohr, quantum theory "because it was a purely theoretical thesis, a heretical novelty for the Harvard physics department." In 1925 he finished his first major work, *Quantum Principles and Line Spectra,* only to find that it had been made obsolete by the new quantum mechanics before it was off the press. Undaunted, Van Vleck quickly became one of the first Americans to master the radical new concepts. His book *Electric and Magnetic Susceptibilities,* published in 1932, served as a basic text in quantum mechanics for generations of graduate students. Van Vleck's subsequent contributions influenced the work of chemists and astronomers (who have found in interstellar space the absorption lines of some of the molecules he studied theoretically) as well as solid-state physicists.

Anderson wrote his Ph.D. dissertation under Van Vleck and joined Bell Laboratories in 1949. He has made important contributions to the theoretical understanding of ferroelectricity, ferromagnetism and antiferromagnetism, magnetic resonance, superconductivity and superfluidity, and the electrical behavior of amorphous materials. Anderson's 1958 paper, "Absence of Diffusion in Certain Random Lattices," caught the attention of Mott, who invited him to lecture at the Cavendish Laboratory at Cambridge. From 1967 to 1975 Anderson taught at Cambridge in the fall term; he now divides his time between Bell Laboratories and Princeton.

Mott, director of the Cavendish Laboratory from 1954 to 1971, has been a pioneer in developing the electronic theory of metals. According to Anderson, it was Mott who first saw how to use concepts in Anderson's 1958 paper to make sense of experimental data that had been puzzling investigators of solid-state phenomena, particularly in amorphous materials. For example, Mott explained why nickel oxide, which ought to be a metallic conductor according to the standard band theory of crystals, is actually an insulator. Mott's refined theory takes electron-electron interactions into account.

The prize in chemistry went to Ilya Prigogine of the Free University of Brussels, who is also director of the Statistical Mechanics and Thermodynamics Center of the University of Texas at Austin. Classical thermodynamics has the limitation that it deals only with reversible processes and transitions between states of equilibrium. It does not deal, for example, with a thermocouple, in which an electric current flows between two junctions maintained at different temperatures. Lars Onsager received the Nobel prize in chemistry in 1968 for developing exact methods for dealing with such phenomena.

Prigogine extended thermodynamic theory to embrace systems that are even further from thermodynamic equilibrium than a thermocouple. He has demonstrated that under such conditions new forms of ordered structures can exist. He calls them "dissipative structures" to stress that they exist only by exchanging energy with the surrounding environment. When dissipative structures are close to equilibrium, their order tends to be destroyed; when the order is far from equilibrium, however, it is maintained and new structural pat-

VO

The Standard of Giving.

Give Seagram's V.O.

Bottled in Canada. Preferred throughout the world.

CANADIAN WHISKY. A BLEND OF CANADA'S FINEST WHISKIES. 6 YEARS OLD. 86.8 PROOF. SEAGRAM DISTILLERS CO., N.Y.C. GIFT-WRAPPED AT NO EXTRA CHARGE.

terns may appear. In short, Prigogine has shown mathematically that it is possible for order to be created out of disorder. The application of his theorems to living systems is being explored.

The prize in economic sciences was awarded jointly to Bertil Ohlin, professor emeritus of economics at the Stockholm School of Economics, and to James E. Meade, formerly professor of economics at Cambridge. Considered a child prodigy, Ohlin was made a full professor of economics in 1924 at the age of 25. Between 1944 and 1967 he served in the Swedish Parliament as head of the Swedish Liberal Party. In his major work, *Interregional and International Trade* (1933), he explored the factors that determine, on the one hand, the pattern of foreign trade and the international division of labor and, on the other, what effect foreign trade has on the allocation of resources, price relations and the distribution of income. More than 30 years ago Ohlin foresaw that modern welfare states would have difficulty with the problem now called "stagflation": the tendency of wages and prices to rise before full employment is achieved.

Meade became widely known among economists in the 1930's for his editing of the League of Nations' World Economic Surveys. During World War II he shared with John Maynard Keynes responsibility for directing the British economy. Meade's major work is a two-volume treatise, *The Theory of International Economic Policy,* published in 1951 and 1955.

All Ears

The search for intelligent life elsewhere in the universe is a "timely and feasible" undertaking "with substantial potential secondary benefits" that can be started now with only modest resources and can be expanded later to a much larger scale, if that turns out to be necessary and desirable. Although "intrinsically an international endeavor," such a project is unlikely to be taken seriously "until one big nation, such as the U.S., seizes the initiative and invites serious participation by others." Even if international cooperation were slow to materialize, the search for extraterrestrial intelligence "remains a ... worthwhile U.S. endeavor" and one that "could become a significant milestone in the history of our civilization."

Those are the principal conclusions to emerge from a series of workshops conducted over the past two years by an interdisciplinary group of experts from various U.S. research centers, representing such diverse fields as astrophysics, radio astronomy, planetary science, geochemistry, biology and electrical engineering. The workshops, chaired by Philip Morrison, professor of physics at the Massachusetts Institute of Technology, were sponsored by the Ames Research Center of the National Aeronautics and Space Administration. The report of the group's findings, *SETI: The Search for Extraterrestrial Intelligence,* has just been published by NASA.

In his preface to the SETI report Morrison summarizes the objectives of the Science Workshops on Interstellar Communication: "To examine systematically the validity of the fundamental criteria and axioms associated with a program to detect extraterrestrial intelligent life; to identify areas of research in the astronomical sciences, and in other fields, that would improve the confidence levels of current probability estimates relevant to SETI; to enumerate the reasons for undertaking a search, the values and risks of success, and the consequences of failure; to explore alternative methods of conducting a search; to select, in a systematic way, preferred approaches; to indicate the conceptual design of a minimum useful system as required to implement the preferred approaches; to delineate the new opportunities for astronomical research provided by the system and their implications for system design; to outline the scale and timing of the search and the resources required to carry it out; to examine the impact of conducting a search, and the impact of success or failure in terms of national, international, social and environmental considerations; and to recommend a course of action, including specific near-term activities."

In keeping with its charge the study group devotes a major section of its report to laying out the background and rationale of a SETI program and to examining the implications of such a program. In particular the report discusses the significance of the detection of intelligible signals from an extraterrestrial civilization and of the information that might be contained in such signals. The report also presents extracts of "the most interesting and significant" of the group's discussions and a selection of detailed technical arguments about aspects of the proposed SETI project.

The authors of the report make clear at the outset that what they are recommending is strictly a one-way transmission: "We do not intend to send any signals out to add to those that have already gone out from our TV transmitters and our powerful radars," they write. "Rather, we want to listen, to search all the directions of space ... to seek possible signals. Perhaps it will be only an accidental signal, as we have made ourselves.... Or perhaps there is a deliberate signal, a beacon for identification, or even a network of communication. There seems no way to know without trying the search."

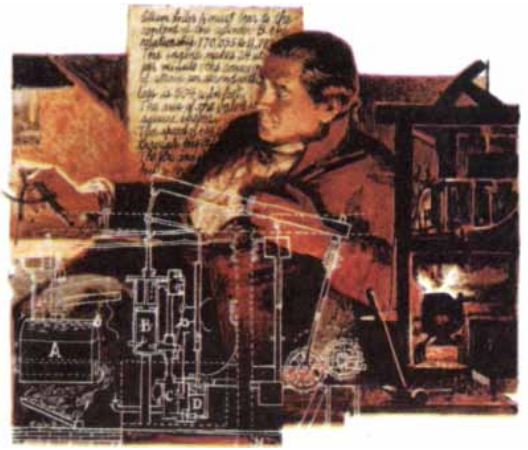
The group proposes that the search for signals from "our interstellar neighbors" concentrate initially on the microwave part of the radio spectrum, using existing radio telescopes on a part-time basis. "If we but equip existing radio telescopes with low-cost state-of-the-art receiving and data processing devices, we will have both the sensitivity to explore the vicinity of nearby stars for transmitters similar to Earth's, and to explore the entire Galaxy for more powerful signals, or for signals beamed at us. Such explorations, even should they yield negative results, would decrease our uncertainty concerning whether intelligent life transmitting powerful signals may lie beyond our solar system. At the very least, it would be of great interest and some importance either to know we have near neighbors, or to be reasonably confident no nearby transmitting civilizations exist. If, *after* we have made such modest searches, it seems important to us to embark upon a more ambitious SETI program, ... the experience we will have gained will prove not only invaluable, but essential."

The group thus concurs in the original proposal put forward by Morrison and Giuseppe Cocconi in 1959 that currently available radio-astronomy technology provides an adequate means of establishing contact with extraterrestrial civilizations, particularly at microwave frequencies near the spectral line of hydrogen. (This comparatively quiet region of the electromagnetic spectrum, now generally accepted as the prime spectral band to be searched for interstellar signals, has in recent years been dubbed the "water hole," since it is bounded at the low-frequency end by the hydrogen line at 1,420 megahertz and at the high-frequency end by the hydroxyl lines between 1,612 and 1,720 megahertz.)

Beginning in 1960 there have been several deliberate attempts, both in the U.S. and in the U.S.S.R., to detect signals of extraterrestrial origin with the aid of radio telescopes, all with negative results. The present report points out, however, that in most cases "only a few select objects have been observed, at a few discrete frequencies using relatively wide bandwidths and at only moderate to low sensitivities." Morrison remarks: "It is clear that the task has been barely begun."

In recommending that the U.S., through NASA, play a leading role in the search for extraterrestrial intelligence, the members of the SETI study group disavow "the pursuit of narrow national advantage." On the contrary, they argue, the U.S. "has frequently demonstrated the will and foresight to take the initiative in programs of worldwide benefit. The U.S. space program has provided not only excitement and scientific knowledge, but numerous practical satellite services not for this country alone, but for the whole world. It is in this same spirit of providing a focal point for international cooperation and support that we feel the U.S.

Thomas Savery patented the first steam engine before James Watt was born. But it remained for Watt, while strolling in a Glasgow park, to figure out how to triple the amount of work a coal-fired steam engine could do. He thus became the father of a radically redesigned engine that made a critical stride in efficiency. He even figured out a way to measure the work an engine could do. He called it *horsepower.*



Mr. Watt, until you took that walk, the steam engine was just a toy.

Look where the science of energy is today.

Man is still trying to make the most of his energy sources: the sun, wind, the atom, coal, gas, water, oil, and the earth's heat.

Take the sun. Every 40 minutes, the earth receives enough solar energy to meet man's needs for a year. The trick is to collect and store it — economically.

Consider coal. America's 400-year supply can be burned, or turned into gas and oil. Cleaning up coal's act and converting it at acceptable cost are problems yet to be solved.

Rockwell is addressing a variety of problems like this in a variety of ways through a variety of technologies.

Our annual report talks about those efforts. If you'd like a copy, just write Rockwell International, Dept. 815S, 600 Grant Street, Pittsburgh, PA 15219.

Rockwell valves up to 15 feet high and 25 tons heavy are in nearly every U.S. nuclear power plant. In order to close quickly enough, valves like this must generate 1,000,000 pounds of thrust — as much thrust as rockets produced when they put America's first satellites into orbit.

Today, 2% of all America's energy comes from nuclear power. By 2000 AD, that number must reach 25-30% to meet our energy needs. Rockwell equipment and technology are involved in today's plants — and in a new breeder that could produce as much electricity from 1.5 tons of uranium as today's plants do from 150 tons.

One Rockwell technology in development extracts oil from coal. Another extracts gas from coal. A third lets power plants burn high-sulphur coal without contaminating the air.

According to American Petroleum Institute estimates, the U.S. has less than a 20-year supply of known oil and gas reserves. In present and future fields, you will find our McEvoy oil-field equipment helping to bring out those precious and diminishing energy reserves with a minimum of waste.

Rockwell International

...where science gets down to business

can and should take the initiative" in the search for extraterrestrial intelligence.

The members of the NASA interstellar-communication workshop who signed the report were, in addition to Morrison, Ronald N. Bracewell (Stanford University), Harrison S. Brown (California Institute of Technology), A. G. W. Cameron (Harvard College Observatory), Frank D. Drake (Cornell University), Jesse L. Greenstein (Cal Tech), Fred T. Haddock (University of Michigan), George H. Herbig (University of California at Santa Cruz), Arthur Kantrowitz (AVCO Everett Research Laboratory), Kenneth I. Kellerman (National Radio Astronomy Observatory), Joshua Lederberg (Stanford), John S. Lewis (M.I.T.), Bruce C. Murray (Jet Propulsion Laboratory of Cal Tech), Bernard M. Oliver (Hewlett-Packard Company), Carl Sagan (Cornell) and Charles H. Townes (University of California at Berkeley). The group's report was edited by Morrison with the aid of John Billingham and John H. Wolfe of the Ames Research Center.

Failed Hypothesis

The low-fat, low-cholesterol diet promoted for a generation in an attempt to reduce the incidence of coronary heart disease is based on an unproved hypothesis and is ineffective, according to George V. Mann of the Vanderbilt University School of Medicine, who conducted a broad review of the epidemiological, clinical and experimental evidence and has reported his conclusions in *The New England Journal of Medicine*. Cholesterol has long been implicated in the formation of the atherosclerotic plaque, the lesion of the artery wall that can cause coronary disease: cholesterol is present in plaques, the blood-cholesterol level is correlated with the size of plaques, and feeding cholesterol to some laboratory animals produces lesions that resemble obstructive plaques in human beings. The leap from these observations to a diet regime was made largely on the basis of Ancel B. Keys's 1953 report that in six selected countries the prevalence of coronary heart disease was correlated with food-fat supply. In a few years, Mann writes, "that fragile hypothesis" had been transformed "into treatment dogma." Dissenting investigators tended not to receive research funding "because the peer-review system rewards conformity and excludes criticism."

Mann cites four kinds of "evidence that the diet-heart hypothesis is wrong." The first has to do with dietary habits and blood-fat levels. Detailed interviews with some 1,000 individuals enrolled in the well-known Framingham, Mass., study failed to show a relation between their dietary practices and cholesteremia (cholesterol in the blood); diet-recall interviews with 2,000 people

enrolled in another community study showed no relation between diet and blood-fat levels.

Second, U.S. trends in mortality and fat intake fail to support the hypothesis. The overall mortality rate has been quite constant in the period since 1950; during this period better methods have been developed for coronary care, leaving no room for an effect of any dietary changes. As a matter of fact, the U.S. intake of polyunsaturated fats, which according to the hypothesis should reduce blood cholesterol, has doubled since 1900 as the coronary-disease rate has soared. There is no evidence that cholesteremia diminished during the 1960's and early 1970's, when "diet-heart propaganda was full blast."

Third, the results of several clinical trials indicate that "no diet therapy has been shown effective for the prevention or treatment of coronary heart disease."

Fourth, clinical trials with drugs that reduce cholesteremia show no measurable effect on coronary disease. Since the effect of the drugs on cholesteremia is about twice as great as the effect of the most rigorous low-fat diet, "if the drugs do not help, the diets are unlikely to do as well."

The fact remains that cholesterol levels are higher in developed societies than they are in less developed ones; the data indicate that "the prevailing hypercholesteremia is contributory [to obstructive atherosclerosis] and that a major reduction would make a difference." Why, then, is the cholesterol level high in developed societies? Feeding studies (adding eggs to the diet, for example) show that it is not because of higher intake; in any case intake accounts for only 10 to 20 percent of the body's cholesterol supply, the remainder being accounted for by cholesterol synthesis within the body. Other studies, however, show that the cause is not increased synthesis. Nor is it a diminished rate of excretion of the products of cholesterol.

The explanation must be that the breakdown of cholesterol is impaired in such a way that it is accomplished only after a higher blood level of the cholesterol has been attained, leaving a larger steady-state pool of cholesterol in the body. Mann suggests that the critical step is likely to be the conversion of cholesterol to 7α-hydroxycholesterol. That conversion could be impaired by certain oxidation products of cholesterol itself or by "trans" fatty acids, which are found in the hydrogenated vegetable oils promoted for their anticholesteremic effect. Other candidates are vitamin D (the level of which may be high in the diet of developed societies) and carbon monoxide.

While efforts are being pressed to identify the impairing agent is there anything to be done? Mann writes: "The most impressive array of epidemiologic evidence suggests that fit and active people are spared the complications of atherosclerosis." Although there have been no systematic studies of this effect ("a commentary on the lost generation of diet-heart enthusiasts"), exercise is probably helpful, if perhaps only when it is pushed beyond some threshold level of exertion.

Creation at Michigan State

The coeditor of a biology textbook that espouses the biblical view of creation is now teaching a course along those lines at Michigan State University. It is apparently the first such course offered at a publicly supported university. The course is Natural Science 135; undergraduates can (but do not have to) take it to fulfill part of the university's stipulation that 12 credits (three or four courses) in the natural sciences are a prerequisite for graduation.

The course is taught by John Newton Moore, who is coeditor with Harold Schultz Slusher of Christian Heritage College in San Diego of *Biology: A Search for Order in Complexity*. Several states have authorized the use of the book in public schools. Last spring a judge in Indiana overruled the Indiana textbook commission on this score, finding that the book seeks to promote "fundamentalist Christian doctrine in the public schools" and thus violates the principle of separation of church and state laid down in the first amendment to the Constitution. The judge acted in a suit brought by the Indiana Civil Liberties Union.

The Michigan Civil Liberties Union has made no move against Natural Science 135; indeed, a spokesman said the MCLU was unaware of the course. Moore describes his course as comparing evolution and creationism. Some members of the Michigan State faculty have complained informally about the course, arguing that it has no place in a department devoted to the natural sciences. The president of Michigan State, Clifton R. Wharton, Jr. (whose appointment as chancellor of the State University of New York is due to take effect early in 1978), upheld the offering of the course. "The university has a responsibility to confront its students with a wide range of ideological positions," he said, "provided the university does not demand that any student accept a particular viewpoint.... The argument that Professor Moore violates the concept of separation of church and state seems not to apply here. We do not have a situation in which there is control, or attempted control, over the education programs by the church."

Murder in the Middle Ages

That slow decay of memory called nostalgia invariably results in the good old days being perceived as better than the bad new ones. From time to time, however, the study of history provides therapy. Working with English circuit-court records for the period between A.D. 1202 and 1276, James Buchanan Given of the University of Michigan at Dearborn has collected evidence indicating that the homicide rate in late medieval times was higher than the contemporary U.S. national rate of 9.7 murders per 100,000 population (1974) and often exceeded that of such high-crime areas as Miami, Fla., where the average rate from 1948 to 1952 was more than 15 murders per 100,000.

Given's data have recently been presented in his book *Society and Homicide in Thirteenth-Century England*, published by the Stanford University Press. In summary they cover 21 court records of the 13th century from the counties of Bedford, Kent, Norfolk, Oxford and Warwick and from the cities of London and Bristol. Over that period 3,492 persons were accused before the courts of the murder of 2,434 persons. Some 50 percent of the accused, a total of 1,758, had fled after the murder and remained unlocated at the time of trial; nearly 22 percent of the victims had been killed by persons unknown. More than 1,000 of the victims had been killed by one person; 735 others had been killed by two or more persons.

The murder weapon was most often the knife, an implement that in 13th-century England was carried by virtually all adults and even by many children. Of the 455 victims whose cause of death was specified in the court records, some 30 percent had been stabbed. The other weapons favored were in descending order the cudgel (22 percent), the axe (14 percent) and a stone (4 percent). For the rest it was apparently whatever had come to hand: a pitchfork, a scythe, a spade, a mattock, a stool, a trivet and even a piece of firewood.

Kent was the county with the highest homicide rate (23 per 100,000). Medieval London ranks below modern Miami with a rate of 12 per 100,000; medieval Bristol, with a rate of four per 100,000, falls between modern Philadelphia (5.7) and Milwaukee (2.3). Some strong qualitative similarities between 13th-century England and 20th-century America are noted by Given. For example, among 550 Philadelphia homicides with identified murderers 25 percent of the murderers belonged to the same family as the victim. Of 159 similar homicides in Kent 60 percent of the victims were either the husband (32) or the wife (64) of the murderer. Perhaps the strongest similarity of all is that both the murderers and their victims were usually poor. Of the 1,724 persons found guilty and subject to chattel confiscation, more than 60 percent had no chattels to yield. An additional 30 percent owned property having a total value of 20 shillings or less.

The Epidemiology of Influenza

The phenomenon of genetic recombination between human and animal strains of the influenza virus may be responsible for the appearance of new subtypes such as the virus that caused the great pandemic of 1918–19

by Martin M. Kaplan and Robert G. Webster

Influenza has been called the last great plague of man. Certainly the pandemic of 1918–19, which raced around the world in three waves, killing upward of 20 million people and leaving enormous social and economic disruption in its wake, ranks as one of the most devastating afflictions of any kind in human history. Over the past 250 years or so at least 10 and perhaps as many as 20 lesser influenza pandemics have swept the globe, interspersed with a much larger number of milder, more localized epidemics. What accounts for the wide variation in the extent and severity of the different outbreaks of the disease?

It has been known for many years that influenza is a viral infection and that the various appearances of the disease in human populations are caused by different strains of the influenza virus. There is at present no completely satisfactory explanation of the origin of the pandemic strains of the virus in man, but an accumulating body of evidence points to the important role played by the influenza viruses of other species, including domestic animals and migratory birds. In recent years it has become increasingly clear that the existence of a vast reservoir of influenza-virus strains in these creatures presents a continuing and evidently ineradicable threat to mankind.

History of the Disease

The highly contagious, acute respiratory illness now known as influenza appears to have afflicted human beings since ancient times. The individual symptoms and epidemiological traits of the disease are sufficiently characteristic to enable one to identify a number of major epidemics in the distant past. One such epidemic was recorded by Hippocrates, the father of medicine, in 412 B.C., and numerous episodes were described in the Middle Ages.

The term influenza was introduced in Italy early in the 15th century to describe an epidemic that was attributed to the influence of the stars. The term was adopted by the English in the 18th century; during the same period the French named the disease *la grippe*. A colorful and accurate description of the main characteristics of influenza is contained in a letter sent from Edinburgh by Lord Randolph to Lord Cecil in November, 1562:

"Maye it please your Honor, immediately upon the Quene's [Mary] arivall here, she fell acquainted with a new disease that is common in this towne, called here the newe acquayntance, which passed also throughe her whole courte, neither sparinge lordes, ladies nor damoysells not so much as ether Frenche or English. It ys a plague in their heades that have yt, and a sorenes in their stomackes, with a great coughe, that remayneth with some longer, with others shorter tyme, as yt findeth apte bodies for the nature of the disease. The queen kept her bed six days. There was no appearance of danger, nor manie that die of the disease, excepte some olde folkes. My lord of Murraye is now presently in it, the lord of Lidlington hathe had it, and I am ashamed to say that I have byne free of it, seinge it seketh adquayntance at all men's handes."

The first well-recorded pandemic occurred in 1580 and was believed to have originated in Asia; from there it spread to Africa and Europe. Mortality was high in some cities, and doubtless the toll was greatly increased by the customary practice of bleeding the sick to reduce fever. Over the next three centuries, although record keeping was irregular and reporting was often inaccurate, there were a definite number of serious pandemics (along with the intervening epidemics) that historians in general agree on. Retrospective research in the past decade has partially clarified the nature of the 1889 pandemic by testing for influenza antibodies in the blood serum of people who were alive at that time. It was not until the early 1930's, however, that a specific virus was first identified as the cause of influenza, marking the start of a better understanding of the disease.

In terms of numbers of human victims the great pandemic of 1918–19 was unprecedented. Estimates run from a minimum of 20 million killed throughout the world to as much as twice that number. More than 500,000 deaths were recorded in the U.S., and other parts of the world were equally or even more gravely struck. One authority estimated 20 million deaths in India alone, and some parts of Alaska and the Pacific islands lost more than half their population.

There was vast disruption throughout the U.S., and at the height of the epidemic community life in many cities was brought almost to a standstill. Some 25 million clinical cases of influenza were observed during the fall and winter of 1918–19: a fourth of the entire population. In Philadelphia alone during the third week of October, 1918, there were 4,600 deaths from influenza. In almost all major cities theaters and other public gathering places were shut down, and hospitals were overflowing and lacking in medical services. Previously healthy adults sometimes sickened and died within 24 hours. Entire families were stricken, often with no one to care for them, in spite of a remarkable outpouring of volunteer services throughout the country. Grotesque remedies abounded, but in the end the only effective treatment was good nursing care.

The U.S. armed forces were badly hit, slowing to a trickle the delivery of American troops on the Western Front in the summer of 1918. Deaths from influenza in the U.S. armed forces in 1918 amounted to 43,000, about 80 percent of the total number of American battle deaths in the war. The slowdown and eventual failure of the last German offensive in the spring and summer of 1918 was believed by some German generals to have been caused largely by influenza.

It is still not clear why the pandemic was so lethal. Certainly secondary bacterial infections causing pneumonia and other serious conditions accounted for many, if not most, of the deaths. Such infections can today be treated effectively with antibiotics. Another important

factor could have been a marked increase in the virulence of the virus itself during the first phases of the pandemic in the spring and summer of 1918. So far all attempts to clarify these questions have failed, including the microscopic examination of tissue specimens from victims of the disease and the exhumation in the 1950's of bodies buried in the frozen soil of Alaska in a search for the virus strain involved.

The name given to the 1918–19 pandemic was "Spanish flu," a misnomer that has persisted to this day in spite of the fact that most accounts place the first human cases elsewhere. Apparently Spain acquired the dubious distinction because of the practice of censorship by the Allied and German armies and the unwillingness of the authorities to admit to the widespread incapacitation of the troops. It is difficult to ascertain the geographic area first affected, since the pandemic came in three waves: in the spring of 1918, the fall of 1918 and the first months of 1919. Close attention was not focused on the disease until the "killer" wave of the fall of 1918 had struck. Some authorities ascribe the origin of the first wave of the epidemic to China in March, 1918, but much more dependable information can be found in the records of the U.S. Army. These records pinpoint the first cluster of influenza cases among troops stationed at Fort Riley in Kansas on March 11, 1918. That the disease spread throughout the world from that single site cannot be demonstrated conclusively, but subsequent evidence has shown that the hypothesis is not improbable.

It was because of the dread experience of the 1918–19 pandemic that apprehension was great among U.S. public-health officials when another outbreak of the disease, apparently involving the same strain of the influenza virus, made its appearance at Fort Dix in New Jersey in January, 1976.

The Search for a Cause

For centuries men speculated wildly on the cause of influenza; the stars, the weather and poisonous gases from swamps were implicated in turn. (As late as 1894 Charles Creighton, an eminent British epidemiologist, insisted that influenza was not contagious.) By the end of the 19th century, however, the microbiological concept of infectious disease had taken firm root, preparing the ground for the discovery of a bacillus in the throats of many influenza patients. This bacillus, *Haemophilus influenzae* (also known as Pfeiffer's influenza bacillus, after Richard F. J. Pfeiffer, a German bacteriologist), remained for many years the leading suspect as the causative agent of influenza. The discovery of the true viral cause came in the late 1920's, when a strain of the virus was first found in pigs. A related strain was finally isolated from a human patient in 1933.

A number of historical accounts of the disease mention the interesting coincidence of influenzalike diseases in animals, particularly in horses, immediately preceding or accompanying influenza epidemics in man. The wide variety of microbial agents other than influenza viruses that are known to cause upper-respiratory disease in animals confused the picture. It is now certain that virus-caused outbreaks of the disease have occurred among animals many times in the past. It was not until the fall of 1918, however, that a close correlation between human and animal influenza was reliably established. J. S. Koen of Fort Dodge, Iowa, a veterinarian then working as an inspector for the U.S. Depart-

GAUZE MASKS, adopted as a public-health measure in many parts of the U.S. as the second "killer" wave of influenza swept across the country in the fall of 1918, are seen here being worn by a muster of Seattle policemen in December of that year. The influenza pandemic of 1918–19 was the most deadly in history, killing at least 20 million people around the world, including more than 500,000 Americans.

ment of Agriculture, reported that a new disease that had appeared in pigs in the Middle West was strikingly similar to and coincident with human influenza among farm families. He concluded that they were one and the same disease.

After long research into the transmissibility of influenza between pigs a team of veterinarians from the Bureau of Animal Industry, led by C. N. McBryde, succeeded in 1928 in transmitting the disease by means of unfiltered mucus from the respiratory tract of stricken pigs, but an effort to transmit it with filtered material failed. Following up these findings, Richard E. Shope of the Rockefeller Institute of Comparative Pathology in Princeton, N.J., was able to show that the virus could be transmitted between pigs with filtered material, and he reported this and subsequent work in 1931.

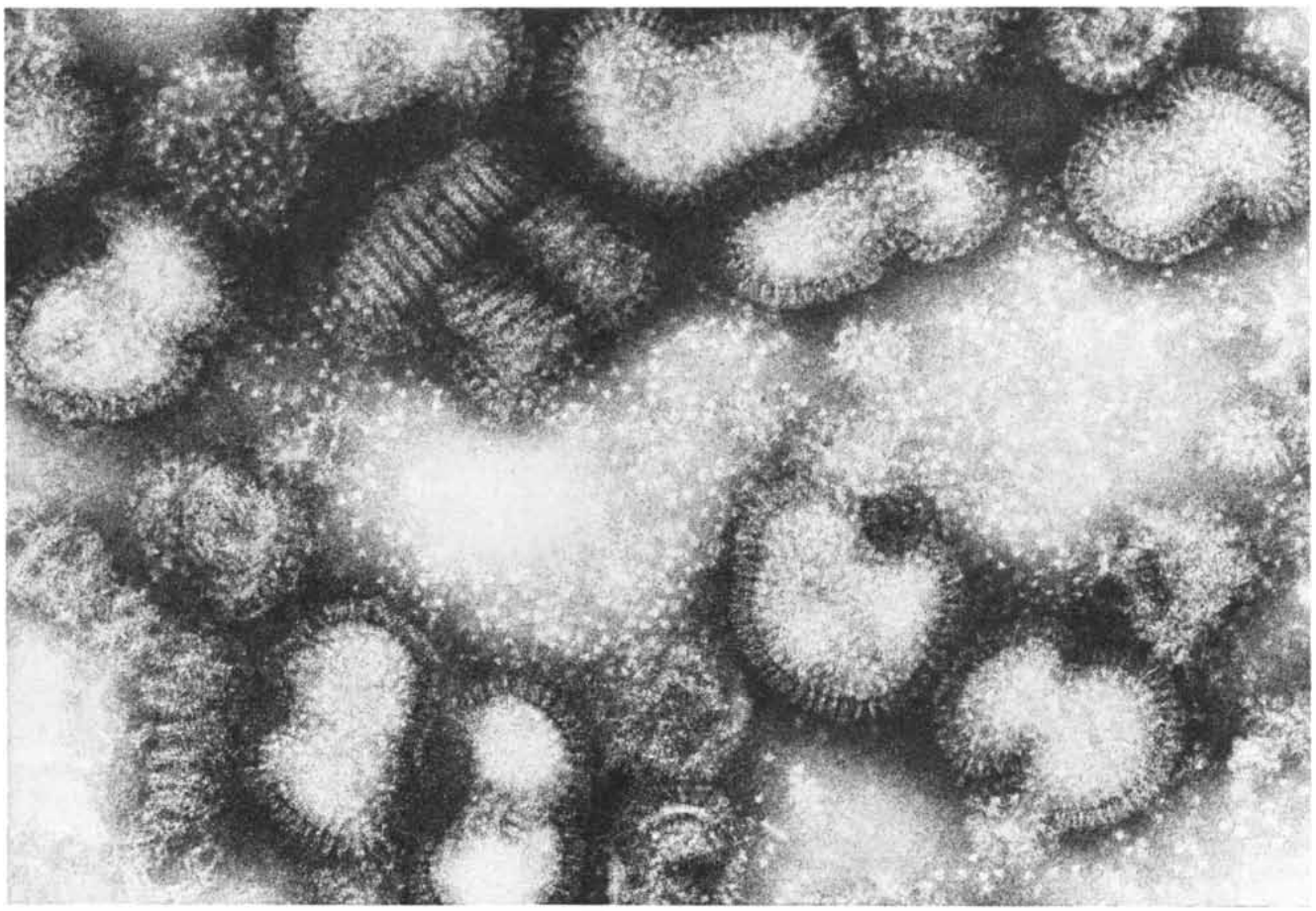
Shope's work was noted in Britain, where another attempt was made to isolate a virus during an epidemic of human influenza in 1933. Wilson Smith, Christopher Howard Andrewes and P. P. Laidlaw, working at the National Institute for Medical Research in London, finally succeeded by inoculating a filtrate of human throat washings into the nose of ferrets, which was the same route used by Shope in his pig experiments. The British team later succeeded in producing pneumonia in mice with infected nasal material from ferrets.

Rapid progress followed the demonstration that influenza virus could be transmitted to ferrets and mice. In 1940 ferrets were experimentally infected with a second type of influenza virus from man. The second human strain was designated influenza *B*, to distinguish it from the first type found, which became known as influenza *A*. (A third type, influenza *C*, was isolated from man in 1949.) Also in 1940 F. M. Burnet of Australia found that influenza viruses would multiply in the cells lining the allantoic cavity of the developing chicken embryo, and a year later George K. Hirst of the Public Health Research Institute of New York observed that influenza-infected fluid from chicken embryos would agglutinate, or clump, the red blood cells of chickens. This hemagglutination disappeared on warming, suggesting the presence of an enzyme on the virus that caused the virus and the red cells to dissociate.

These developments set the stage for further progress. The availability of high concentrations of influenza viruses obtained from developing chicken embryos led to the development of inactivated vaccines for man. The hemagglutination reaction could be inhibited by specific antibodies in the blood serum of man or animals infected or vaccinated with influenza viruses, so that a simple method became feasible for distinguishing between different strains of influenza virus and for measuring the body's immunologic response to a given strain.

The Changing Influenza Virus

An electron micrograph of the influenza virion, or virus particle, reveals a sphere approximately 100 nanometers in diameter. The surface of the particle



VIRIONS (virus particles) of the swine influenza virus are shown enlarged approximately 300,000 diameters in this electron micrograph made by Gopal Murti of St. Jude Children's Research Hospital in Memphis, Tenn. The particles are typically spherical and measure about 100 nanometers across; their form may vary widely, however, and some may even take the shape of long filaments. The spikelike projections that cover the outer surface of the intact particles are molecules of hemagglutinin (*H*) and neuraminidase (*N*). There are on the average some 500 *H* spikes and 100 *N* spikes on each spherical particle. The *H* spike enables the virus to attach itself to cells in the body; after an infection specific *H*-spike antibodies are formed that prevent reinfection with the same strain of influenza virus. The *N* spike, an enzyme, is thought to be responsible for dissolving the linkage between the virions and the infected cells, thereby facilitating the spread of the virus from cell to cell. Specific antibodies are also formed for the *N* enzyme, but they are less important in providing protection against future infection. Some of virions are disrupted, exposing inner components. White specks are liberated *H* and *N* spikes.

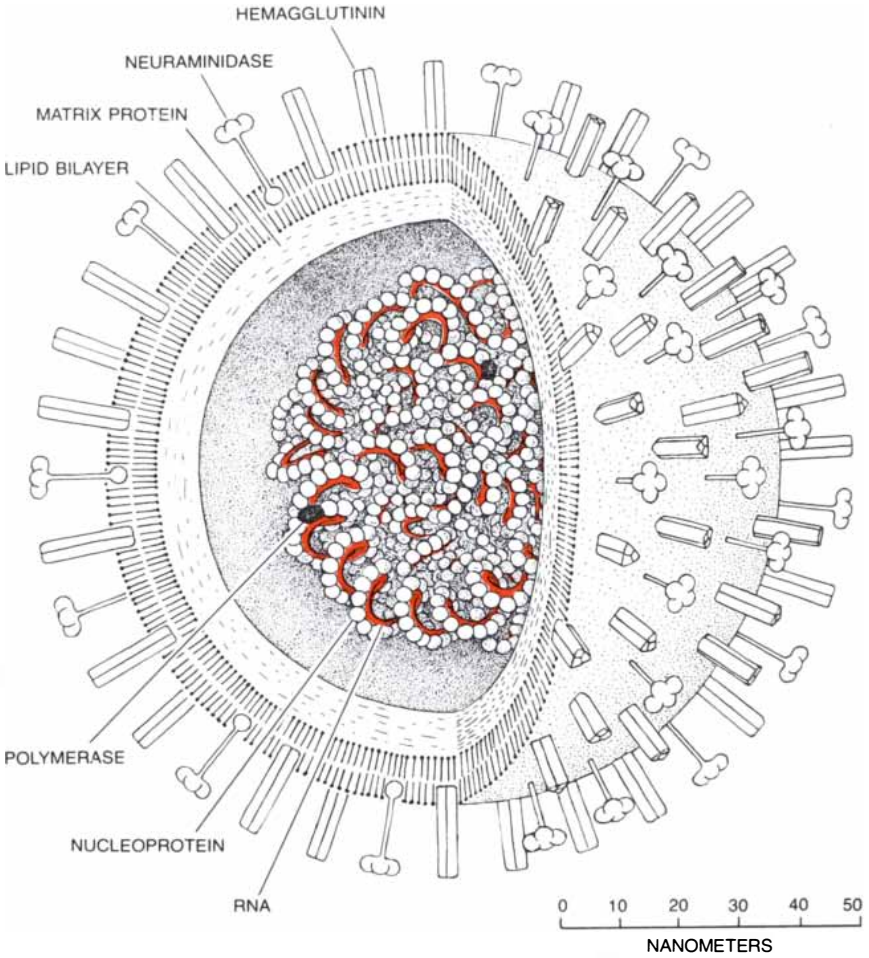
is covered with "spikes": densely arranged radial projections. These surface spikes are of two different kinds, one kind of spike that combines with red blood cells and causes them to agglutinate (hence the name hemagglutinin, usually abbreviated *H*) and another that dissolves the linkage between the hemagglutinin spike and the red blood cell. The second kind of spike is an enzyme, neuraminidase (*N*).

The hemagglutinin causes influenza viruses to attach themselves to cells. After infection antibodies to the *H* spike are formed, preventing reinfection with the same strain of influenza virus. Considerable antigenic variation is observed in the hemagglutinin antigen.

The *N* spike on the virion is completely different in appearance and function from the *H* spike. The neuraminidase enzyme may be responsible for getting the virus out of infected cells; it is known to facilitate the spread of the virus from cell to cell. Antibodies are also formed against the *N* enzyme after infection, but they are of less importance in providing protection from infection. The neuraminidase molecule also shows antigenic variation, although it is less variable than the hemagglutinin molecule. Thus both the *H* and the *N* spikes are highly variable proteins, and antibodies to these proteins distinguish different strains of influenza viruses.

Influenza viruses differ from most other animal viruses in that the RNA, or ribonucleic acid, that contains their genetic information is replicated and included in the virion as eight separate single-strand segments. This segmentation of the RNA means that genetic recombination, or reassortment, can occur readily during mixed infection with different influenza *A* strains. The recombination of RNA segments is probably of key importance in accounting for major antigenic variations of influenza viruses. Each RNA segment has been shown to include the genetic information required to code for a single virus protein; eight proteins are synthesized by the virus in the infected cells.

In the virion each of the eight RNA segments is associated with four other molecules: a nucleoprotein and three polymerase enzymes. The exact arrangement of these molecular complexes is not understood at present, but it is apparently responsible for the helical internal structures seen in the electron micrographs of some virions. The two major proteins in the core of the virion, the nucleoprotein and the matrix protein, are very similar in influenza viruses of the same type; in other words, they are type-specific for the influenza *A, B* and *C* viruses. Antibodies produced after infection do not distinguish between the nucleoproteins from different influenza *A* viruses, that is, between subtypes within the same type. The antibodies to these proteins are not thought to be important in providing protection from infection.



CUTAWAY DIAGRAM of an influenza virion reveals its remarkable external and internal construction. The *H* and *N* spikes are embedded in a lipid, or fatty-acid, bilayer that surrounds the core of the virus particle. Matrix protein molecules line the underside of the lipid bilayer and surround the core. Inside the core there is a helical complex of molecules, consisting of RNA (ribonucleic acid) in association with nucleoprotein and polymerases (enzymes that initiate replication). The RNA (*color*), which contains the virus's genetic information, is divided into eight segments of differing lengths. Antibodies to the ribonucleoproteins and the matrix proteins formed in the body after infection are specific for types *A, B* and *C* influenza virus, but they do not distinguish between subtypes, or strains, within each type and they are thought to be not important in protection against reinfection. After the virus is adsorbed by a cell the lipid in the virion bilayers fuses with the cell's lipid bilayer either at the surface of the cell or in small vesicles within the cell. Once the virus core has entered the cell, complementary copies of the eight RNA segments of the virus are made with the aid of the polymerase from the incoming particle. The complementary RNA in turn directs the synthesis of proteins and RNA for new virus particles. The actual sequence of events inside the infected cell, including the sites of viral replication and the function of the cell nucleus during replication, is poorly understood at present.

Influenza viruses display two kinds of antigenic variation in their main *H* and *N* antigens. The first kind of change, called antigenic drift, consists of a series of minor alterations within a group of similar strains. The second variation is called antigenic shift to connote an abrupt and major change in the composition of either the hemagglutinin or neuraminidase antigens (or both), which, by convention among virologists, are designated *H*0, *H*1, *H*2, *H*3, *N*1, *N*2 and so forth. Both kinds of variation are observed in influenza *A* viruses, but only antigenic drift has been detected in influenza *B* viruses. Influenza *C* virus is rarely isolated, and it has been much less studied. Influenza *A* virus has been isolated from many species of lower animals as well as from man and is the only type that causes pandemics in man. For this reason the remainder of this discussion deals with the Type *A* viruses.

Since 1933, when the first human influenza virus, now designated *H*0*N*1, was isolated, antigenic drift and antigenic shift have occurred in influenza *A* viruses. Minor changes (antigenic drift) have been seen in both of the surface-spike proteins, particularly in the hemagglutinin molecule, giving the mutant viruses a selective advantage in the otherwise immune host population. For example, a variant isolated in 1947 was designated *H*1*N*1, to signify a distinctive change in the *H* spike. Recent work,

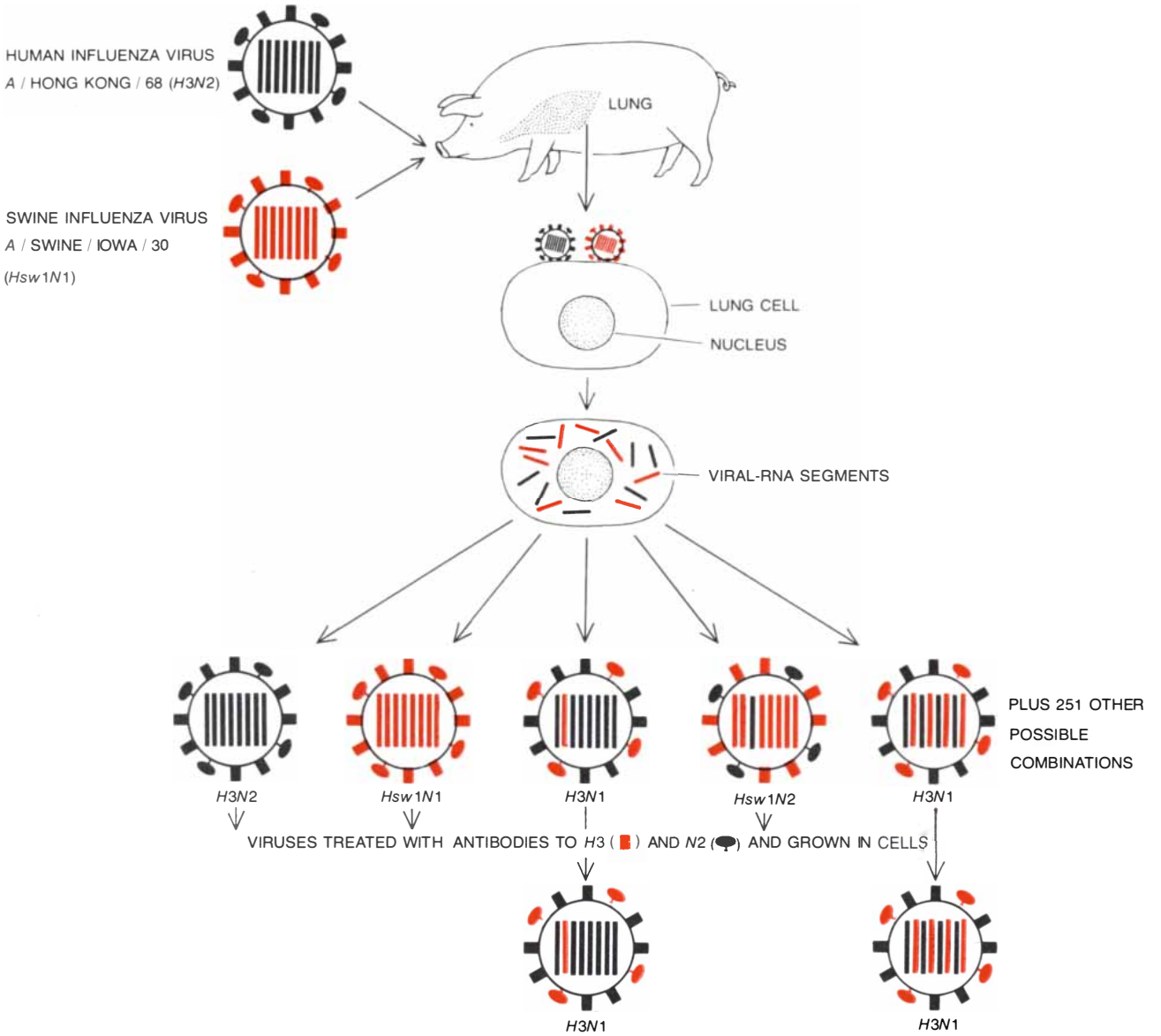
however, has shown that these particular viruses (*H0N1* and *H1N1*) belong to the same type.

The first major change (antigenic shift) detected after 1933 was recorded in 1957, when a new pandemic strain appeared. This virus was the first of the Asian strains, and both the *H* and the *N* surface antigens were completely unlike those of the earlier strains. The surface-spike antigens were designated *H2N2* to indicate that they were different from the preceding *H1N1* virus. The Asian virus underwent antigenic drift until 1968, when the Hong Kong subtype appeared. The Hong Kong virus possessed a new *H* spike, but the *N* spike was the same as that in the Asian strain; this virus was accordingly designated *H3N2*. The Hong Kong virus has since shown antigenic drift, producing the variants known as the England, Port Chalmers, Scotland and Victoria strains, each strain named after the geographic area in which it was first isolated.

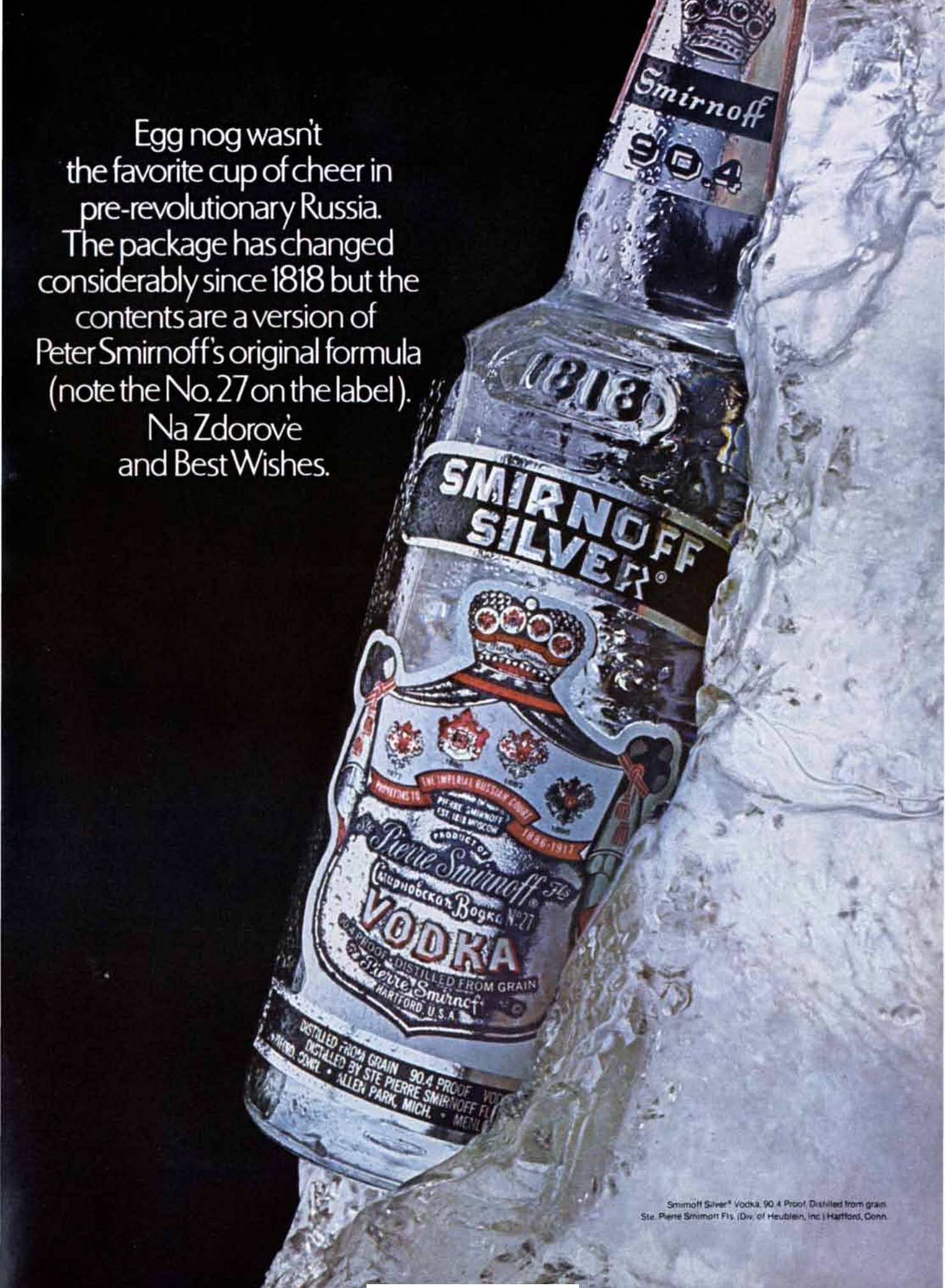
After the appearance of a new influenza virus in man the preceding strain disappears almost completely from the human population. Thus in 1977, while the Victoria variant of the Hong Kong, or *H3N2*, group was still circulating in man, the earlier groups of viruses from 1957–68 (designated *H2N2*) and from 1933–57 (*H0N1* and *H1N1*) could not be found in man.

Why Pandemics?

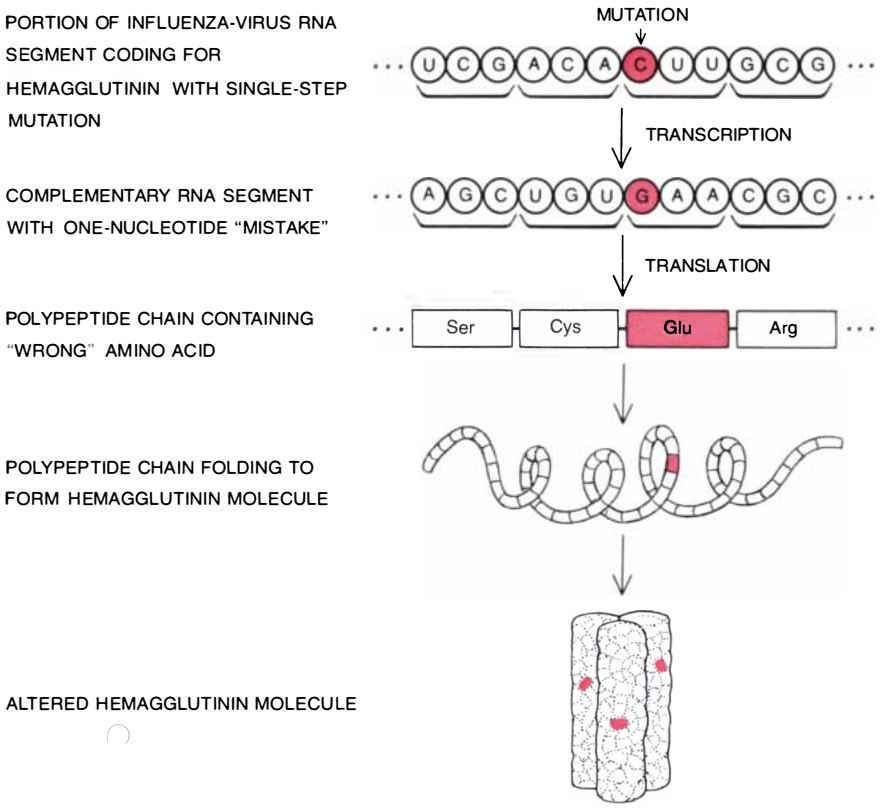
Since 1918 several pandemics have



RECOMBINATION of the genetic material from two different Type *A* influenza viruses is represented schematically in this diagram. (The various strains of the virus are labeled, by convention among virologists, according to the sequence in which major variations in the two main antigenic components of the virion, hemagglutinin and neuraminidase, first make their appearance. In man, for example, the variant *H* and *N* surface antigens are designated *H0*, *H1*, *H2*, *H3*, *N1*, *N2* and so forth. For strains that appeared first in other animals, such as swine, appropriate identifying names or letters, in this case *sw*, are added to the coded label.) Here the two influenza viruses, one from man (*H3N2*) and the other from swine (*Hsw1N1*), are shown being inoculated into the nose of a pig. The inoculation results in the simultaneous infection of a single lung cell with the eight separate RNA segments from each virus. Once inside the cell the viruses multiply, and the 16 different RNA segments can be recombined in many ways during the "packaging" of the new virus particles. In the presence of antibodies to the neuraminidase of the human strain and to the hemagglutinin of the swine strain all the resulting viruses will be neutralized except for those possessing the *H3N1* combination of surface antigens. The *H3N1* influenza virus may contain many different combinations of RNA's that include those coding for the two surface proteins, but their rearrangements with the remaining six segments can vary and some of the recombinants may cause infection in pigs. The diagram illustrates an actual experiment carried out by one of the authors (Webster) with Charles H. Campbell at the U.S. Department of Agriculture's isolation facility on Plum Island off Long Island. *H3N2* strain was first detected in human beings in Hong Kong in 1968; *Hsw1N1* strain was first isolated from swine in Iowa in 1930.



Smirnoff Silver® Vodka. 90.4 Proof. Distilled from grain.
Ste. Pierre Smirnoff Fils. (Div. of Heublein, Inc.) Hartford, Conn.



ANTIGENIC DRIFT, one of two kinds of antigenic variation displayed by influenza viruses in their main *H* and *N* antigens, consists of a series of minor genetic alterations within a group of similar strains. One such alteration is depicted here. The short portion of a viral-RNA segment shown at the top contains the genetic information that codes for the hemagglutinin molecule; the segment contains a single-step mutation (*color*). The RNA, represented simply as a string of lettered beads, is actually a chain of nucleotides each of which is made up of a ribose sugar, a phosphate group and one of four organic bases: adenine (*A*), uracil (*U*), guanine (*G*) and cytosine (*C*). The viral RNA is copied by the polymerase in the virion to yield a complementary RNA strand, *A* generally pairing with *U* and *G* with *C*. In this case, however, a "mistake" has been made: the mutant *C* pairs with a *G*. As a result the three-letter "triplet" coding for a specific amino acid inserts instead a "wrong" amino acid in the polypeptide chain of the *H* protein. The new amino acid may in turn interfere with the folding of the polypeptide chain, distorting the *H* molecule, or it may appear in the antigenic region of the molecule, producing a minor antigenic variant that cannot be neutralized by antibodies to the parent influenza virus.

occurred, but none has approached the 1918 pandemic in severity. Meanwhile localized epidemics of a relatively mild nature have continued to be interspersed between the pandemics. In epidemiological terms these events are characterized by waves of disease through the human population attributable to a lack of immunological protection conferred by preceding strains to which the population has been exposed.

The drifting strains within each antigenic subtype are closely related, although the range of drift (within a subtype such as *H*2 or *H*3, for example) can be fairly extensive. Hence a vaccine prepared from an early *H*3 strain (one that appeared, say, in 1968) does not protect as well as one prepared from a later strain (one that appeared, say, in 1975). For this reason viruses isolated from human influenza patients, particularly after an upsurge of clinical disease, are continually subjected to antigenic analysis so that vaccines incorporating the new strain can be prepared when needed.

A reasonable explanation for this antigenic drift is that it results from the immunological selection of a naturally occurring mutant that is not neutralized by antibodies to the earlier strain. The process has been referred to as immunological pressure in order to denote the neutralization by a specific antibody of identical virus particles, thereby favoring the selection of mutant forms. It is easy to demonstrate antigenic drift in the laboratory using immunological selection against the earlier strain. Burnet, Alick Isaacs (the discoverer, with Jean Lindemann, of the virus inhibitor known as interferon) and their co-workers first accomplished this in 1949 in chicken eggs containing developing embryos. They mixed a strain of influenza virus with its specific antiserum and then inoculated the mixture into the allantoic sac of the embryonated eggs. The antiserum neutralized all the virus except for a mutant form, which then became the predominant virus on further subculture in eggs.

In viral replication (and microbial replication in general) mutants appear at a frequency of about one in a million particles, depending on the viral strain and the conditions of culture. Such mutants of influenza virus are considered to be "point mutations" that might affect only one of the nucleotide building blocks of the RNA, with the result that a difference of a single amino acid building block in the *H* protein molecule results from the translation of the changed code of the RNA.

Antigenic shift, on the other hand, constitutes a much more radical change in the composition of the RNA molecule, probably involving the recombination of different segments of RNA, each of which represents a gene, or thousands of nucleotides in sequence. It is most unlikely that the phenomenon of antigenic shift occurs simply by mutation and selection from immunological pressure. It seems still less likely that a double shift in both major *H* and *N* antigens, as was seen in the 1957 Asiatic strain (from *H*1*N*1 to *H*2*N*2), occurred as a result of a mutation, or even recombination, of strains existing solely within the human population. The reasons for our doubting this hypothesis, still held by some virologists and epidemiologists, will become evident below.

Influenza in Animals

Until 1955 animals other than pigs and certain experimental laboratory animals were not known to be susceptible to infection by influenza virus. Beginning that year the recognition that the viruses causing influenza in horses, ducks and other animals belonged to the same group as the human influenza *A* virus spurred some interest in such infection. In July, 1957, soon after the onset of a new pandemic of Asian influenza and a vague report from China that concurrent outbreaks of influenza in pigs had been found, one of us (Kaplan), at that time a member of the staff of the World Health Organization (WHO), undertook to explore the intriguing possibility that the human strain might have emerged from an animal reservoir.

A worldwide survey of pigs and horses was begun under the auspices of the WHO to determine the status of influenza infection in these animals before and after the pandemic struck local animal populations. In 1958 Chu Chi-Ming of the National Vaccine and Serum Institute in Changchun described the epidemiological characteristics of the 1957 epidemic in China and placed its origin in early February, 1957, in the area of western Kweichow and eastern Yunnan. He also thought that the radically different character of the new 1957 strain with respect to the formerly prevalent strain suggested an animal reservoir as the origin of the new strain.

Since July, 1957, the WHO has stimulated and coordinated studies of animal influenza in different parts of the world

DC-10.
When you have a choice...

To many people, "all airplanes look alike," but more and more discerning travelers are seeing and choosing the McDonnell Douglas DC-10. More than 100,000 passengers board DC-10s each day.

In addition to its size, the most distinguishing characteristic of the DC-10 is the engine perched high in the tail, where it is superbly efficient.

DC-10s are nearly everywhere. Thirty-four airlines fly more than 4,900 flights a week to 150 cities on six continents. In fact, DC-10s fly to more places, more often, than any other wide-cabin jet.

When you fly, don't settle for less. Whether your journey is a short trip, or a long flight across an ocean, you'll find the DC-10 as gracious inside as it is graceful outside. Tell your airline or travel agent, "When I have a choice, I prefer the DC-10."

The DC-10
MCDONNELL DOUGLAS

EQUAL OPPORTUNITY IN PROFESSIONAL CAREERS. SEND RESUME: BOX 14526, ST. LOUIS, MO. 63178

This close.

This far.

Shoot every 1½ seconds.

All automatically.

No other camera in the world

Reproduced from original unretouched SX-70 photographs. ©1977 Polaroid Corporation "Polaroid" and "SX-70"®

Daytime with flash.

Nighttime without.

And in minutes...

Sharp, clear color.

does it all.

Polaroid's SX-70 Alpha 1.

The 1978
Never before has there been

LeBaron 2-Door

A personal car inside and out.

Should you be looking for a sleek, personal car this year, make Chrysler LeBaron your choice. It's the first of a totally new class of automobiles. LeBaron 2-door comes with many personal appointments of luxury. Overhead lighting. Velour seats. And soft leather seating is available on the Medallion model. In short, LeBaron gives you the pleasures of a luxury car at a price that makes it a pleasure to own.

LeBaron 4-Door

Luxurious and affordable.

The nice thing about LeBaron is the way it satisfies your taste for luxury but still has regard for your budget. LeBaron 4-door, for example, is elegant both inside and out. The luxurious LeBaron pictured at your right is equipped with a padded vinyl roof, available wire wheel covers, plush velour seats, and much more. In fact, it comes with many of the same comfort and convenience features you'd find on $12,000 sedans. But LeBaron can be yours for many thousands of dollars less.

LeBaron Town & Country

There are many new wagons, but there is still only one Chrysler.

If you have a family, or just like the feel of a wagon, this could very well be your choice for 1978. LeBaron Town & Country is a lighter, more efficient size luxury wagon. But don't let that fool you. It has 90 percent as much passenger space as last year's full-size Town & Country. And though LeBaron is smaller outside, it's still big on luxury inside.

Chryslers. such a luxury of choice.

Cordoba

The ultimate personal car.

Cordoba has been the choice of many discerning car buyers. In fact, it's the most successful Chrysler ever introduced. And that's quite an accomplishment. But then, Cordoba is quite an automobile. This year's Cordoba is newly styled to make it the ultimate personal luxury car. From its classic lines, to its impeccable luxury, to its affordable price, Cordoba has the utmost consideration for quality.

Newport

So much Chrysler for so little.

There are those of us who demand full-size comfort. Then, there are those of us who demand value. Somehow, Newport has managed to combine both. For 1978, Newport offers you standard comfort features like power steering and brakes, steel-belted radials, automatic transmission, rich foam cushions, even the Electronic Lean Burn engine. And the price? Well, let's put it this way. Newport is our most affordable full-size Chrysler.

New Yorker

The feel of luxury and a sense of value.

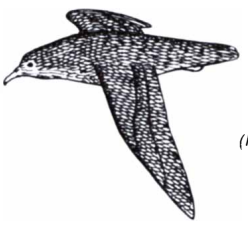
Traditionally, New Yorker's always had a knack for bringing pleasure to even the most fastidious car owner. This year's New Yorker may be the finest yet. It literally surrounds you in comfort. And the most surprising comfort feature of all, is New Yorker's very reasonable price.

But whatever type of luxury you want to Buy or Lease for 1978, the choice is up to you. At your Chrysler-Plymouth Dealer.

CHRYSLER
A PRODUCT OF CHRYSLER CORPORATION

MALLARD DUCK
(ANAS PLATYRHYNCHOS)

ARCTIC TERN
(STERNA PARADISAEA)



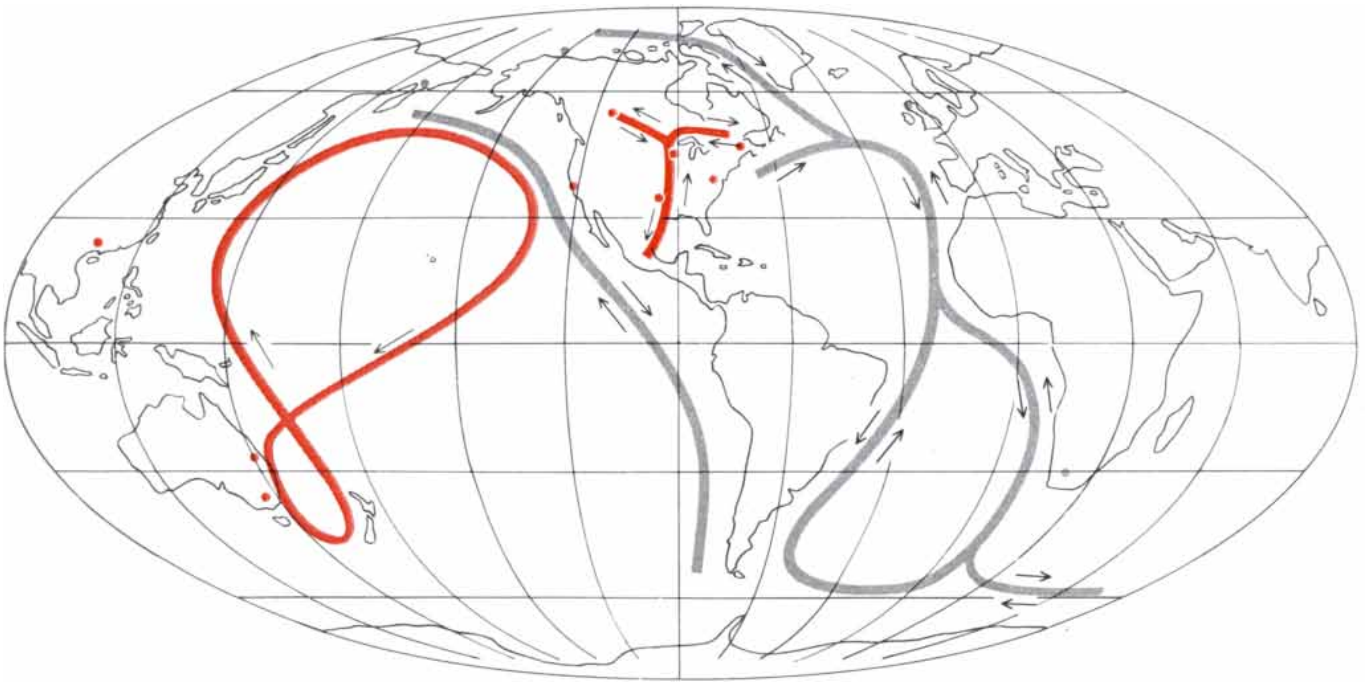
SHEARWATER
(PUFFINUS TENUIROSTRIS)

MIGRATING BIRDS may play an important role in the spread of new influenza virus strains among other species of animals. The migratory patterns of three such avian species known to have influenza viruses associated with them are indicated on this map of the world. Influenza viruses bearing the surface antigens designated *Hav*6N2 have been iso-

to throw some light on the role of animals in the epidemiology of human influenza. Much of the information presented here was obtained in these studies. The principal aims of these studies have been to determine (1) whether or not lower animals are of any importance as a prime source of the major antigenic components in the influenza *A* viruses that cause the recurring pandemics in man, and (2) if animals are important in these pandemics, whether it is because of the emergence of an animal strain with a slight change that can spread in man, because recombinations occur in nature between animal and human strains leading to emergence of a new strain, or because of both.

Influenza in Pigs

In 1941 Shope described a highly complex biological cycle of swine influenza he believed represented the natural course of the disease in swine. The cycle of infection involved the pig lungworm (*Metastrongylus apri*), the common earthworm (*Lumbricus terrestris*) and a precipitating factor such as sudden inclement weather, all combined with infection with the swine variety of Pfeiffer's bacillus. Shope's pig-lungworm-earthworm-pig cycle was later confirmed experimentally by other workers.



lated in recent years from mallard ducks (*Anas platyrhynchos*) in the Canadian province of Alberta and from the same species during its migrations in Wisconsin and Arkansas; an identical virus was isolated in 1965 from turkeys in Massachusetts, where it caused mild respiratory illness (*dark colored dots*). Similarly, influenza viruses labeled *Hav5Nav2*, which were isolated in 1961 from common terns (*Sterna hirundo*) in South Africa and from avian species in eastern U.S.S.R. and the islands off Alaska, are related to a virus (*Hav5N1*) found in 1959 in domestic chickens in Scotland (*gray dots*). The migratory species known to link this area is the arctic tern (*Sterna paradisaea*). Finally, an influenza virus isolated from shearwaters (*Puffinus pacificus*) on the Great Barrier Reef off Australia has been found in chickens in Hong Kong, in turkeys in California and in wild ducks in Delaware (*light colored dots*). A related species of shearwater (*Puffinus tenuirostris*) that migrates near these areas could transmit the viruses to shorebirds, which in turn may transmit them to other species.

An analogous hypothesis, that of a widely seeded virus in man awaiting a precipitating event, is still attractive to some epidemiologists.

Subsequent observations have shown that Shope's complex biological cycle is not necessary as an interim reservoir for swine influenza. In the mid-1960's Dionyz Blaškovič's group, including one of us (Kaplan), carried out a series of transmission experiments in pigs at the Czechoslovak Academy of Sciences' Institute of Virology in Bratislava. This group proved that pigs free of parasites could be infected and could transmit infection to pigs in contact with them for up to three months after experimental infection. In addition J. Mensik and his colleagues at the Institute of Veterinary Research in Brno demonstrated parasite-free sow-to-offspring transmission, which resulted in the death of the piglets. Bernard C. Easterday and his colleagues at the University of Wisconsin extended these findings in studies of the natural history of the disease in the Middle Western U.S. and other parts of the country.

It has now been established that naturally occurring swine influenza affects pigs in every state of the U.S. and that virus can be recovered from different pigs during all months of the year. Although the clinical manifestations of the disease are observed primarily in the fall and winter season in the north-central part of the U.S., virus is evidently circulating among herds of swine throughout the year. Swine influenza virus (labeled *Hsw1N1*) has been isolated from pigs in Britain, Czechoslovakia, Kenya, Poland, the U.S.S.R., Hong Kong, Italy and West Germany, but the clinical disease has been observed only sporadically in these countries.

Surveys of pigs in many countries before and after the 1968 pandemic of Hong Kong virus (*H3N2*) in man clearly revealed the passage of that strain of virus from man to swine. The *H3N2* strains that have drifted in man since 1968 have also been isolated from swine, the most recent of which (*A*/Victoria/75) was found in pigs in Hong Kong and in the U.S. in 1976. (The 1968 Hong Kong strain has also been isolated from dogs in the U.S. and the U.S.S.R., and it has caused disease in chickens on the Kamchatka peninsula of the U.S.S.R. and in calves in the central region of the U.S.S.R. The strain isolated from calves was shown to cause acute respiratory infection in experimental calves on Plum Island, the isolation facility of the U.S. Department of Agriculture off Long Island. For unknown reasons natural infections of chickens and calves with the Hong Kong strains of virus have not been observed elsewhere.)

The Fort Dix Incident

In January, 1976, the *Hsw1N1* swine strain, identical with those found in pigs in other parts of the U.S. in recent years, was isolated from a soldier who had died of influenza at Fort Dix. This strain was also recovered from five other cases of influenza among service personnel at Fort Dix, although at the same time influenza caused by the *A*/Victoria/75 strain was infecting troops in the camp. Serological investigations showed that the Fort Dix swine strain had infected some 500 personnel at Fort Dix, but apparently it did not spread further.

Previously human infection with *Hsw1N1* had been observed only in 1961 in a farm worker in Czechoslovakia and in 1974 in a Minnesota farm boy who died of Hodgkin's disease (which inhibits the immunological protective system, so that the swine virus was probably an incidental infection). Just before the Fort Dix episode, however, *Hsw1N1* infection in an eight-year-old boy and members of his family was noted in Wisconsin in the fall of 1975, coincident with an outbreak of swine influenza in pigs on the same farm. Since February, 1976, human infection with *Hsw1N1*

has been detected by isolating the virus from sick people on two farms in Wisconsin and one in Minnesota where pig infection was also present.

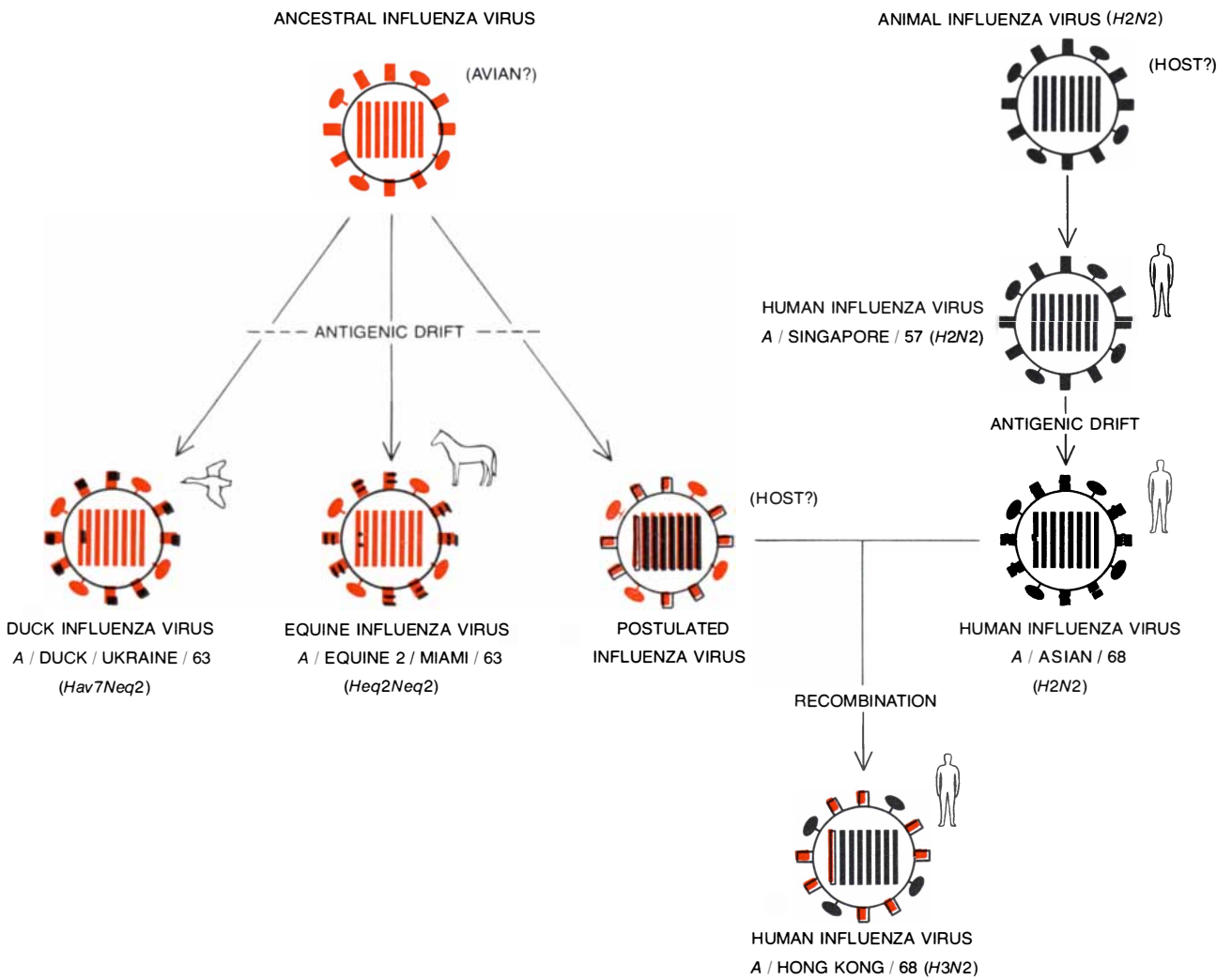
It is likely that pig-to-man transmission has occurred many times but was not detected by routine virological and epidemiological investigations. Fortunately these infections appear to have been self-limiting. In view of the relation of the *Hsw*1*N*1 virus to the presumed 1918 pandemic strain, however, the great concern of public-health authorities about the Fort Dix incident is hardly surprising. There was always the possibility that if the 1976 strain had acquired the capacity to spread in a human population along with the character of virulence, a pandemic like that of 1918 could result.

The question naturally arises of whether the 1918 pandemic strain emerged from pigs (or from some other animal reservoir) or whether it infected pigs for the first time in that year, as was maintained by veterinarians in the Middle West at the time. The question may never be answered, but the hypothesis of emergence from pig to man in 1918 is given credence by the Fort Dix epidemic and the other human infections cited above. It is also supported by the localized nature of swine influenza in the U.S., a situation that indicates the existence of a long-established, well-adapted host-parasite cycle in an animal population.

It is particularly intriguing that the first definitely observed cluster of human influenza cases in March, 1918, was at Fort Riley, and that this overcrowded camp drew many of its recruits from young farm workers of the Middle Western states. It is entirely possible that a similar event (the transmission of swine virus to man) took place in Fort Dix in 1976, the difference being that the 1918 virus somehow acquired the capacity to disseminate widely and, graver still, to become highly virulent, properties that cannot be described as yet at either the biological or the molecular level.

Influenza in Horses and Birds

Equine influenza occurs widely in many parts of the world, and it is particularly troublesome in racehorses and riding horses. Until the equine influenza virus was isolated in 1956 the term influenza was applied loosely to several diseases that we now know can be caused by different agents. Many outbreaks of "horse influenza" were described in the 19th century and were linked to human epidemics, but present knowledge indicates that the two equine influenza strains now prevalent, *Heq*1*Neq*1 and



PROPOSED ORIGIN of the 1968 Hong Kong strain of human influenza virus is outlined. According to the hypothesis advanced by the authors, an ancestral influenza virus (possibly avian) gave rise to a related series of animal viruses, including one virus isolated from horses in Miami in 1963 and another isolated from ducks in the Ukraine the same year. Five years later another virus in this family that was related to both the equine and the duck strains recombined with the then current Asian, or *H2N2*, strain of the human virus, giving rise to the Hong Kong, or *H3N2*, strain. Thus the new human strain acquired the hemagglutinin (*H3*) gene from the postulated intermediate animal strain and retained both the neuraminidase (*N2*) gene and the genes required to cause disease in man from the human Asian strain.

Heq2Neq2, are not highly infective for man. Both strains can cause disease concurrently in the same stable, although usually only one strain is encountered at a time.

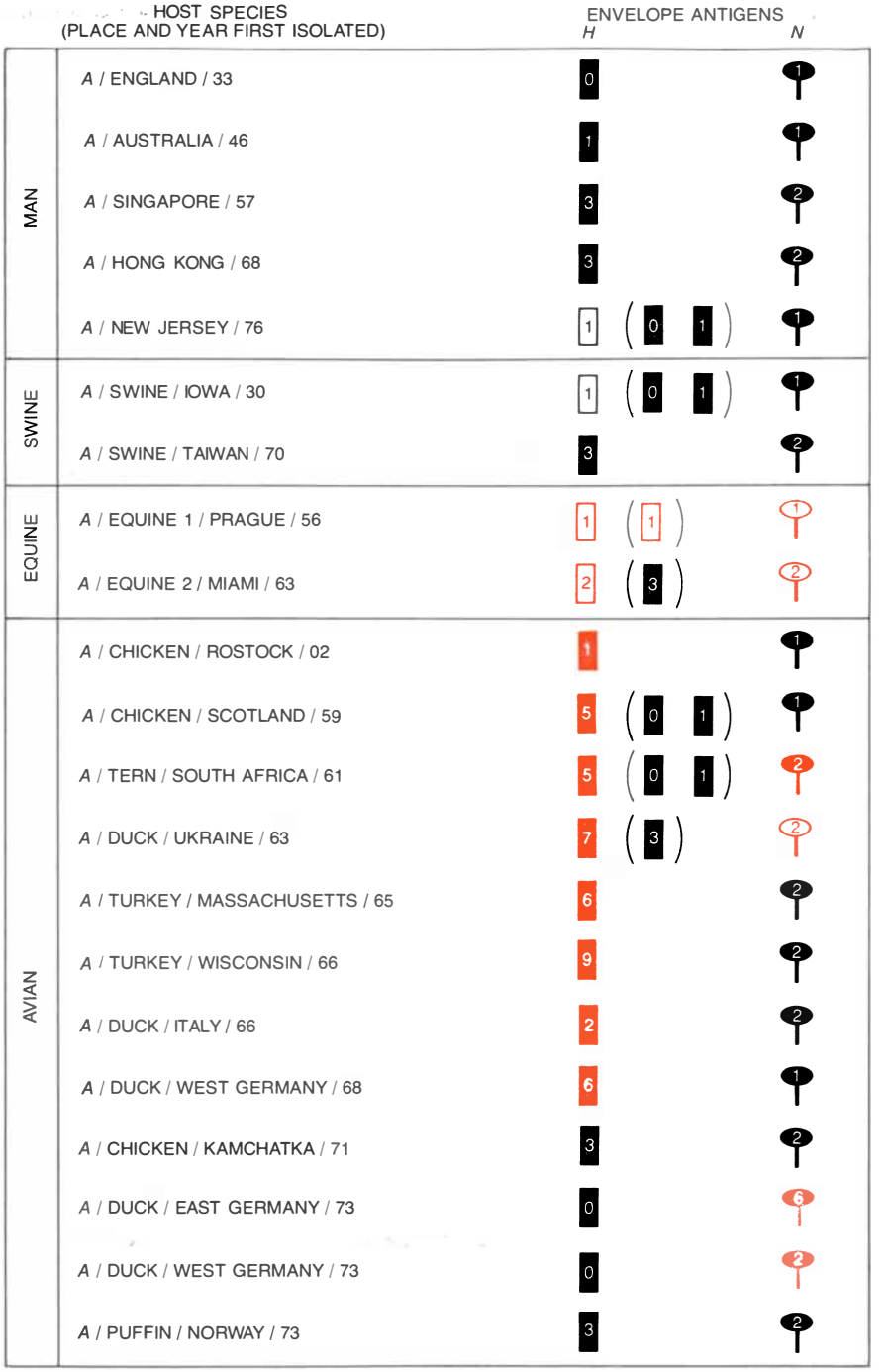
Of special interest are the results of analyses performed in recent years with the blood serum obtained from people alive in 1889. These analyses indicate that a strain containing the *H* antigen closely related to the Hong Kong antigen of 1968 (*H3*) and the *N* antigen from the second equine strain infected people at that time. This antigen, *Heq2*, was first identified in an influenza virus strain obtained from racehorses in an outbreak of the disease in Miami in 1963. It is also related to an avian *H* antigen (*Hav7*) isolated from a duck in the Ukraine in 1963, five years before the *H3* antigen appeared in the human pandemic of 1968.

Another curious fact is that in the 1960's *Heq2* antibodies were detected in the serum of Mongolian horses, an isolated animal population that has no contact with horses from other countries. Mystery still surrounds the question of how an apparently long-standing infection of horses in Mongolia could have reached Miami. A possible explanation is that it was carried by migratory birds.

Following the identification in 1955 of the virus causing fowl plague as belonging to the human influenza *A* type, hundreds of influenza viruses have been isolated from a wide variety of domestic and wild bird species. These have been grouped into nine different *H* subtypes (*Hav*) and six *N* subtypes (*Nav*). Some of these include the major *H* and *N* antigens related to human strains.

Avian influenza ranges from highly lethal infections in chickens and turkeys to inapparent infections in these and other avian species that harbor the same strains. Domestic ducks and quail often manifest influenza infection by coughing, sneezing and swelling around the beak, with variable mortality. Wildlife species and most domestic fowl show little or no signs of disease.

The wide dispersal of influenza infection in some species of migratory birds is particularly noteworthy. For example, Easter day and his colleagues have sampled various migratory duck species along the upper Mississippi flyway and have found virus in up to 30 percent of the specimens taken. Similar studies of juvenile mallard ducks by Virginia Hawkins and her colleagues in Canada have found a 25 percent incidence of influenza virus in samples from the cloaca of the birds. Influenza virus has been isolated from puffins by Geoffrey Schild's British team working on the islands off Norway, and by Graeme Laver of Australia from shearwaters on the Great Barrier Reef. D. K. Lvov and his colleagues of the D. I. Ivanovsky Institute of Virology in Moscow have isolated numerous strains from a large variety



RELATIONS between strains of influenza *A* virus isolated from various species, including man, have been established by blood-serum analyses performed in conjunction with other tests designed to detect antibodies to the major antigenic proteins of the influenza virion: hemagglutinin, neuraminidase and ribonucleoprotein. The influenza virus isolated from man in 1957 (*A*/Singapore/57), for example, possesses a hemagglutinin (*H*) antigen similar to an influenza virus isolated from ducks in East Germany in 1973, whereas the neuraminidase (*N*) on the influenza virus isolated from man in 1957 is similar to the neuraminidase on a virus isolated from turkeys in Massachusetts in 1965. The table shows that each of the antigens on the influenza viruses from man can be found on influenza viruses from lower animals. In recent years biochemical analyses of separated proteins and RNA segments have served to improve considerably the accuracy of such comparisons between components derived from different viruses. The accuracy of the comparisons is being further enhanced by ongoing work on the amino acid structure of certain molecules, particularly hemagglutinin. This table, which is based on serological analyses of the *H* and *N* antigens, shows that the *H* antigen is more variable and displays much greater antigenic drift than the *N* antigen. For example, the *H* antigen of the 1968 Hong Kong series of human strain (*A*/Hong Kong/68) and that of the 1930 Iowa swine strain (*A*/swine/Iowa/30) have drifted markedly, whereas their *N* antigens have been stabler. The *A*/New Jersey/76 strain, isolated from soldiers at Fort Dix in New Jersey early last year and subsequently found in farm workers in several other states, is identical with the many strains isolated from swine throughout the U.S. during the same period. The swine strains have evidently drifted from the *A*/swine/Iowa/30 strain. The numbered antigen symbols indicate identities or near-identities. (Those in parentheses represent fairly close relations.)

of wild birds during several expeditions to many parts of the U.S.S.R.

Migrating wild birds are probably responsible for the spread of avian influenza viruses in the world, and they may also act as a source of the viruses that give rise to the sporadic outbreaks of influenza infection in domestic poultry. For example, an influenza virus isolated from chickens in Scotland in 1959 is related through its surface *H* antigen to a virus found in common terns (*Sterna hirundo*) in South Africa in 1961. Antibodies to the South African tern virus were then detected in wild birds on the Pacific coast of the U.S.S.R. in 1971, and recently Easterday has isolated the same virus from birds caught in Alaska. Several avian species may be involved in the more localized spread of this virus, but the arctic tern is the only migrating bird known to link these areas.

Similarly, an influenza virus isolated from the shearwaters (*Puffinus pacificus*) on the Great Barrier Reef off Australia has been isolated from domestic turkeys in California (where it caused overt respiratory disease), from wild black ducks on the Atlantic flyway in Delaware and from domestic chickens in Hong Kong. The only bird known to migrate through these areas is the short-tailed shearwater (*Puffinus tenuirostris*). This species is exclusively oceanic, however, so that if it is involved in the transmission of influenza viruses, it is through contact with other birds, which then carry the viruses inland.

Another example of the transmission of influenza viruses by migrating birds can be seen in North American mallard ducks. An influenza virus found in domestic turkeys with respiratory disease in Massachusetts in 1965 has since been isolated from migratory mallards in Canada, Wisconsin and Arkansas. In the mallard duck no sign of the disease is caused by this virus.

The possibility that influenza viruses are transmitted between wild birds and domestic birds and eventually to mammals including man seems at first to be rather unlikely, particularly if the transfer is only by the respiratory route. Recent studies by one of us (Webster), however, have shown that many of the influenza viruses of ducks multiply in the cells lining the intestinal tract and are shed in high concentrations into water. These viruses will remain viable for at least four weeks in water at a low temperature (four degrees Celsius) and for five days in water at room temperature (20 degrees C.). It is therefore likely that influenza is a waterborne infection, at least among avian species, facilitating the transmission from wild ducks to domestic animals, and perhaps even to man through untreated water. Until now, however, only one definite case of human infection with an avian influenza-virus strain (fowl plague) has been reported.

It seems reasonable to conclude that from an evolutionary point of view the family of influenza viruses originated in the bird kingdom, an animal group some 100 million years old that now appears to live in reasonable harmony with its influenza viruses.

Origin of Human Pandemic Strains

There are at present at least three possible explanations for the origin of the pandemic strains of influenza viruses in human beings: (1) A human influenza virus recombines with an influenza virus from lower animals and acquires a completely new surface protein while retaining the capacity to cause disease in man; (2) an influenza virus of lower animals is transmitted to man and acquires the capacity to cause disease in man, and (3) the existing strains in man undergo multiple rapid changes by mutation to give rise to surface antigens that are unlike the earlier strains.

We believe there is good evidence for the first two possibilities but little or no evidence for the third. The most convincing evidence for the role of recombination in the origin of a new human pandemic strain of influenza comes from studies by Laver and one of us (Webster) of the Hong Kong strain that was first detected in southeast China in 1968. As we have noted, the Hong Kong strain acquired a new *H* surface spike that was designated *H*3, but it retained the earlier *N*2 antigen and the capacity to cause disease in man. Immunological and biochemical studies suggest that the amino acid sequence of the polypeptide chains making up the Hong Kong (1968) *H* spike is quite different from the sequence of the chains making up the spike in the Asian (1957) family of influenza viruses. These studies also indicated that the *H* molecule on the Hong Kong strain is closely related to the *H* molecule on influenza viruses isolated in 1963 from ducks in the Ukraine and to the *H* molecule on the equine virus (*Heq2Neq2*) isolated in the same year from horses in Miami. This evidence, taken together with the experimental demonstration that two different influenza *A* viruses can recombine when they are simultaneously introduced into the same animal, provides support for the first possibility given above, that of a human influenza virus recombining with an influenza virus from lower animals and acquiring a completely new surface protein. It should be stressed, however, that the evidence is circumstantial, since the Hong Kong strain itself was not isolated from animals before it was isolated from man. A comparison of the other viral proteins from the human, duck and horse strains of influenza virus and studies of the RNA sequence of these viruses are required before we can be sure of the origin of these viruses. Work along these lines is

SCIENCE/SCOPE

A new short-range missile seeker technology -- based on an M-band frequency of 94 GHz -- has been shown to penetrate adverse weather better than electro-optical or infrared seekers. It also delivers better resolution than does conventional radar. A prototype 94 GHz seeker, developed by Hughes under joint Air Force-Navy funding, has undergone laboratory and tower testing in active and passive modes against tanks and trucks. Helicopter captive flight tests have been completed at the Naval Weapons Center, China Lake, California.

Results indicate that the M-band provides a better match of resolution and penetrating characteristics for use in fog, rain, heavy clouds, battlefield smoke or dust than any other portion of the spectrum. Though still developmental, the 94 GHz seeker is projected as a small, relatively inexpensive, terminal guidance unit for short-range missiles, guided projectiles or longer range weapons equipped with a mid-course guidance system.

Negotiations are underway for two additional satellite programs for meteorological service to the U.S. and Japanese Space Agencies. NASA has selected Hughes to build, test and deliver three Geostationary Operational Environmental Satellites (GOES-D, E and F). The heart of the spacecraft in this series will be a Visible Infrared Spin-Scan Radiometer Atmospheric Sounder (VAS) to produce day and night cloud cover photos of an Earth-portion from a geostationary orbit plus vertical temperature soundings of various atmospheric levels. The instrumentation payload also will contain a Space Environment Monitoring system consisting of three separate sensors designed to monitor solar emission activities.

Hughes Industrial Products Division in Southern California has opportunities for Electronic Engineers, Physicists, Technicians and Glassblowers with experience in CRT's, laser applications, microelectronics, infrared detectors, silicon crystal growth, industrial automation, image and display products, and liquid crystal displays. Please send resume to: Hughes Aircraft Company, Professional Employment, 6155 El Camino Real, Carlsbad, CA 92008.

New laser instrument application with forty-millionths-of-an-inch accuracy foretells a revolution in manufacturing standards. Such unprecedented precision was produced in Hughes radar fabrication shop through a linkup of a laser interferometer (used to measure distances) with a computer-controlled Ekstrom Carlson machine, designed to contour and profile aircraft radar antenna quads.

Automatically compensating for changes in temperature and air pressure, the interferometer lets a machinist position a part within one/ten-thousandths of an inch -- double the machine's original accuracy. In this unique application, accurate measurements to 0.000040-inch have been recorded. Changes in height, depth and width can be monitored and corrected to the virtual elimination of error. Quality control inspectors can cross-check findings against their standard measurement techniques. And the increased accuracy is seen as reducing the need for total inspection of parts after fabrication and the number of rework parts.

now being actively pursued in several laboratories.

Evidence for the second possibility, that of direct transmission of influenza viruses from lower animals to man, comes from the recent isolation of swine influenza viruses from man in the Fort Dix episode and from pigs and man in three instances on the same farms in Wisconsin and Minnesota between 1974 and 1976. These latter incidents provide the first unequivocal evidence of the direct transmission of animal influenza viruses to man. It is also possible that the Asian strain ($H2N2$) in 1957 arose from animals; so far, however, the $H2$ and $N2$ antigens have been found only on different animal influenza viruses isolated after 1957. This fact, of course, does not rule out the possibility of the direct transmission of a virus from animals to man at that time, because studies of influenza in animals were initiated by the WHO only after 1957. As for the third possibility, that of multiple rapid changes by mutation of a human strain to produce a virus that is radically different in its surface spikes, evidence is completely lacking at this time, but the hypothesis cannot be entirely ruled out as a possible explanation.

Anticipating Future Pandemics

Thus the available information indicates that antigenic shifts giving rise to pandemic strains of influenza virus may result either from recombination or from direct transmission of the virus from animals. The ultimate criterion for an influenza virus's being able to cause a pandemic is that it must have the dual capacity to pass easily from human being to human being and to cause disease. Pandemics are most easily explained by the phenomenon of recombination between an influenza virus from (or within) an animal reservoir and a virus already in man, with the acquisition of new surface-spike proteins and the retention of the genes responsible for spreading and producing disease. At this time we do not know which of the eight gene combinations of the influenza virus specify transmissibility. Virulence is almost certainly polygenic (requiring two or more RNA segments). Hence the swine influenza viruses that appeared in man in 1976 apparently lacked the properties for becoming a pandemic strain in man, but we cannot yet explain why at the gene level.

Vaccines against influenza are at present only partially effective in preventing the disease in human beings and do not limit the broad spread of influenza in epidemics and pandemics. This situation is in marked contrast to that of smallpox, which has a very special set of epidemiological characteristics, such as the absence of significant animal reservoirs and the existence of long-term immunity through natural or vaccinal infection. These characteristics make it relatively easy to prevent the extensive spread of the disease.

Even if the efficacy of the present inactivated influenza vaccines were to be vastly improved, drawbacks such as laborious and costly production processes and problems of mass application within a short period (exemplified by the recent experience with swine-influenza vaccine in the U.S.) would necessarily restrict use of the vaccines to a few privileged countries. Research is under way in a number of countries, including the U.S., the U.K., China and the U.S.S.R., to develop modified live-virus vaccines that will be effective, safe and genetically stable. The indications are that such a vaccine would also be cheap and relatively easy to produce and administer. The process of achieving such a combination of properties, however, will be an arduous one, particularly with regard to safety and efficacy.

Where the antigenic drift of a virus strain has been marked, and in all cases of antigenic shift, it is essential to incorporate in the vaccine any new antigens that have appeared. This calls for highly specialized and time-consuming procedures, with the result that several months must necessarily elapse before vaccines can be made available in quantity. It also calls for an efficient global surveillance program to monitor the emergence of new strains, an activity now being coordinated by the WHO.

The present mode of vaccine production involves recombining the major H and N antigens with strains that give high yields in chicken embryos; the virus harvest is then inactivated. A wise course for the future would be to prepare and store under refrigeration recombinants of high-yielding strains with all the new H and N antigens detected in man or animals. Then if a new strain were to become epidemic in character, vaccines could quickly be produced from the "banked" recombinants. This approach is being taken in some countries, but it is complicated by the continual discovery of new antigenic types of influenza virus in animals, indicating the existence of a large pool of virus strains.

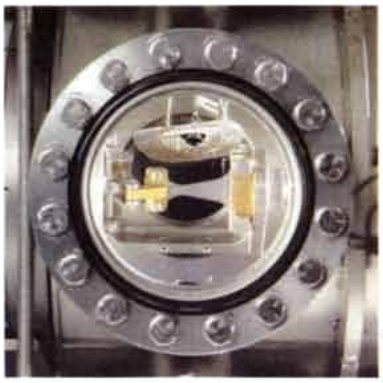
It is impossible to predict the dimensions of the continuing threat represented by the vast and ineradicable reservoir of influenza viruses that exists in domestic animal species and birds. What is most needed now is to expand our knowledge in order to determine whether there are any natural restrictions on gene combinations, in other words to determine whether there are only some strains of influenza virus that can predictably cause disease in man. It is certain that influenza pandemics will occur in the future, but in view of our advancing understanding of the disease and its ecology it seems unlikely that these new plagues will be as devastating as those of the past.

Forecasting fatigue life... in minutes.

Like the human body, metal too gives early warning signs of failure. One of these, oxide cracking, appears surprisingly early.

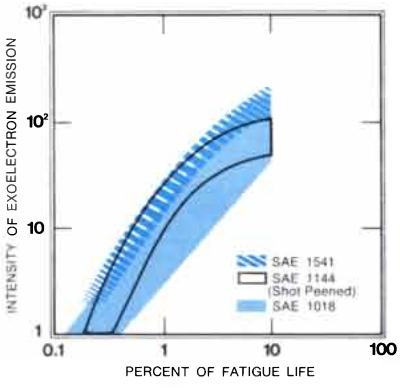
Here at the General Motors Research Laboratories, our physicists are detecting oxide cracks by exoelectron emission measurements. As a result, they have been able to predict—after testing a sample for less than 1% of its fatigue life—when the metal will fail.

How did the basic technique evolve? In exploratory studies, ultraviolet light was used to stimulate exoelectron emission from metal samples being fatigued in a vacuum. Intensity measurements and observations in a photoelectron microscope led to these discoveries:



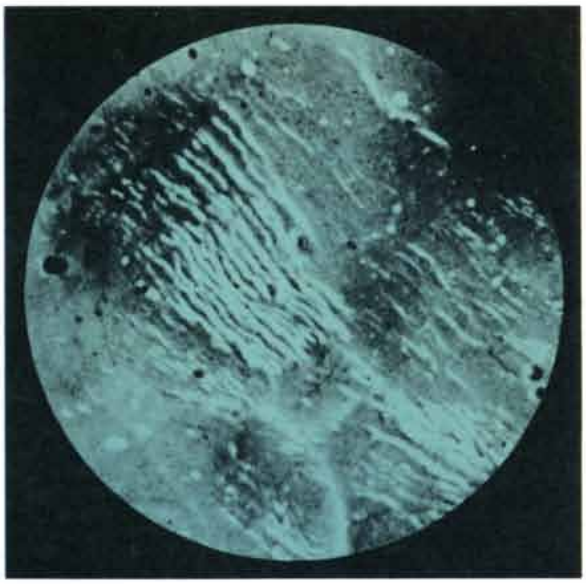
- Exoelectrons flow only from bare metal areas where microcracks have developed in the natural oxide coat, such as the cluster of streaks at right (magnification: 250X).
- The oxide cracks are caused by fatigue damage in the underlying metal.
- Exoelectron emission rate correlates with percent of fatigue life (the graph shows the growth of emission intensities for three types of steel fatigued under different loading conditions).

This correlation is the heart of the prediction. It enables the fatigue life of samples to be forecast in minutes versus weeks for accelerated life tests of components.



What role will the new technique play when perfected? It probably won't replace conventional life testing. But it will serve to establish component designs earlier in the game. And that means a healthy cut in development time.

Predicting fatigue life by exoelectron emission: A scientific approach to a very old engineering problem.



General Motors Research Laboratories
Warren, Michigan 48090

The Retinex Theory of Color Vision

A retina-and-cortex system (retinex) may treat a color as a code for a three-part report from the retina, independent of the flux of radiant energy but correlated with the reflectance of objects

by Edwin H. Land

The scientific tradition of simplifying the conditions of an experiment has left us until recently without a satisfactory explanation of how the eye sees color in everyday life. Paradoxically the modern technology of color photography has reinforced the belief that the colors discerned by Newton in the spectrum are, with minor qualifications, the colors of the world around us. We know, for example, that if we use daylight color film when we take a picture in the light shed by an ordinary tungsten-filament lamp, the picture will turn out to have a strong reddish cast. That, we say, is because the rays from the tungsten filament are too "red," never asking how we ourselves can move constantly in and out of tungsten-lit worlds without experiencing any change in the color of familiar objects: apples, lemons, strawberries, bread, human faces (the tones of which are so hard to get right on a television screen).

How, then, does the eye deal with the excess of "red" in a tungsten-lit room? As I hope to demonstrate in this article, the eye, in determining color, never perceives the extra red because it does not depend on the flux of radiant energy reaching it. The eye has evolved to see the world in unchanging colors, regardless of always unpredictable, shifting and uneven illumination. How the eye achieves this remarkable feat has fascinated me for many years.

In 1959 I described in these pages a series of experiments in which a scene created by the superposition of two black-and-white transparencies, one projected through a red filter and the other projected without a filter (that is, in white light), conveys to the eye nearly the gamut of colors present in the original scene [see "Experiments in Color Vision," by Edwin H. Land; SCIENTIFIC AMERICAN, May, 1959]. To produce such "red-and-white" images the picture projected through the red filter is taken through a red filter and the picture projected in white light is taken through a green filter. It would be expected that the superposed image on the projection screen could generate only red, white and various shades of pink. Actually one sees a picture remarkably similar to the full-color photograph reproduced on the opposite page. In the red-and-white photographic projection peppers are green, radishes and strawberries are red, the orange is orange, the lemon and bananas are pale yellow, the wood cutting board and knife handle are brown and the design on the plate is blue.

The challenge presented by our early red-and-white experiments led us step by step over a 20-year period to an explanation of how the visual system is able to extract reliable color information from the world around us, a world in which virtually every scene is lighted unevenly, in which the spectral composition of the radiation falling on a scene can vary enormously and in which illumination as brief as a lightning flash suffices for the accurate identification of color. If the nature of the responses of the photoreceptors in the retina of the eye even approximated what most of us were taught in school, functioning primarily as intensity-level meters with peaks in three different parts of the spectrum, we would be continually confusing one color with another. An object that looked yellow in one part of our field of view might look green or gray or even red when moved to a different part of the field. The fact remains that objects retain their color identity under a great variety of lighting conditions. This constancy is not a minor second-order effect but is so fundamental as to call for a new description of how we see color.

The visual pigments are photosensitive molecules that respond to a wide band of light frequencies. The three pigments in the cone cells of the retina cover the visible spectrum in three broad, overlapping curves. The pigment with a peak sensitivity at a wavelength of 440 nanometers responds in some degree to the entire lower-frequency half of the visible spectrum. Each of the other two pigments responds to almost two-thirds of the visible spectrum, the two being offset at their peaks by barely 30 nanometers, with their peak sensitivities located at 535 and 565 nanometers [*see upper illustration on page 110*].

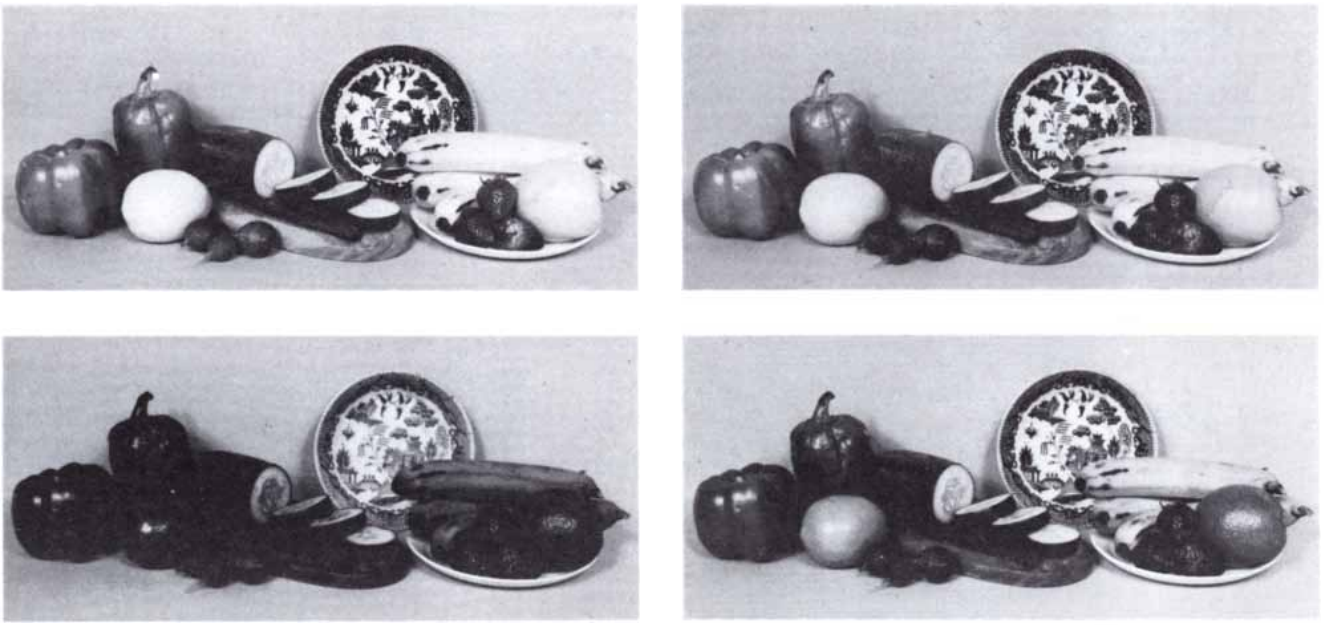
In this discussion the names of colors—"red," "green," "blue" and so on—will be reserved for the color sensation we have when we look at the world around us. In short, only our eyes can categorize the color of objects; spectrophotometers cannot. This point is not a trivial one because many people viewing some of our experiments for the first time will identify something as being red or green but will then ask, as if their eyes were being fooled, "What color is it really?" The answer is that the eye is not being fooled. It is functioning exactly as it must with involuntary reliability to see constant colors in a world illuminated by shifting and unpredictable fluxes of radiant energy.

Since I believe the study of color in fully colored images is best begun by examining images that are completely devoid of and completely uncomplicated by the experience of color, let me describe that experience in some detail. The hypersensitive system based on the rod cells in the retina functions at light levels as much as 1,000 times weaker than the systems based on the cone cells do, so that it is possible to answer the interesting question: What colors will one see if only the rod system is activated? One procedure is to put on a pair of tightly fitting goggles equipped with neutral-density filters that reduce the incident light by a factor of 30,000. After one has worn the goggles for about half

an hour objects in a room illuminated to the typical level of 20 foot-candles will become visible. The effective illumination in the room will thus be 1/1,500 foot-candle. As one looks around the room the familiar colored objects will be seen devoid of color, exhibiting instead a range of lightnesses from white to black, much as they would appear in a black-and-white photograph taken through a green color-separation filter. In other words, the reds will appear very dark, the greens lighter, the blues dark, the whites light and the blacks very dark.

In this colorless world one finds that the nature of the image is not determined by the flux of radiant energy reaching the eye. The illumination can easily be arranged so that there is more flux from a region that continues to look very dark than there is from a region that continues to look very light, whether these regions are three-dimensional objects or artifacts contrived with a montage of dark and light pieces of paper. The paradox immediately arises

STILL LIFE was used to make the four black-and-white images presented below. The reproduction of the still life above was made by conventional processes of color photography and photoengraving to show the reader what the colors of the original objects in the scene were. The black-and-white images were made with film-filter combinations that closely duplicate the separate wavelength sensitivities of the four systems of photoreceptors in the retina of the eye: the three systems of cone cells and the hypersensitive system of rod cells.



BLACK-AND-WHITE IMAGES OF STILL LIFE were taken with four different film-filter combinations, creating what the author calls retinex records. The picture at the top left was taken with a film whose spectral response was altered so that it matched the spectral sensitivity of the long-wave cone pigments in the eye. This photograph enables the observer to see a colorless image that approximates the image produced by the long-wave cones by themselves. The picture at the top right shows the same scene as it would be viewed by the middle-wave cone pigment. The picture at the bottom left is the scene as it would be viewed by the short-wave cone pigment. The picture at bottom right corresponds to the image seen by the rods. Unlike cone images, which cannot be viewed independently, images produced by the rod pigment can be studied in isolation at very low light levels, without interference from much less sensitive cone systems.

that each of the objects, the pieces of paper for example, whether dark or light or in between, maintains its lightness without significant change as it is moved around the room into regions of higher or lower flux. Light papers will be seen as being light and dark papers simultaneously as being dark, even with the same flux coming from each of them to the eye. Strong gradients of flux across the field will be apparent only weakly, if at all.

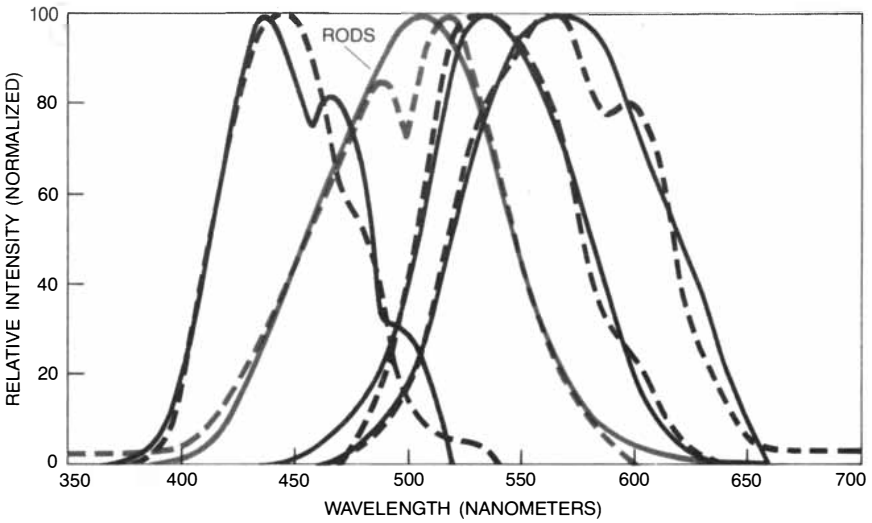
Furthermore, in an intricate collage of areas of various lightnesses, sizes and shapes, the lightness of a given element does not change visibly as it is relocated in any part of the collage and associated with a new arbitrary surround. When a small area is totally surrounded by a large area, the lightness of the small area will change somewhat depending on whether the large area is darker or lighter than the small one. In general, however, the impressive fact is that the lightness of a given area is not appreciably modified by the immediately surrounding areas, nor is it modified by the still larger areas surrounding them.

Although I have been describing a colorless world as it is seen by the hypersensitive receptors of rod vision, all the observations about the stability of lightness values can readily be reproduced with a montage of white, black and gray papers viewed at ordinary light levels. If, for example, a square of matte-surface black paper or, better still, black velvet is placed at one side of such a montage and a square of white paper is placed at the opposite side several feet away, with an assortment of light and dark papers scattered in between, one can place a strong light source close enough to the black square so that it sends more radiant energy to the eye than the white square, remote from the light; yet the black square will continue to look black and the white square white. In fact, with the montage still strongly illuminated from one side either the black square or the white one can be moved to any other part of the montage without a significant change in its appearance.

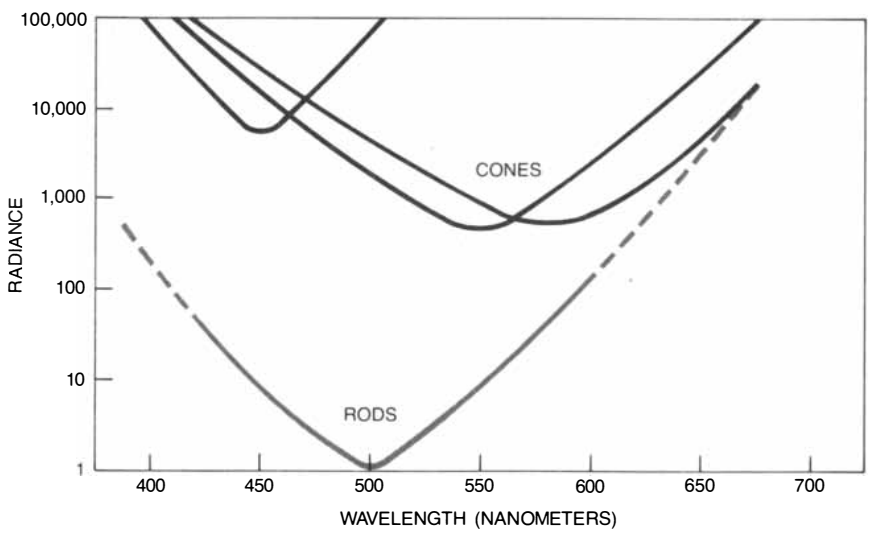
This remarkable ability of the eye to discover lightness values independent of flux, so convincingly demonstrated when only a single photoreceptor system is operating, is the rock on which a satisfactory description of color vision can be built. The first response of the visual system is for the receptors to absorb the light falling on the retina. Whereas the initial signal produced in the outer segment of the receptor cell is apparently proportional to the light flux absorbed by the visual pigment, the final comprehensive response of the visual system is "lightness," which shows little or no relation to the light flux absorbed by the visual pigment.

The processing of fluxes to generate lightnesses could occur in the retina, or in the cerebral cortex, or partially in both. Since we are uncertain of the location of the mechanisms that mediate these processes, I have coined the term retinex (a combination of retina and cortex) to describe the ensemble of biological mechanisms that convert flux into a pattern of lightnesses. I shall therefore use the term throughout this article in referring to these biological mechanisms. I shall also reserve the term lightness to mean the sensation produced by a biological system. Although the rods can be stimulated at light intensities below the cone threshold, the cones cannot be stimulated without exciting the rods. For cones we must study the lightness images produced by each individual set of receptors using retinex photography, as I shall explain below, or learn the properties of lightness images from model calculations based on spectroradiometric measurements.

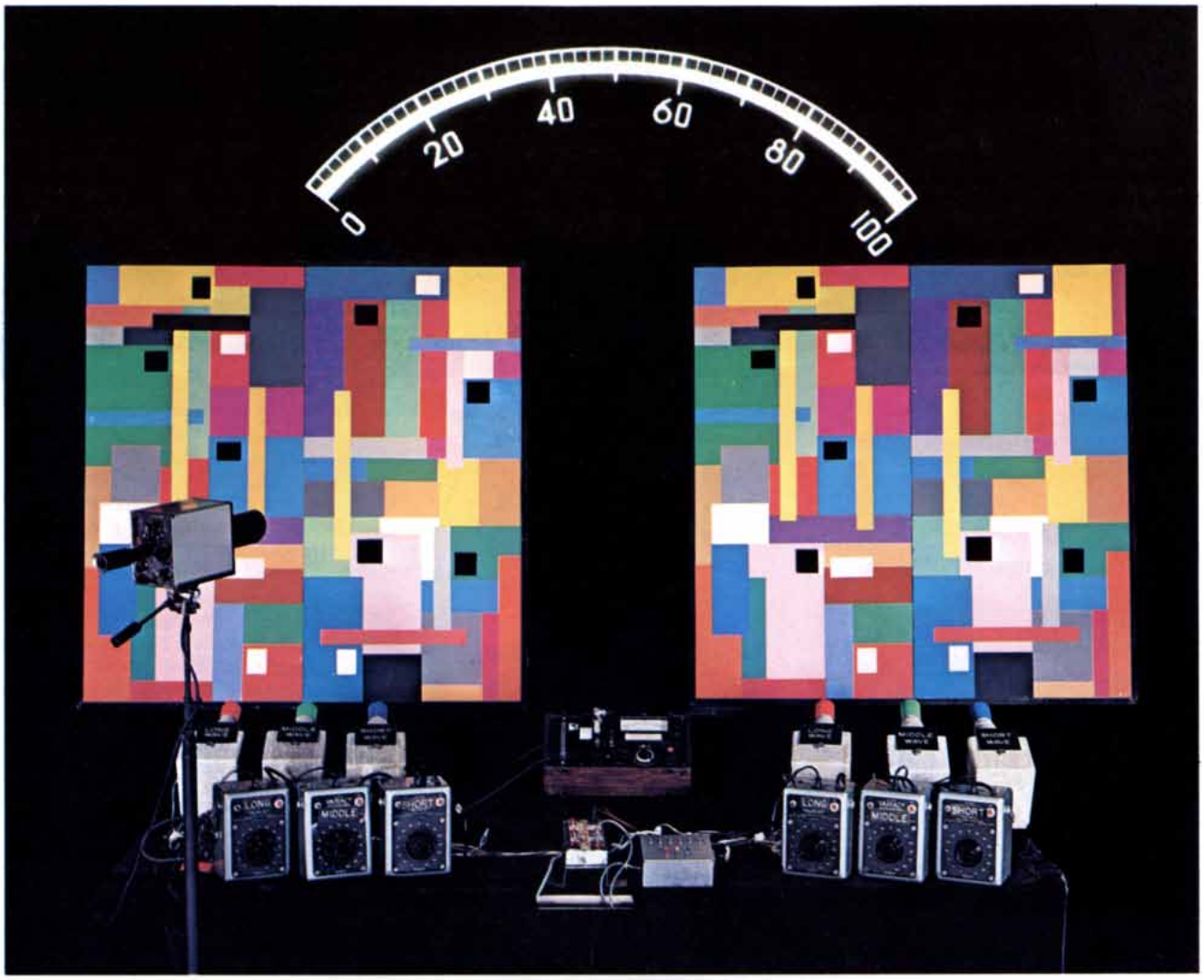
Now that we know that at low light levels an isolated receptor system generates an image in terms of lightness that is completely free of color, might it be possible to bring one of the cone systems into operation along with the hypersensitive system, so that only the



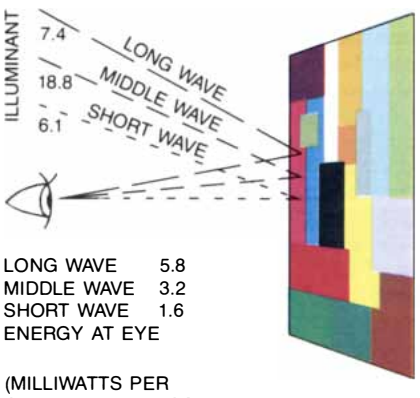
NORMALIZED SPECTRAL SENSITIVITIES OF FOUR VISUAL PIGMENTS (*solid lines*) span the visual spectrum in overlapping curves. Curve that peaks at about 500 nanometers corresponds to sensitivity of rod pigment. Other three curves represent cone pigments. Broken lines show sensitivities of the film-filter combinations that were selected to match the sensitivities of the four retinal pigments and used to make the black-and-white retinex records in the illustration at bottom of preceding page. Cone curves are adapted from work of Paul Brown and George Wald of Harvard University. The rod curve is standard scotopic luminosity curve.

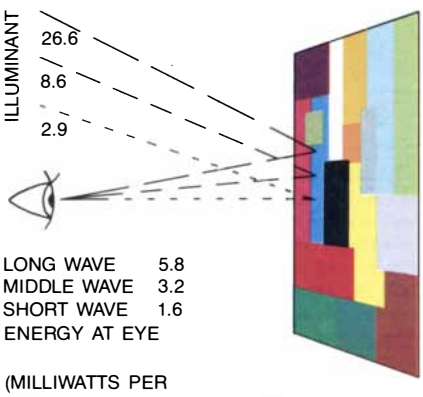


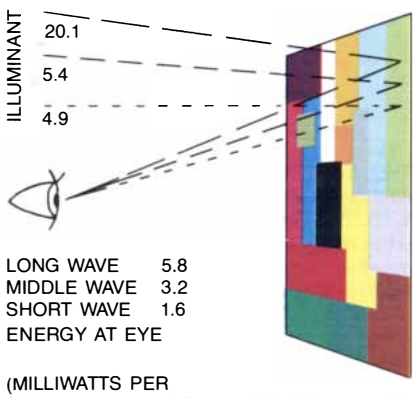
THRESHOLD RESPONSES OF RETINAL RECEPTORS vary by large factors. The hypersensitive rod system provides vision at radiance levels about 1,000 times weaker than the light levels needed to activate the cone systems. It has been shown in author's laboratory that multicolored scenes exhibit nearly their normal range of colors when they are viewed at light levels so adjusted that only rod system and one cone system, the long-wave system, are responding.



"COLOR MONDRIAN" EXPERIMENT employs two identical displays of sheets of colored paper mounted on boards four and a half feet square. The colored papers have a matte finish to minimize specular reflection. Each "Mondrian" is illuminated with its own set of three projector illuminators equipped with band-pass filters and independent brightness controls so that the long-wave ("red"), middle-wave ("green") and short-wave ("blue") illumination can be mixed in any desired ratio. A telescopic photometer can be pointed at any area to measure the flux, one wave band at a time, coming to the eye from that area. The photometer reading is projected onto the scale above the two displays. In a typical experiment the illuminators can be adjusted so that the white area in the Mondrian at the left and the green area (or some other area) in the Mondrian at the right are both sending the same triplet of radiant energies to the eye. The actual radiant-energy fluxes cannot be re-created here because of the limitations of color reproduction. Under actual viewing conditions white area continues to look white and green area continues to look green even though the eye is receiving the same flux triplet from both areas.







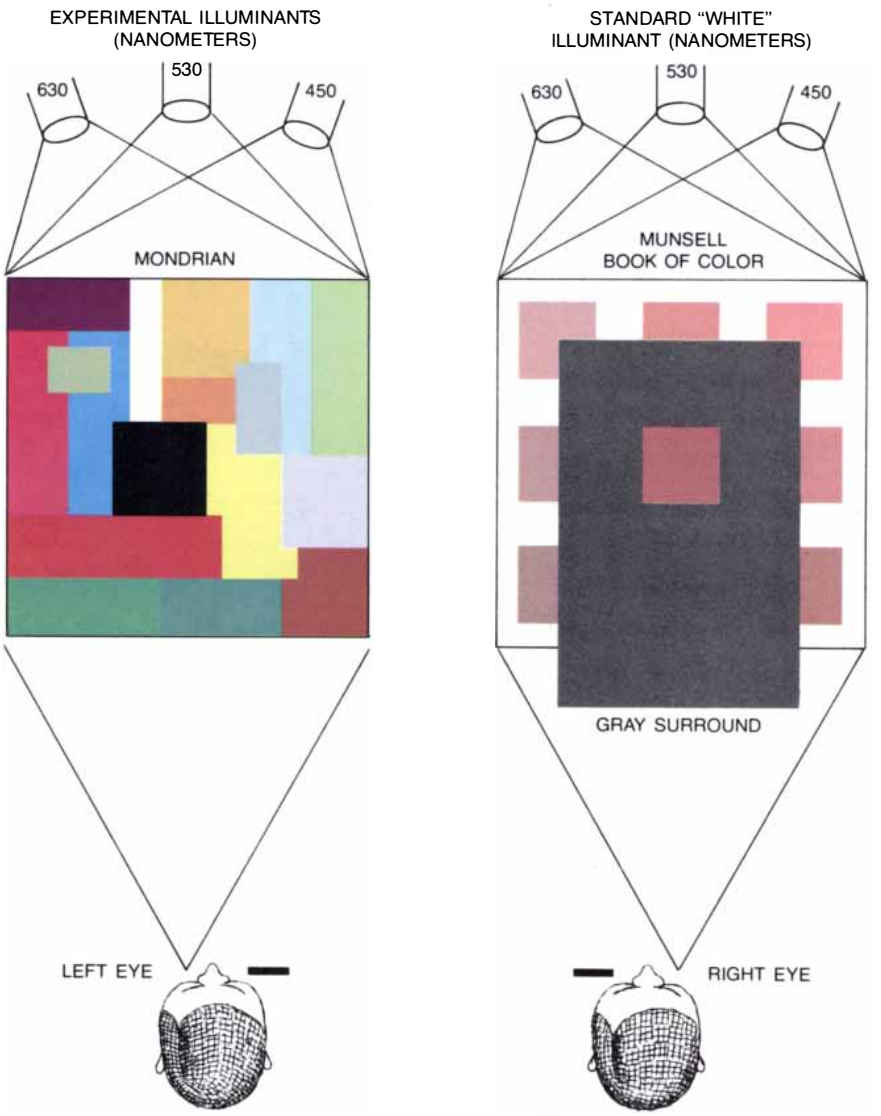
IDENTICAL ENERGY FLUXES AT THE EYE provide different color sensations in the Mondrian experiments. In this example, with the illuminants from the long-wave, middle-wave and short-wave illuminators adjusted as indicated, an area that looks red continues to look red (*left*), an area that looks blue continues to look blue (*middle*) and an area that looks green continues to look green (*right*), even though all three are sending to the eye the same triplet of long-, middle- and short-wave energies. The same triplet can be made to come from any other area: if the area is white, it remains white; if the area is gray, it remains gray; if it is yellow, it remains yellow, and so on.

completely colorless system and one other were functioning? This two-receptor experiment has been carried out and provides a powerful confirmation of the ideas derived from all our binary work with red-and-white images and subsequent ternary studies with multicolored displays seen under various illuminants. The experiment, rapidly becoming a classic, was devised by my colleagues John J. McCann and Jeanne L. Benton.

McCann and Benton illuminated a color display with a narrow wave band of light at 550 nanometers. The light level was raised just above the amount needed to make the display visible to the dark-adapted eye, thus ensuring that only the hypersensitive system was operating. They then added a second narrow-band illuminant at 656 nanometers, with its level adjusted so that it was just sufficient to activate the long-wave receptor system but not the middle-wave system. Under these conditions only two receptor systems, namely the rods and the long-wave cones, were receiving enough light to function.

The resulting image exhibited a remarkable range of color, enabling an observer to assign to each area in the display the same color name it would have if it were illuminated above the cone threshold. The result is reminiscent of the multicolored images produced by the red-and-white system. The demonstration explicitly confirms our early proposition that the lightness information collected at two wave bands by separate receptor systems is not averaged, point by point and area by area, but is kept distinct and is compared. We know that the rod system does not produce a colored image when the image is seen by itself, and we know that the long-wave light alone cannot produce an image with a variety of colors. The combination, however, gives rise to a wide variety of colors, namely reds, yellows, browns, blue-greens, grays and blacks.

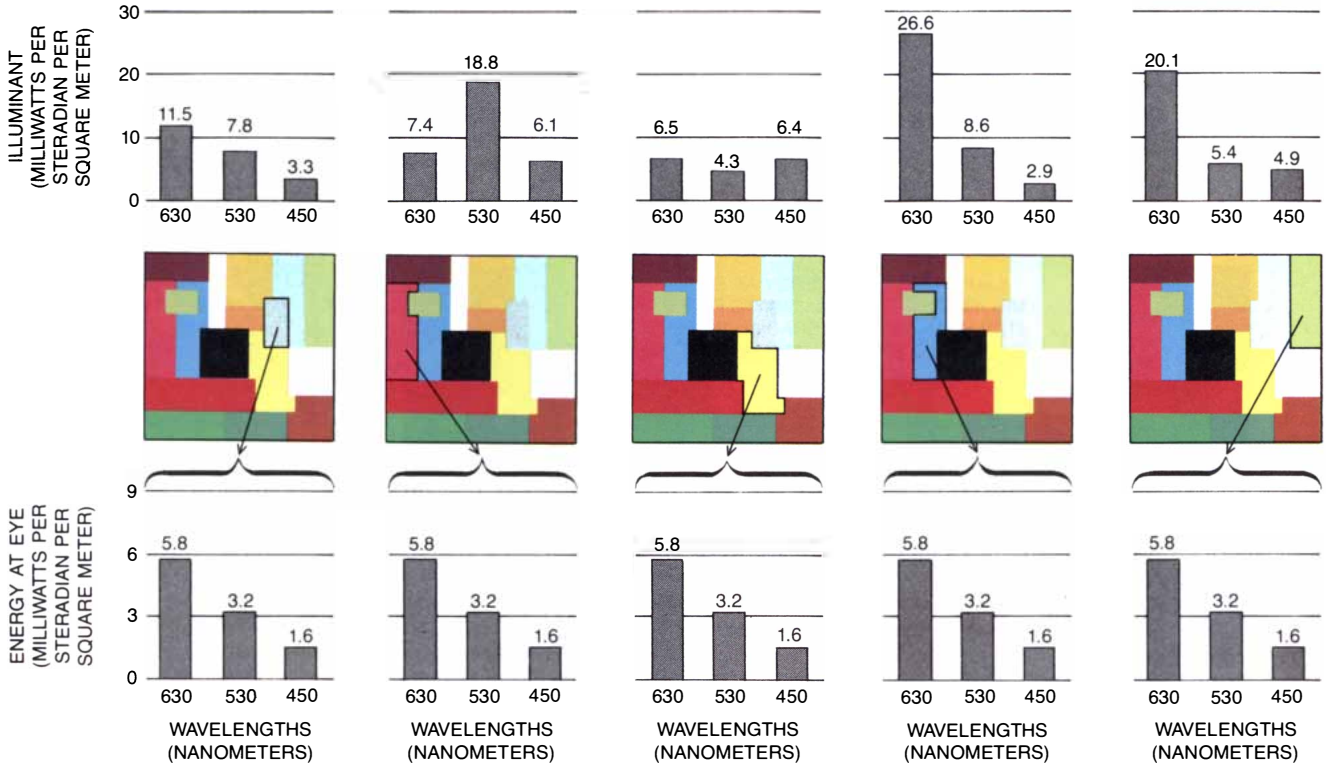
What, then, accounts for the color? The emergence of variegated colors can be ascribed to a process operating somewhere along the visual pathway that compares the lightnesses of the separate images on two wave bands, provided by the two independent retinex systems. The two-receptor experiment makes it plausible that when three independent images constituting the lightnesses of the short-, middle- and long-wave sets of receptors are associated to give a full-colored image, it is the comparison of the respective lightnesses, region by region, that determines the color of each region. The reason the color at any point in an image is essentially independent of the ratio of the three fluxes on three wave bands is that color depends only on the lightness in each wave band and lightness is independent of flux.



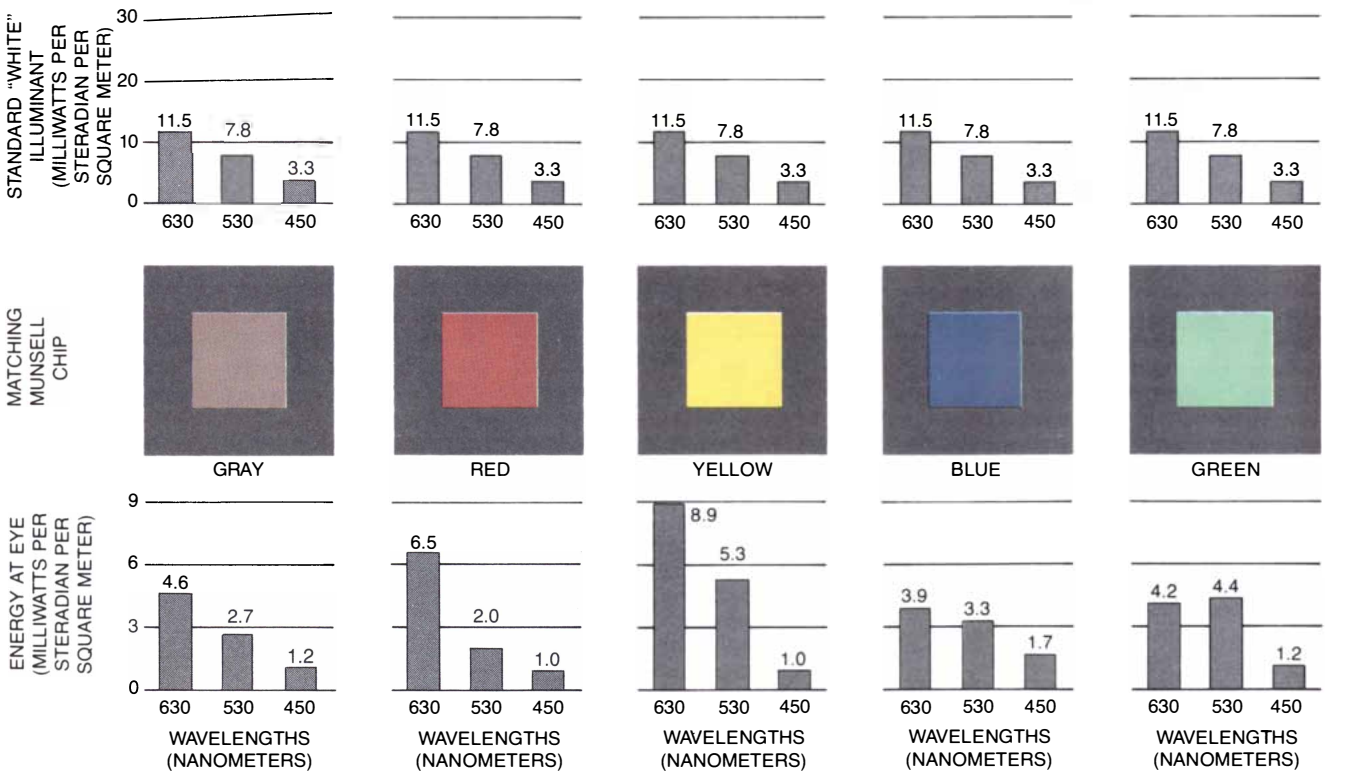
COLOR-MATCHING EXPERIMENT uses a simplified Mondrian of 17 color areas (*left*) and a standard color reference, *The Munsell Book of Color,* which contains 1,150 color "chips" (*right*). The Mondrian is illuminated with three narrow-band light sources: one at 630 nanometers (long-wave light), one at 530 nanometers (middle-wave light) and one at 450 nanometers (short-wave light). The ratio of the three illuminants can be adjusted so that the triplet of energies reflected to the eye from any chosen area will exactly equal the triplet that previously reached the eye from some other area. In this experiment five areas, gray, red, yellow, blue and green, were selected in sequence to send the same triplet of energies to the eye. In each of the five consecutive parts of this experiment the observer selected from the Munsell book the chips that came closest to matching the 17 areas of the Mondrian. The Munsell book was illuminated throughout the experiment with a constant spectral mixture of three narrow-band lights adjusted at the outset so that the white Munsell chip appeared the "best white." The experiment was set up so that the observers used one eye for viewing the Mondrian and the other eye for viewing chips. Gray paper with an opening was used to provide chips with a constant surround.

As we have seen, the spectral sensitivities of the visual pigments overlap broadly. If we illuminated a scene with the entire range of wavelengths to which a single visual pigment is sensitive, we would see a large variety of colors because more than one retinex system would respond. With the help of filters and appropriate film emulsions, however, we can isolate the lightnesses that would ordinarily be incorporated into the sensation of color. We call black-and-white photographs made for this purpose retinex records.

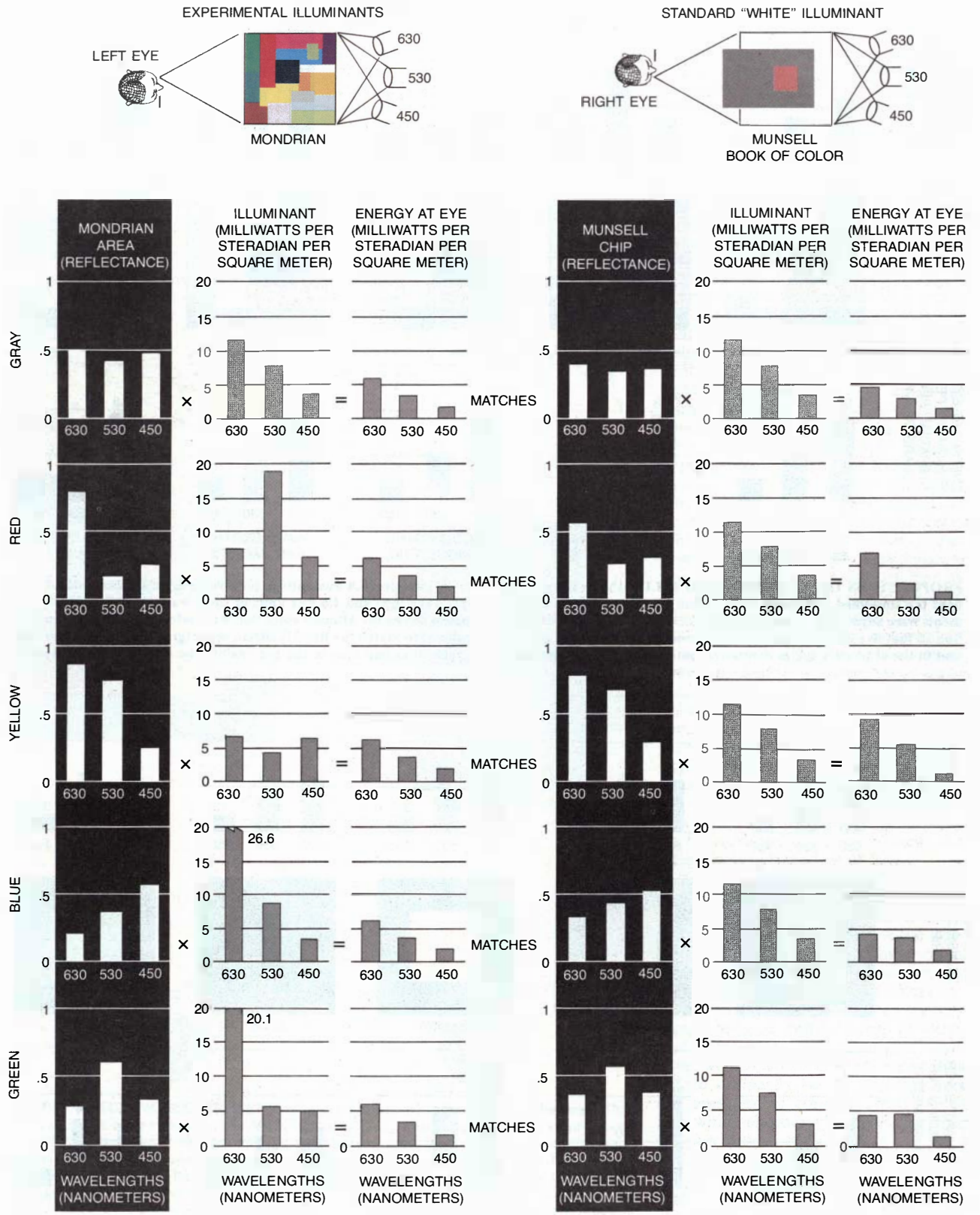
The photographic technique, making use of silver emulsions, performs two functions. First, the system provides spectral sensitivities that are the same as those of the visual pigments. Second, it generates black-and-white pictures for a human observer to examine. It is the human visual system that converts the photographic pattern deposited in silver into lightness. Ideally we should like



PROPORTIONS OF NARROW-BAND ILLUMINANTS used to light the simplified Mondrian in the Munsell-chip matching experiments were adjusted as is shown by the bars at the top of this illustration so that five different areas of the Mondrian (*indicated by arrows*) sent to the observer's eye in successive matching trials the same triplet of energies: 5.8 flux units of long-wave light, 3.2 flux units of middle-wave light and 1.6 flux units of short-wave light. The illustration below shows the Munsell chips that were selected in the constant illuminant to match the five Mondrian areas (gray, red, yellow, blue and green) that had sent to the eye exactly the same triplet of energies.



MUNSELL CHIPS SELECTED BY OBSERVERS to match the five Mondrian areas that had sent identical triplets of energy to the eye are reproduced. The Munsell book was illuminated with a constant spectral mixture of narrow-band illuminants (*bars at top*) and the chips were viewed within a constant gray surround. The energy that was sent to the eye by the selected Munsell chips is shown by the bars at the bottom of the illustration. It is evident that the match between the Mondrian areas and the Munsell chips is not made on the basis of the flux of radiant energy at the eye of the observer. What does cause the two areas to match is described in the illustrations that follow.



FURTHER ANALYSIS OF MATCHING EXPERIMENT begins to identify the basis on which the visual system makes the color match between the Mondrian area and the Munsell chip without regard to the flux each member of the pair sends to the eye. The efficiency with which a given area in the Mondrian reflects light in each of the three wave bands (*first column*) multiplied by the amount of energy striking that area in each of the wave bands (*second column*) yields the energy triplet that reaches the eye (*third column*). The three columns at the right contain comparable data for the Munsell chips selected as a match for the Mondrian areas. Whereas illustration at bottom of the preceding page shows that the eye does not match colors using a "meter" that measures triplets of energies at the eye, this illustration shows that when a match is made, it is the reflectances of two areas that correspond, as is shown in first and fourth columns.

our observer to examine the black-and-white pattern with only one set of cones, reporting the lightnesses appropriate to that set. At any point in the black-and-white pattern, however, the reflectance is essentially the same throughout the visible spectrum. Therefore with a black-and-white photograph we stimulate all the receptors with the same information, that is, with the energies that would be absorbed by a single visual pigment. If we assume that all the retinex systems process information in an identical manner, we can propose that sending this identical information to several sets of receptors is the same as sending it to only one receptor, thereby enabling us to see what the image would look like if it were possible to isolate it.

On page 109 the reader will see three black-and-white pictures taken through retinex filters that simulate the response of the three cone pigments. The strawberries and radishes, for example, are light on the long-wave record, darker on the middle-wave record and darkest on the short-wave record. Although the orange and lemon are about as dark as the strawberries and radishes on the short-wave record, they are nearly as light on the middle-wave record as they are on the long-wave record. On the printed page the distinctions are subtle. To the eye viewing an actual full-color scene the subtle distinctions provide all the information needed to distinguish countless shades and tints of every color.

After the three lightnesses of an area have been determined by the three retinex systems no further information is necessary to characterize the color of any object in the field of view. Any specific color is a report on a trio of three specific lightnesses. For each trio of lightnesses there is a specific and unique color.

The limitations of color photography make it impossible to show the reader the demonstrations readily accomplished in our laboratory, which dramatically reveal the independence of perceived color from the flux reaching the eye. What the reader would see would be two boards four and a half feet square identically covered with about 100 pieces of paper of various colors and shapes. In order to minimize the role of specular reflectance the papers have matte surfaces and, except for black, have a minimum reflectance of at least 10 percent for any part of the visible spectrum. In these displays, which we call "color Mondrians" (after the Dutch painter to whose work they bear a certain resemblance), the papers are arranged so that each one is surrounded by at least five or six others of different colors [*see top illustration on page 111*].

Each of the identical Mondrians is illuminated by its own set of three projectors equipped with sharply cutting band-pass filters (not retinex filters): one at 670 nanometers embracing a band of long waves, one at 540 nanometers embracing a band of middle waves and one at 450 nanometers embracing a band of short waves. The amount of light from each illuminating projector is controlled by a separate variable transformer. In addition the illuminating projectors have synchronized solenoid-activated shutters to control the duration of illumination. There is a telescopic photometer that can be precisely aimed at any region of either Mondrian to measure the amount of radiation reflected from any point and therefore the amount of flux reaching the eye. The output of the photometer is projected on a scale above the Mondrian, where it can be seen by those taking part in the demonstration.

The demonstration begins with the three illuminating projectors turned on the Mondrian on the left; the Mondrian on the right remains dark. The variable transformers are set so that the entire array of papers in the left Mondrian are deeply colored and at the same time the whites are good whites. This setting is not critical. Then, using one projector at a time and hence only one wave band at a time, we measure with the telescopic photometer the energy reaching the eye from some particular area, say a white rectangle. The readings from the white area (in milliwatts per steradian per square meter) are 65 units of long-wave light, 30 units of middle-wave light and five units of short-wave light. We have now established the three energies associated with that sensation of white.

We turn off the three projectors illuminating the color Mondrian on the left. On the right we turn on only the long-wave projector. We select a different area of unknown color and adjust the long-wave light until the long-wave energy coming to the eye from the selected area is the same as the long-wave energy that a moment ago came from the white paper in the Mondrian on the left, 65 units. We turn off the long-wave projector and separately adjust the transformers controlling the middle- and short-wave projectors, one after the other, so that the energies sent to the eye from the selected area are also the same as those that came from the white area on the left. We have not yet turned on all three light sources simultaneously, but we know that when we do so, the triplet of energies reaching the eye from the selected area of still unknown color will be identical with the triplet that had previously produced the sensation white.

When we turn on the three illuminants, we discover that the area in the Mondrian on the right is green. We now illuminate the Mondrian on the left with its illuminants at their original settings so that both Mondrians can be viewed simultaneously. The white area on the left continues to look white and the green area on the right continues to look green, yet both are sending to the eye the same triplet of energies: 65, 30 and five in the chosen units.

We turn off the illuminants for both Mondrians and select some other area in the left Mondrian and sequentially adjust the energies reaching the eye from it so that they are the same as the energies that originally gave rise to the sensation of white and also gave rise to the sensation of green in the right Mondrian. When we turn on all three projectors illuminating the left Mondrian, we see that this time the selected area is yellow. The triplet of energies reaching our eye is the same one that had previously produced the sensations of white and green. Again, if we wish, the yellow and green can be viewed simultaneously, with yellow on the left and green on the right.

We can continue the demonstration with other areas such as blue, gray, red and so on. It is dramatically demonstrated that the sensation of color is not related to the product of reflectance times illumination, namely energy, although that product appears to be the only information reaching the eye from the various areas in the Mondrians.

In order to demonstrate that the color sensations in these experiments do not involve extensive chromatic adaptation of retinal pigments the projectors are equipped with synchronized shutters so that the Mondrians can be viewed in a brief flash, a tenth of a second or less in duration. Regardless of the brevity of observation the results of the demonstrations are not altered. Thus one can say that neither chromatic adaptation nor eye motion is involved in producing the observed colors. Finally, the very essence of the design of the color Mondrian is to obviate the significance of the shape and size of surrounding areas, of the familiarity of objects and of the memory of color. Curiously, from time to time there is a casual attempt to adduce what is called color constancy as an explanation of these demonstrations. Clearly color constancy is only a compact designation of the remarkable competence that is the subject of this article.

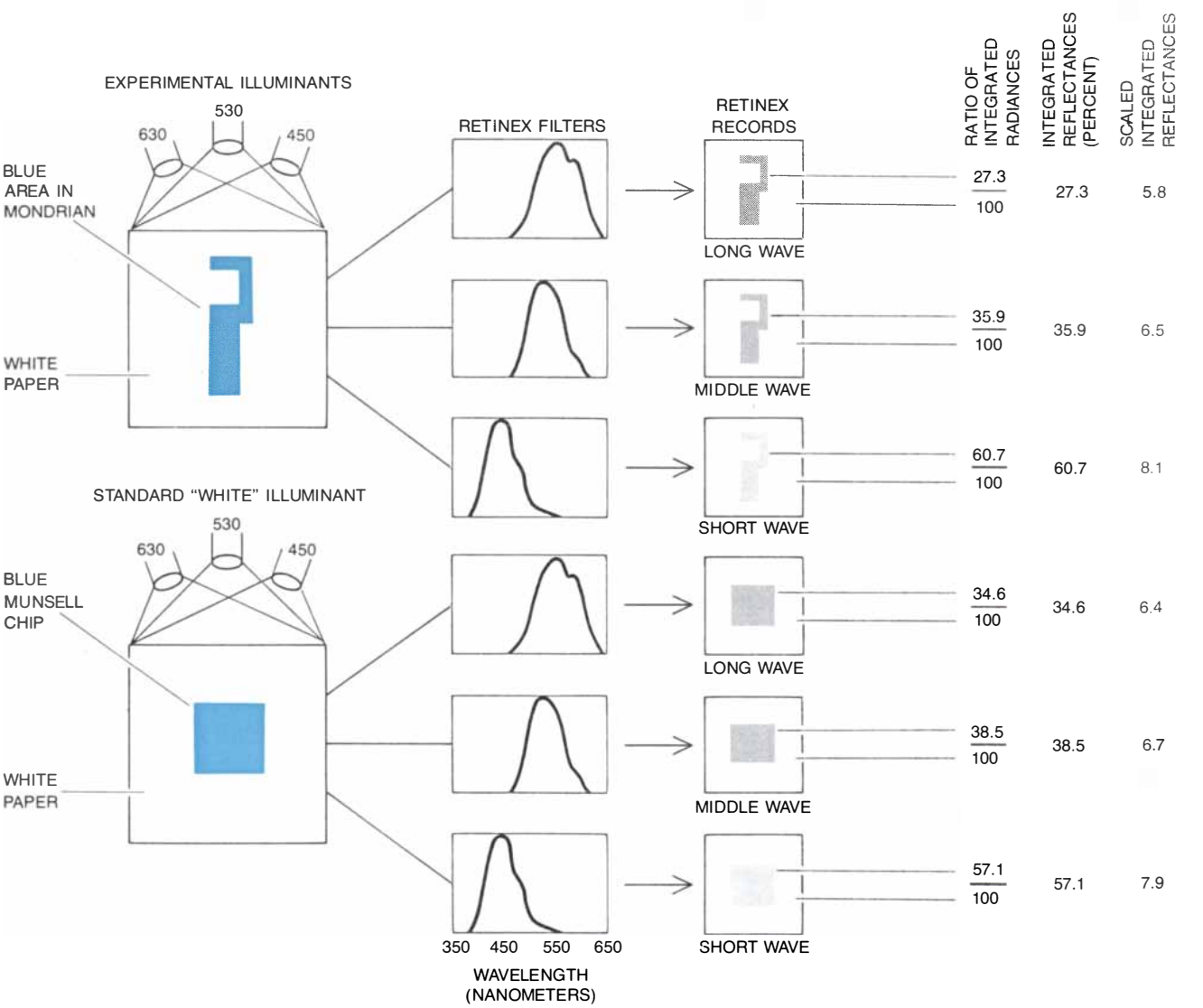
The mystery is how we can all agree with precision on the colors we see when there is no obvious physical quantity at a point that will enable us to specify the color of an object. Indeed, one can say the stimulus for the color of a point in an area is not the radiation from that point. The task of psychophysics is to find the nature of the stimulus for that color.

Here let us remember that what the eye does unfailingly well is to discover

lightness values independent of flux. We saw this to be true for a single receptor system, the rod system, operating alone and for the three cone systems operating collectively when they viewed an array of white, gray and black papers. Let us now illuminate the colored Mondrian array with light from just one of the three projectors, say the projector supplying long-wave light, and observe the effect of increasing and decreasing the flux by a large factor. We observe that the various areas maintain a constant rank order of lightness. If, however, we switch the illumination to a different wave band, say the middle wave band, the lightnesses of many of the areas will change; many of the 100 or so areas will occupy a different rank order from lightest to darkest. Under the short-wave-band illuminant there will be yet a third rank order. Specifically, a red paper will be seen as being light in the long-wave light, darker in middle-wave light and very dark in short-wave light. A blue paper, on the other hand, will be light in short-wave light and very dark in both middle- and long-wave light. Papers of other colors will exhibit different triplets of lightnesses. When we conducted such experiments nearly 20 years ago, we were led inevitably to the conclusion that the triplets of lightnesses, area by area, provided the set of constancies we needed to serve as the stimuli for color, independent of flux.

It is evident that the lightnesses exhibited by a given piece of colored paper under illuminants of three different wave bands is related to the amount of energy the paper reflects to the eye at different wavelengths. Let us now examine, by means of a particular experiment, how such reflectances can be related step by step to perceived lightnesses and how, in the process, the radiant flux that reaches the eye—the ultimate source of knowledge about lightness—finally becomes irrelevant to the sensation of color.

In our laboratory McCann, Suzanne P. McKee and Thomas H. Taylor made a systematic study of observers' re-



ROLE OF REFLECTANCE and its psychophysical correlate, lightness, in guiding the eye to match Munsell chips with Mondrian areas was examined with the help of retinex filter-photomultiplier combinations that match the spectral sensitivity of the cone pigments. Under each combination of illuminants (*top*) the integrated radiance, or flux, in each retinex wave band of a Mondrian area was compared with the integrated radiance of a sheet of white paper. The ratio of integrated radiances yields the integrated reflectance of the Mondrian area, expressed here in percent. For the matching Munsell chip a set of ratios was similarly determined (*bottom*). The final step in deriving a physical equivalent of lightness is the scaling, or spacing, of integrated reflectances to be consistent with the spacing of lightness sensations. This transformation is explained in the illustration on the opposite page. The scaled values appear in the column at the right.

sponses to a simplified color Mondrian with areas of 17 different colors. They asked the observers to match the 17 areas one at a time under different illuminants with colored squares of paper that had been selected from a standard color-reference book, *The Munsell Book of Color* and that were viewed under a constant "white" illumination.

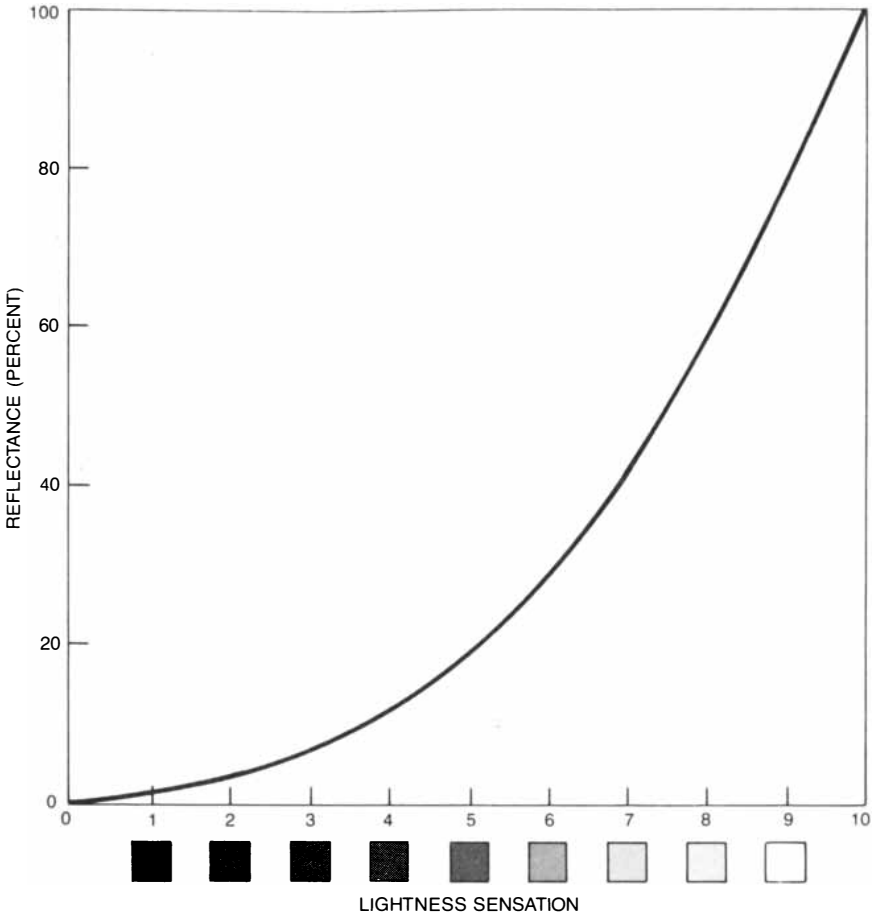
The illuminants on the Mondrian were adjusted in five separate matching experiments so that five different areas (gray, red, yellow, blue and green) sent to the eye an identical triplet of radiances. The observer began by selecting a matching Munsell "chip" for each of the 17 areas in the Mondrian when the gray area in the Mondrian sent a particular triplet of energies to the eye. Another set of 17 matching Munsell chips was selected when the same triplet was later sent to the eye by a red area in the Mondrian, and the same was done for yellow, blue and green areas under illuminants that supplied the same triplet of energies.

The illustrations on page 113 show the details of the experiment and the five different Munsell colors the observers selected to match the five areas when each area sent to the eye precisely the same triplet of energies. In spite of the constancy of the energy reaching one eye from the Mondrian, each observer, using the other eye, selected Munsell chips that were gray, red, yellow, blue and green.

The constant illumination used in viewing the Munsell book was a triplet of illuminants at three wavelengths that observers judged to produce the "best" white. The actual triplet of wavelengths reaching the eye from the whitest paper in the Munsell book was 11.5 units of long-wave light, 7.8 units of middle-wave light and 3.3 units of short-wave light. The illuminants supplied energy in narrow bands with peaks at 630 nanometers, 530 nanometers and 450 nanometers. A similar triplet of narrow-band illuminants were mixed in various proportions to illuminate the Mondrian.

At this point the reader might ask: Would not a single gray area exhibit a pronounced change in color if the surrounding papers had reflected light of widely differing spectral composition? Could these changes in color account for the results of the Mondrian experiments? The answer to the questions is that no manipulation of surrounding papers in the Mondrian is capable of making the gray paper match the red, yellow, blue and green Munsell papers selected by the observers in the Mondrian experiment.

McCann, John A. Hall and I have examined the matter further by repeating the Mondrian-Munsell experiment in various ways so that the average spectral composition of the light reaching the eye from the Mondrian and its surround remains the same regardless of the spectral composition of the light needed to establish a constant triplet from area to area. We have done this in one case by surrounding the entire Mondrian with brightly colored papers selected in such a way that they exactly offset the average mixture of wave bands from the Mondrian itself and, more dramatically, by cutting the 17 areas of the Mondrian apart and placing them well separated on the backgrounds of offsetting color. Neither arrangement has any significant effect on the Munsell chips chosen to match the various areas of the Mondrian.



SENSATION OF LIGHTNESS is plotted on an equal-interval scale. Observers are shown a sheet of white paper (*9*) and a sheet of black paper (*1*) and are then asked to choose a sheet of paper whose shade of gray lies halfway between the two. The selection is the gray labeled *5*. Similar selections are made to determine the locations of midpoints between *1* and *5* and between *5* and *9* and so on until the equal-interval scale is filled. The end values *0* and *10* are extrapolations. The curve is then plotted by measuring the reflectances of the various papers selected by the observers. The curve makes it possible to convert values of integrated reflectance into values of scaled integrated reflectance, as is given in illustration on opposite page.

Let us return, then, to the search for the stimulus that guides us so accurately to the correct identification of colors. If it is not a flux of radiant energy at the eye from each point in the field of view, what are the physical correlates of the lightnesses of objects on three separate wave bands, corresponding to the spectral sensitivities of the cone pigments? Can such a precise physical correlate of lightness be demonstrated?

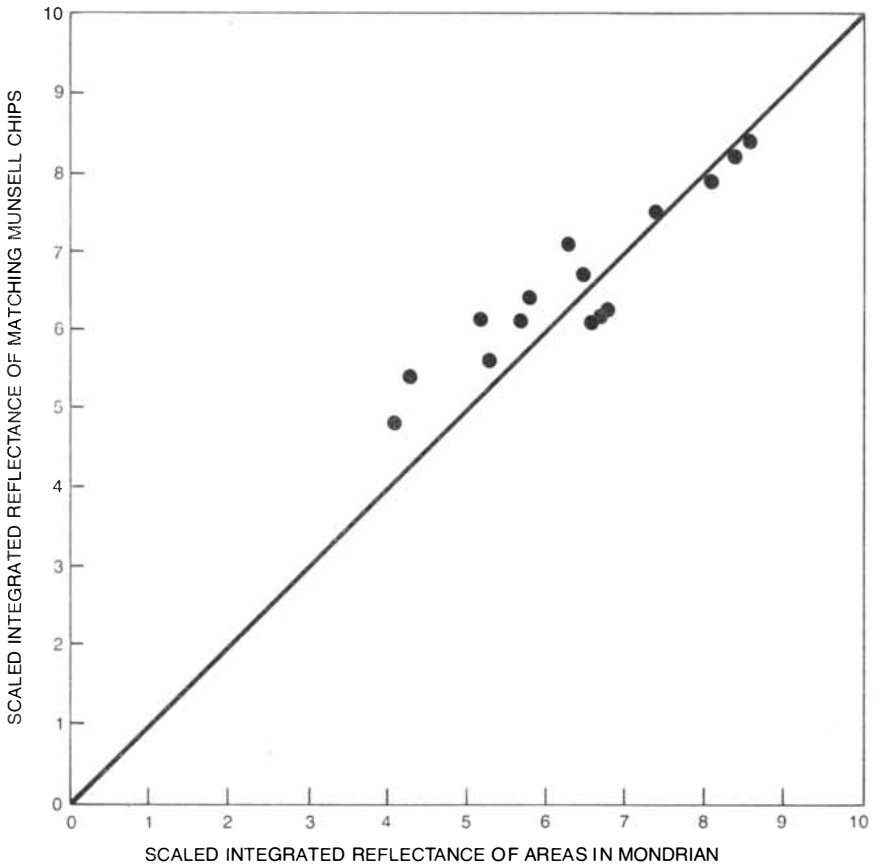
McCann, McKee and Taylor next measured the radiance, or energy at the eye, of the various Mondrian areas and of the matching Munsell chips by using a photomultiplier in conjunction with a version of the retinex filters. Since the retinex-photomultiplier combination integrates the flux of radiant energy over a broad band of wavelengths, the instrument provides a value we call integrated radiance. McCann and his colleagues then obtained the integrated radiances from a large sheet of white paper placed under each of the experimental illuminants that had been used to light the Mondrian in the chip-matching experiments. If the integrated radiance from a Mondrian area is used as the numerator in a fraction and the integrated radiance from the white paper is used as the denominator, one obtains a value for in-

tegrated reflectance, which can be expressed as a percent.

The integrated reflectances for the various Munsell chips are determined in the same manner under the constant "white" illumination. This amounts to measuring the percentage of reflectance using detectors with the same spectral sensitivity as the visual pigments. The results show that the Munsell chip chosen by the eye to match a given Mondrian area will have approximately the same three integrated reflectances as the area. For example, the blue area in the Mondrian has a triplet of integrated reflectances (long-, middle- and short-wave) of 27.3, 35.9 and 60.7 percent. The comparable values for the matched Munsell chip are 34.6, 38.5 and 57.1 percent [*see illustration on page 116*].

Finally, the integrated reflectances are "scaled" so that their equal spacing is consistent with the equal spacing of lightness sensations. The curve for this transformation is shown in the illustration on the preceding page. Using this curve, we see that the blue area in the Mondrian has a triplet of scaled integrated reflectances of 5.8, 6.5 and 8.1, whereas the corresponding values for the matching Munsell chip are 6.4, 6.7 and 7.9. If we study the five areas that successively sent identical triplets of energies to the eye and compare their scaled integrated reflectances with those of their matching Munsell chips, we find that all the values are in excellent agreement. In other words, in the triplets of integrated reflectances we have identified a highly accurate physical correlate of color sensations. The data fall along the 45-degree line that describes the locus of perfect correlation [*see illustration below*].

We have sought a physical correlate for lightness, and we have found that the scaled integrated reflectances of the five areas that sent identical triplets of fluxes to our eyes are the same as those of the matching Munsell chip. This correlation enables us to use scaled integrated reflectances as a measured lightness equivalent. The problem now shifts to one of how the eye derives the lightness that corresponds to the reflectances of objects in each wave band.



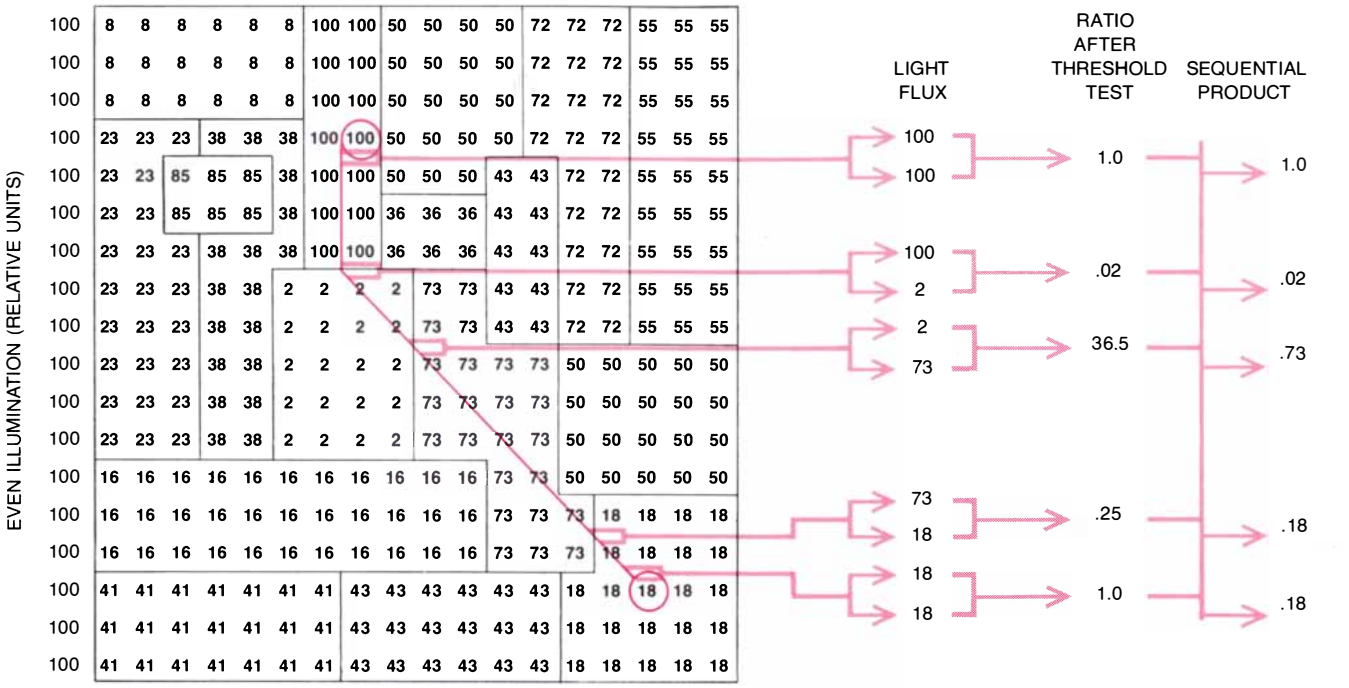
AGREEMENT IN SCALED INTEGRATED REFLECTANCES between Mondrian areas and Munsell chips chosen to match them is summarized for all three wave-band systems. The scaled integrated reflectances of five Mondrian areas and matching Munsell chips were determined as is described in illustration on page 116. In this graph triplets of scaled integrated reflectances of five Mondrian areas that sent identical fluxes to the eye are plotted against scaled integrated reflectances of Munsell chips chosen to match them. Although the dots collectively represent correspondence for all three retinex wave bands, any particular dot denotes the degree of correspondence on one retinex wave band between a Mondrian area and a Munsell chip. Close correspondences show that scaled integrated reflectance is physical correlate of the sensation "lightness," showing precision with which a triplet of lightnesses determines color.

It is one thing to measure a triplet of lightness equivalents using a retinex filter coupled to a photomultiplier; it is quite another for the eye to determine lightnesses in the unevenly lighted world without reference sheets of white paper. I described above the ability of an isolated receptor system—the hypersensitive system of rod vision—to classify objects correctly according to their inherent reflectivity regardless of whether the objects happened to be in a brightly or a dimly lighted region of visual space. The ability of one receptor system to work in this way makes it plausible that the other three systems of normal daytime vision possess the same ability, each system viewing the world through a broad but restricted region of the spectrum, the regions we duplicate with retinex filters. Each system forms a separate lightness image of the world. The images are not mixed but compared. The comparison of lightnesses at each area gives rise to the range of sensations we know as color.

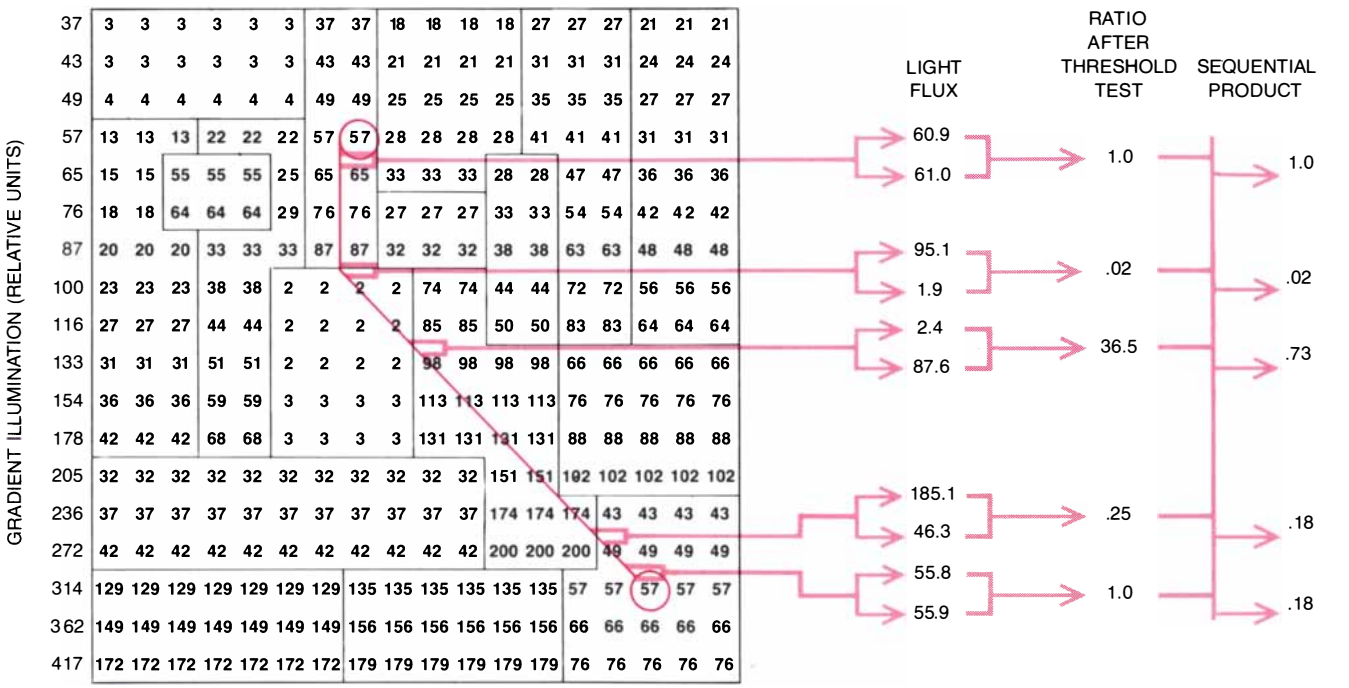
How could the biological system generate a hierarchy and spacing of lightness values given only the flux from each point in a scene and knowing nothing about the pattern of illumination and nothing about the reflectances of objects? The scheme I am about to describe is the most general we have found that surmounts these limitations; its physiological embodiment could take many forms.

Let me begin by pointing out the significance of edges in defining objects or areas in a scene. If a sheet of white paper is lighted strongly from one side, we see no discontinuity in color from one side to the other. Let us now imagine two light detectors positioned to measure the luminance from two different places on the paper. If the illumination is nonuniform, the luminances of the two places will of course be different. As the two detectors are moved closer together the luminances approach the same value and the ratio of the two outputs approaches unity. If, however, the two detectors bridge the boundary between two areas that differ abruptly in reflectance, such as would be the case with even a pale gray square on the white paper, the ratio of the outputs of the two detectors will approach the ratio of the two reflectances. Thus the single procedure of taking the ratio between two adjacent points can both detect an edge and eliminate the effect of nonuniform illumination. If we process the entire image in terms of the ratios of luminances at closely adjacent points, we can generate dimensionless numbers that are independent of the illumination. These numbers give the ratio of reflectances at the edge between adjacent areas; the reflectances themselves are not yet ascertained.

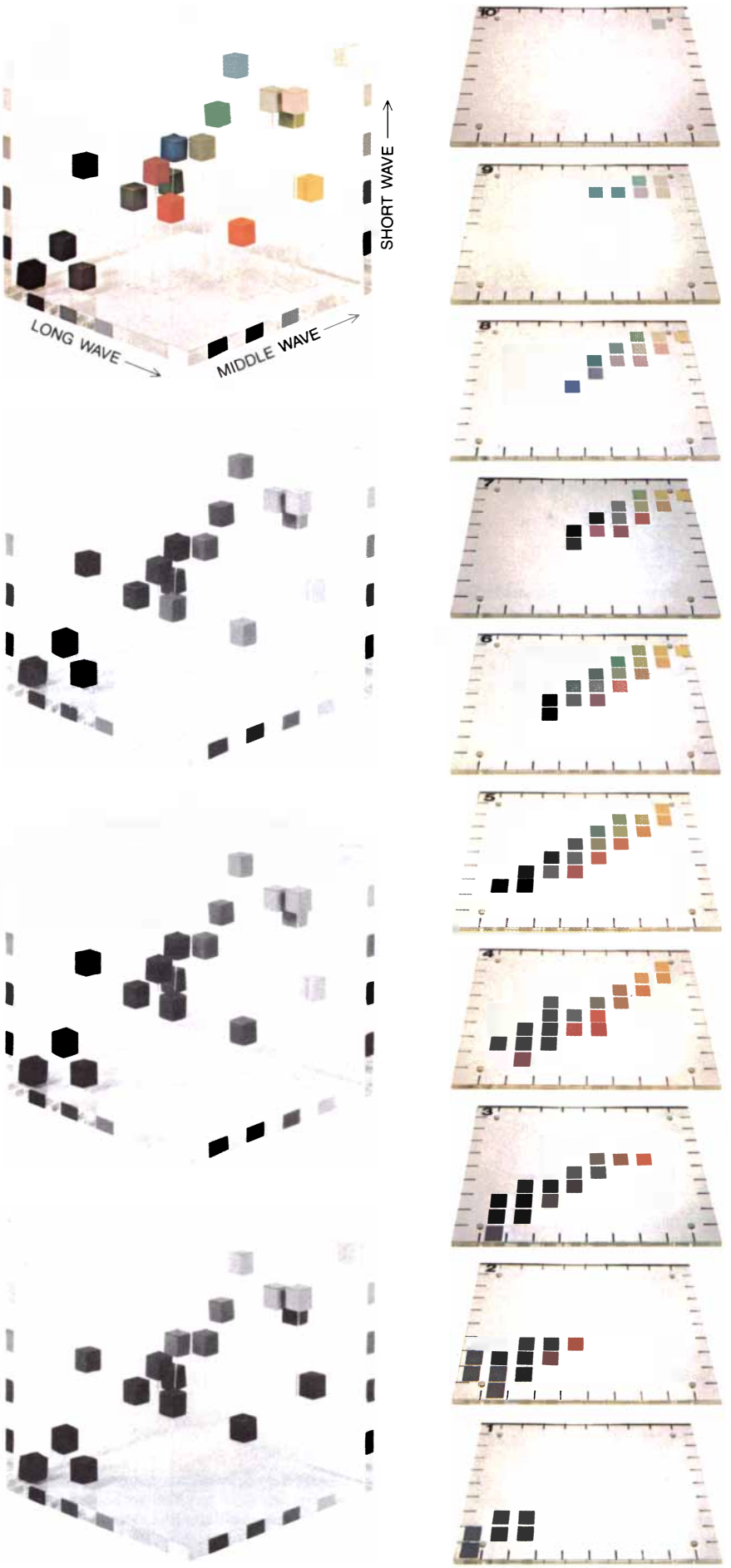
In order to determine reflectances we need to relate all these ratios of reflec-



THE EYE'S METHOD OF DISCOVERING LIGHTNESS in complex images remains to be established. An efficient and physiologically plausible scheme is depicted in this illustration and the one below. The numbers inside the schematic Mondrian represent the long-wave integrated radiances coming from each area of a display that is evenly lighted. The long-wave retinex system independently "measures" the long-wave integrated radiance, point by point, as if it were doing so along an arbitrary pathway (*color*). The flux at each successive closely spaced pair of points is converted into a ratio. This ratio is subjected to a threshold test: any ratio to be regarded as a change must vary from unity by more than some small threshold amount (plus or minus .003 in the computer program). If the ratio does not vary from unity by this amount, it is regarded as being "unchanged" and is set to equal unity. A second threshold-tested ratio along the same pathway is multiplied by the first ratio to give a sequential product that is both the model's response for that point and the signal sent along to be multiplied by the next ratio. When the path crosses an edge between two lightnesses, there is a sharp change in the threshold-tested ratio and hence a similar change in the sequential product. Here the path is started in the white area, where the flux of radiant energy is 100. By the time the path reaches the brown area at the lower right the product is .18. The retinex system has thus determined that the brown area reflects 18 percent as much long-wave energy as the white area. Any other path ending in the brown area would yield the same result as long as it had been through the white area. By averaging the responses for each area, as computed by many arbitrary paths, the long-wave retinex system arrives at a single reflectance value for each area, which designates perceived lightness. Middle- and short-wave retinex systems compute their own sets of lightness values. Comparison of triplet of lightnesses for each area provides sensation of color.



MORE REALISTIC CASE OF GRADED ILLUMINATION is handled equally well by the sequential-product method to arrive at the same reflectance value of .18 for the brown area at the end of the path, even though here the long-wave retinex system receives as much flux from the middle of the brown area (*57*) as it does from the middle of the white area (*57*). The scheme hence provides a means for arriving at computed reflectance independent of flux and without resort to white cards as standards. Precise values of light flux along pathway in this diagram were derived from a computer program that works with 75 values between every two values printed within Mondrian.



tances in the field of view. Given the ratio of luminances at the edge between a first area and a second one, we multiply it by the ratio of luminances at the edge between the second area and a third. This product of ratios approaches the ratio of reflectances between the first and third areas, regardless of the distribution of illumination. Similarly, we can obtain the ratio of reflectances of any two areas in an image, however remote they are from each other, by multiplying the ratios of all the boundaries between the starting area and the remote area. We can also establish the ratio of the reflectance of any area on the path by tapping off the sequential product reached at that area [*see illustrations on preceding page*].

We are now coming close to the answer to the question: How can the eye ascertain the reflectance of an area without in effect placing a comparison standard next to the area? The sequential product can be used as a substitute for the placement of two areas adjacent to each other, thus defining a photometric operation feasible for the eye.

The remaining task is to suggest how the eye can discover the area of highest reflectance in the field of view and then decide whether that area is actually white or some other color. In the model we have proposed, sequential products are computed along many arbitrary pathways that wander through the two-dimensional array of energies on the model's "retina." Since the pathways

COLOR "SOLID" shows the location of all perceivable colors, including white and black, in a three-dimensional color space constructed according to the author's retinex theory. The position of a color in this space is determined not by the triplet of energies at a point but by the triplet of lightnesses computed by the eye for each area. The color photograph at the top left shows the location of representative colors throughout the space. The direction of increasing lightness along each axis is shown by the arrows. The three black-and-white photographs of the color solid were taken with retinex filter-film combinations. They show the lightness values of the representative colors as they would be perceived separately by the eye's long-wave (*top*), middle-wave (*middle*) and short-wave (*bottom*) visual pigments. The set of 10 color pictures at the right represents horizontal planes cut through the three-dimensional color space. Each plane is the locus of colors possible with a constant short-wave lightness. For example, the fifth plane from the bottom shows the variety of color sensations from all possible long- and middle-wave lightness values when those values are combined with a short-wave lightness of 5. The colored squares are samples taken from *The Munsell Book of Color*. In general the blank areas on each plane represent regions where colors could be produced only by fluorescent dyes, if they were produced at all.

The people who developed the 4 hour video home system now present some of the ways it can change your life.

Introducing Omnivision IV™ VHS,™ the Panasonic four-hour video home player/recorder from the people who developed the four-hour video home system—the engineers of our parent company.

Four hours of video recording. Instead of two.

Four hours of taping when you're not at home.

When work rearranges your schedule, rearrange the network's schedule.

So when you know you're going to be stuck at the office, you can also know that four hours of video taping will let you enjoy every moment of the big game...even if it runs into overtime or extra innings.

Two great shows on at the same time? Watch one, tape the other.

Four hours of taping when you are at home. So when the kids want to watch "Charlie Chocolate and His Merry Adventures" on one channel, you can still tape "The Monster that Ate New York" on another channel. And four hours means you can tape the entire movie...even the climax when the Monster eats Yankee Stadium.

Four hours of video taping instead of two means there's virtually no program, no special, no sports event, no movie, no historical event that you can't record on one tape cassette with Omnivision IV.

Which means when a show you want to see runs longer than two hours (and don't the good ones always?) and you're miles away, or fast asleep, you don't have to worry about changing tapes, the way you'd have to with a two-hour system.

Okay, so you're convinced four hours of video taping is better than two hours. But there are a lot of other reasons to choose Omnivision IV VHS:

Catch up with your weekend chores while you catch the weekend game.

With our optional black and white video camera with sound, you can make your own TV shows. And they can be played on any VHS unit—regardless of brand.

You can use the same tape many times. Rerecording automatically erases old material. And to add to your video library of shows, sports and history-making events you've recorded, arrangements are now in progress for services that will allow you to transfer your slides and home movies to VHS™ cassettes. Also prerecorded cassettes on a wide variety of subjects are available.

Omnivision IV™ can work with almost any TV set (although, naturally, we'd like to see it on a Panasonic TV... but that's another ad).

Omnivision IV VHS has Auto-Stop to automatically stop the tape when it runs out. Eliminates the need for you to be there.

Build a video library of great shows, sports and historical events.

To start tape recording when you're not at home, there's an electronic digital clock with a 24-hour timer.

And to find the exact place on the tape, there's memory rewind and a digital tape counter. To make recording convenient there's a remote pause control.

But that's not all: For long life and high performance there are HPF™ heads. Our annealed aluminum die-cast chassis means strength and stability. And a direct-drive motor provides greater picture stability than the belt-drive systems some of our competitors use.

Make a TV show with optional B/W camera and send it to Grandma or anyone with any VHS unit.

Can Omnivision IV VHS change your life? We think so. In fact, we think that the 4-hour video home system is the most important advance in television since television. But see it for yourself at selected dealers.

CAUTION: Unauthorized recording of copyrighted television programs, films, video tapes and other materials may infringe the rights of copyright owners and be contrary to copyright laws.

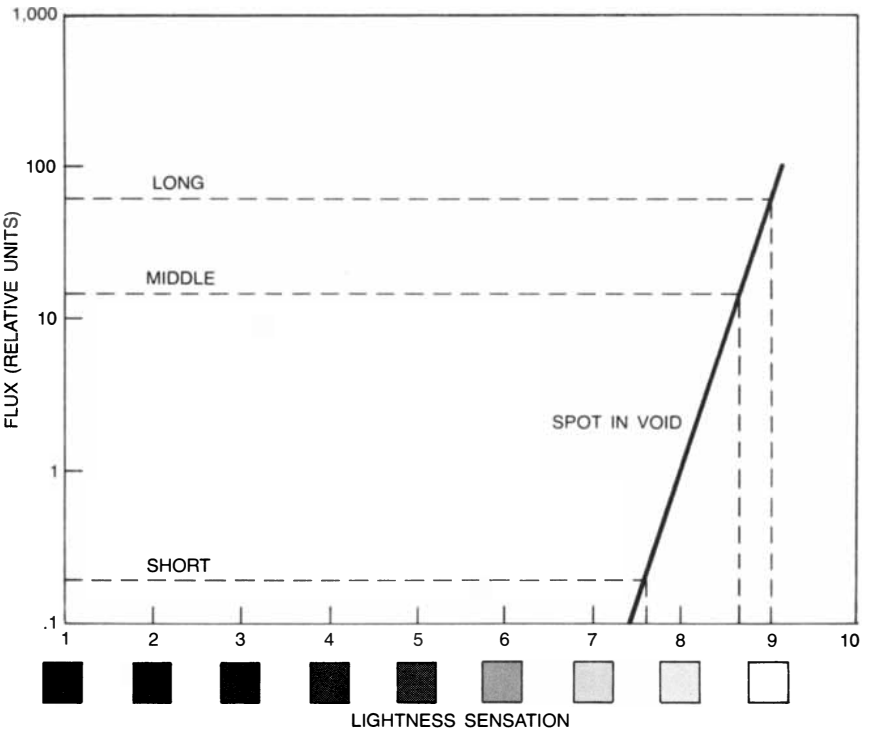
Panasonic
Omnivision IV™ VHS™

can begin anywhere, not just in regions of the highest reflectance, the first value in any sequence is arbitrarily assumed to be 100 percent. Because of this deliberately adopted fiction the sequential product becomes greater than unity whenever the path reaches an area whose reflectance is higher than that of the starting area.

The attainment of a sequential product greater than unity indicates that the sequence should be started afresh with the new area of high reflectance taken as being 100 percent. This procedure is the heart of the technique for finding the highest reflectance in the path. After the path reaches the highest reflectance in the scene, each of the sequential products computed thereafter becomes a fraction of the highest value. A satisfactory computer program has been designed to study the number of paths, their lengths and convolutions, the threshold values for recognizing edges and, perhaps most important, how to utilize all the pathways starting in all areas.

The biological counterpart of this program is performed in undetermined parts of the pathway between the retina and the cortex. The process that corresponds to computing sequential products does not involve the averaging of areas or the averaging of flux. It does, however, call for an arithmetic that extends over the entire visual field. Furthermore, since the relevant phenomena are seen in a brief pulse of light, all the computations and conclusions about lightness must be carried out in a fraction of a second without dependence on eye movement. With a single pulse, eye movement, by definition, is not necessary. With continuous illumination the normal quick motions of the eye probably serve to maintain the freshness of the process.

With our computer model we can obtain a triplet of lightnesses for each area in the color Mondrian that corresponds closely to the lightnesses one would measure with a combined retinex filter and photomultiplier. The color corresponding to any given triplet can be visualized with the aid of the color "solid" we have built, in which the Munsell



SPOT OF LIGHT IN A VOID, that is, a single spot of narrow-band light viewed in an otherwise totally dark environment, has a color that would seem to depend solely on its wavelength. The color can also be explained, however, by the retinex theory in terms of lightness as perceived by the eye's three receptor systems. Psychophysical measurements show that when the eye is presented with a spot of light in a void, the perceived lightness is changed only slightly by very large changes in flux, as is indicated by the straight line. For example, if the spot is composed of a narrow-wavelength band centered, say, at 600 nanometers, the three cone pigments will absorb the flux in quite different amounts because of the shape of their absorption curves. In arbitrary units the long-wave pigment might absorb 80 units, the middle-wave pigment 20 units and the short-wave pigment a few tenths of a unit at most. If these ratios are plotted on the spot-in-a-void curve, the corresponding lightness values are 9 for the long waves, 8.5 for the middle waves and 7.5 for the short. This combination of lightnesses is perceived as a light reddish orange, not ordinarily seen under normal conditions unless surfaces are fluorescent.

colors are located in three dimensions in "lightness-color space" according to their lightness values measured in three wave bands through retinex filters [*see illustration on page 120*].

In normal images the sensation of white light will be generated by any area that is placed at the top of the lightness scale by all three retinex systems. On the other hand, an area that stands at the top of only two of the three lightness scales will be seen as some other color. Hence an area that is at the top of the lightness scale in the long- and middle-wave systems but is surpassed in lightness by some other area in the short-wave system will be seen not as white but as yellow. A similar intercomparison of triplets of lightnesses at the same place within each scene provides the sensation of color, area by area, in spite of unpredictable variations in illumination.

If one looks at black-and-white photographs taken through retinex filters, one sees a dramatic difference in lightness for most objects between the photograph representing the short-wave system and either of the photographs representing the other two systems. And yet it is the comparatively small differences between the long-wave and the middle-wave lightnesses that are responsible for the experience of vivid reds and greens.

Such reliable and sensitive responsiveness to small lightness differences provides the basis for the colors seen under anomalous conditions far removed from those the eye has evolved to see. Two examples of interest are the color of a spot of light in a total surrounding area devoid of light and the spectrum of colors produced by a prism.

One can readily measure the flux at the eye from a spot of light in a void. By changing the flux it is possible to estimate the corresponding change in perceived lightness. What one finds is that the estimated lightness changes only slowly with enormous changes in flux. For example, decreasing the flux by a very large amount will be seen as a very small reduction in lightness. If the spot of light is composed of a narrow band of long wavelength, say 600 nanometers, one can expect all three cone receptors to absorb the radiation in some degree, but significantly more radiation will be absorbed by the long-wave cones than by the other two kinds. When the three values are read on a scale of perceived lightness, the three lightnesses are 9 on the long-wave system, 8.5 on the middle-wave system and 7.5 on the short-wave system [*see illustration on opposite page*]. This combination of lightnesses is seen as a light reddish orange, a color not commonly perceived under ordinary conditions unless the surfaces are fluorescent. The spectrum, a strikingly anomalous display, can be regarded as a series of three laterally displaced

HP measurement and computer advances

The HP-67. The most powerful hand-held programmable HP has ever built. Handles programs of up to 224 steps on a single magnetic card while larger programs—even thousands of steps—are easily handled with several cards. Automatic merging of keycodes increases the HP-67's program capacity. $450*.

The HP-97. All the capability of the HP-67, plus a built-in thermal printer. Small enough to fit very portably into your briefcase. $750*.

The HP-67/HP-97 software library: turnkey solutions to your computational problems

...whether you work in astronomy, surveying, computer science, accounting, thermodynamics, biology, probability, medical science, number theory, machine design, real estate, aviation, marine navigation, solid state devices, behavioral science, dynamics, differential equations, quality assurance, loans, optics, statistics, (or you name it).

We even have a program that converts your calculator into a stopwatch. And one that computes your biorhythm. There is one that tells time from the length of your shadow. And one that solves the classic coconut problem.

Or you can create your own problem solver. With the clear instructions provided in the Owner's Handbook, it is so easy you'll be writing programs the first day. And once you've keyed in a program you can permanently record it on a magnetic card.

Here's how to build your personal program library:

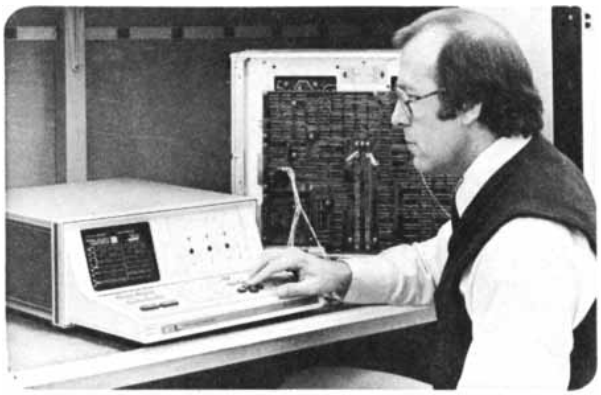
- With your new HP-67 or HP-97 you'll receive a Standard Application Pac containing 15 programs from a variety of fields. These are prerecorded on magnetic cards, which means that in less than two seconds your calculator will be ready to solve complex problems at the touch of a few keys.
- Choose from 10 application pacs each containing an instruction manual and up to 25 prerecorded programs in specific fields such as math, electrical and mechanical engineering, statistics, and surveying. Applications pacs are $35* each.

- Choose from more than 1000 programs in the HP-67/HP-97 Users' Library, containing programs contributed by HP-67 and HP-97 owners as well as by Hewlett-Packard engineers. These programs are not prerecorded, but consist of step-by-step instructions for keying them into your calculator. They're only $3* each.

- Choose from 40 new Users'-Library Solution Books. Each is a collection of 10 to 15 Users' Library programs in fields ranging from aeronautical engineering to thermal and transport sciences. Solutions Books are $10* each.

If you're curious about the range of HP Users' Library programs, send us this coupon. We'll send you a free catalog listing all the contributed programs as well as the contents of each Solutions Book. Our working combination of powerful calculators and comprehensive programs can deliver accurate solutions to your problems—painlessly.

extend your possibilities.



For those who design or debug digital system hardware and software, this new logic analyzer is a major breakthrough.

The new HP 1615 is the first logic analyzer that can make state and timing measurements simultaneously and relate one to the other. It is also the first that can detect timing glitches, strip them from captured data, and use them to trigger further analyses.

Digital machines operate both synchronously and asynchronously. A machine's program is implemented in synchronous mode by system clocks that change the state of the machine. Logic state analyzers—introduced by Hewlett-Packard about four years ago—do a great job of measuring synchronous activity: they monitor program flow and tell the designer when machine execution deviates from the program.

But much of the activity of a digital system is asynchronous—as when the program calls for something to happen between clock pulses, e.g. to move data on an I/O structure or through an address decoder. If such an activity is not carried out on the specified time schedule, the system malfunctions. Timing analyzers are used to measure asynchronous activity.

Even with separate state and timing measurements in hand, the designer does not have a really good view of overall system activity until he can relate any asynchronous fault to the state of the machine at the time of the fault. And that's been an extremely tricky proposition, one that has been left to the designer's ingenuity and intuition.

The HP 1615 performs both state and timing analyses *simultaneously* and *interactively*. Because of this dual capability, it can capture any asynchronous activity that occurs at a specific point in the program; identify which program step a system was executing at the time of malfunction; or monitor handshake sequences at specific points in program execution. Before the introduction of the HP 1615, no single instrument had been able to make these objective measurements.

Then there's the matter of glitches, those digital gremlins that confuse the machine's logic. Because they are transient and unpredictable, problem-causing glitches are extremely difficult to distinguish from harmless ones. Now the HP 1615 does it all: (1) captures system activity, both synchronous and asynchronous, simultaneously, (2) captures related data on surrounding lines, (3) strips the glitches out of this data, and (4) uses the glitches as part of the trigger for state, timing, or combined state and timing analyses.

The HP 1615 can be configured, through simple keyboard entries, as a 24-bit state analyzer, or as an 8-bit timing analyzer, or as a combined 16-bit state and 8-bit timing analyzer. It can trace events in as many as 24 channels at rates to 20 MHz, capturing data in a memory that is 24 bits wide and 256 words deep.

Although most system designers and debuggers will declare the HP 1615 to be priceless, it can be purchased for $6800*.

HEWLETT hp PACKARD

1503 Page Mill Road, Palo Alto, California 94304

For assistance call: Washington (301) 948-6370, Chicago (312) 255-9800, Atlanta (404) 955-1500, Los Angeles (213) 877-1282.

Mail to: Hewlett-Packard, 1503 Page Mill Road, Palo Alto, CA 94304
Please send me further information on

( ) HP-67/HP-97 software library
( ) HP 1615 logic analyzer

Name _____
Company _____
Address _____
City _____ State _____ Zip _____

*Domestic U.S. prices only

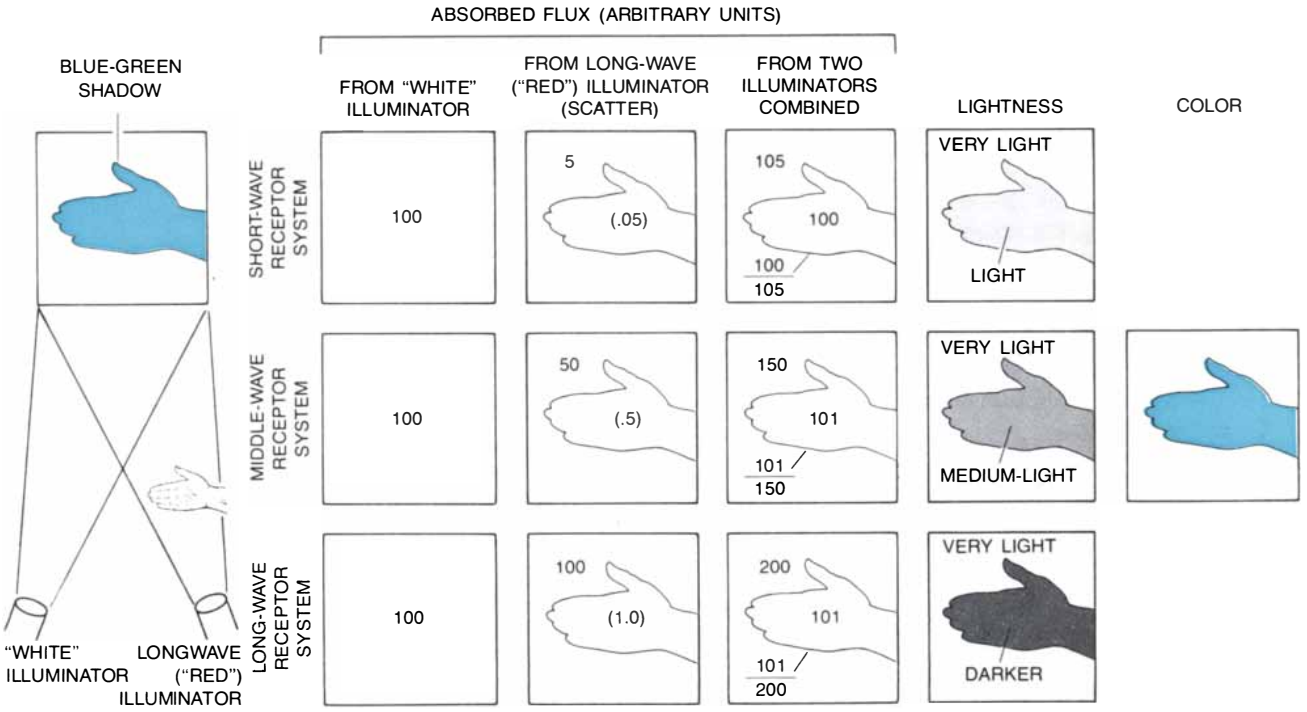
00840

continuous gradients involving both the properties of spots and the properties of areas. From these properties it is possible to predict the colors of the spectrum, whereas it is not possible, as we have seen, to attribute a specific spectral composition to the radiance from a colored area in everyday life.

Perhaps the first observation pointedly relevant to the mechanism of color formation in images is not Newton's spectrum but the phenomenon of colored shadows, described in 1672 by Otto von Guericke. "This is how it happens," he wrote, "that in the early morning twilight a clear blue shadow can be produced upon a white piece of paper [by holding] a finger or other object . . . between a lighted candle and the paper beneath." This important experiment, we now know, depicts an elementary example of generating three different lightnesses on the three receptor systems. A diagram of this experiment with long-wave ("red") light and white light appears below. Here the color of the shadow is blue-green. The diagram shows that the triplet of lightnesses in the shadow corresponds to the blue-green color one would predict for it from its position in lightness-color space.

One can now understand the red-and-white images of our early work as a procedure that carries the colored shadow to a richly variegated family of colors no longer in shadows but in images. The colors seen in a red-and-white projection can be readily predicted by extending the analysis followed in predicting the color of von Guericke's shadow. To demonstrate this point we reproduce on page 128 the "red" and "green" separation images used in making a red-and-white multicolored projection. (In this demonstration no attempt is made to reproduce the colors seen in the actual multicolored image.) The red-and-white projection was photographed through long-, middle- and short-wave retinex film-filter combinations. The three images are reproduced below the pair of long- and middle-wave separation images that were superposed to make the red-and-white image. The significant point is that when the eye views the red-and-white images on the screen with its own retinex system, it is provided with a triplet of lightnesses for each part of the scene that resembles the triplet it would obtain if it viewed the original scene directly. In this important meeting point of the blue-green shadows with the colored images, provided by the red-and-white display, the extended taking and multiplication of ratios determine the lightness of each small area. Finally, all these principles are applied in everyday ternary vision, which creates a distinct lightness image for each of the three sensitive systems and compares them in order to generate color.

The train of interlocking concepts and experiments started 25 years ago with the observation that the relative energies of the red-and-white projectors can be altered without changing the names of the various colors. This observation negated the simplistic explanation in terms of contrast, fatigue and surround and led to the fundamental concept of independent long- and short-wave image-forming systems that ultimately evolved to the concept of three independent retinex systems and to the Mondrian demonstration. The concept of the percentage of available light on each wave band as a determining variable and the technique of measuring it

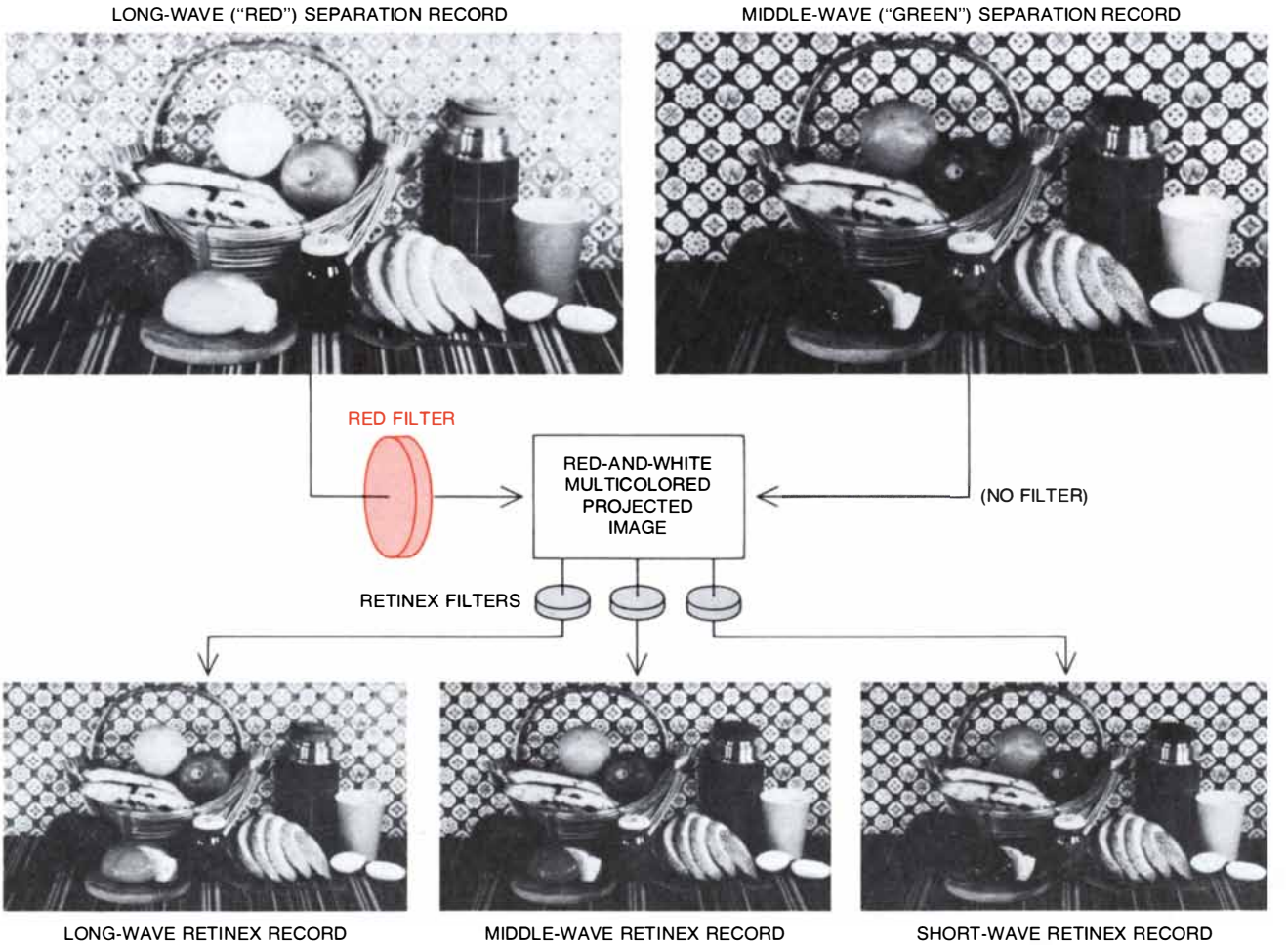


BLUE-GREEN COLORED SHADOW is seen when a hand or some other object is placed in the beam of a projector that is sending long-wave ("red") light to a screen while the screen is illuminated by a beam of white light. The author regards Otto von Guericke's description in 1672 of seeing colored shadows made by candlelight as the first observation pointedly relevant to the mechanism of image and color formation. In the analysis at the right it is assumed that one projector sends white light to the screen. The other projector, equipped with a red filter, sends only long wavelengths to the screen. Assume that the white light contributes 100 arbitrary units of flux to each of the short-, middle- and long-wave receptors. The long-wave flux is absorbed by the three receptor systems in different proportions: 100 units are absorbed by the long-wave system, 50 by the middle-wave system and five by the short-wave system. (A small amount of scattered long-wave light also appears in the shadow.) The third column of boxes shows the combined amounts of flux from both sources absorbed by each receptor system. The fractions represent the ratio at edges of the flux from within the shadow divided by the flux from outside. The fourth column shows the lightness on each receptor system. The lightness of the lightest place in the scene for each receptor system will be near the top of the lightness scale, being determined by the flux of radiant energy in the same way that a spot has its lightness determined by flux. Triplet of lightnesses within the shadow falls in the region of color space that the eye perceives as being blue-green.

evolved to the concept that lightnesses maintain an independent rank order on long- and short-wave bands. This measuring technique in turn evolved from a projected black-and-white image to an arrangement of colored papers in the color Mondrian. The manifest stability and constancy of the lightnesses of all the papers of the Mondrian when a single wave band illuminates it with varying intensity dramatizes the concept that every colored paper has three reflectances on three wave bands and that these reflectances are somehow connected with the biological characteristic: lightnesses.

A black-and-white Mondrian taught that nonuniformity of illumination, size and shape of area and length of edges were basically irrelevant to lightness. What was needed was a far-reaching, edge-reading arithmetic: the sequential product of ratios at edges. For the color Mondrian the ratio at edges was early recognized as requiring a ratio of the integrals of the product at each wavelength of the absorbance of the cone pigment times the reflectance of the colored paper times the illuminants. Separate integrals were taken over the wave bands of the three cone pigments. In a long series of binocular comparison-and-selection observations the quantity satisfying the integral was shown to be impressively well correlated with lightness, particularly after the realization that the scale, or spacing, of the reflectance integral should be made to correspond with the spacing of the biological quantity lightness. This led to the designation "scaled integrated reflectance" as the external partner to which the retinex system relates the internal partner: constructed lightness.

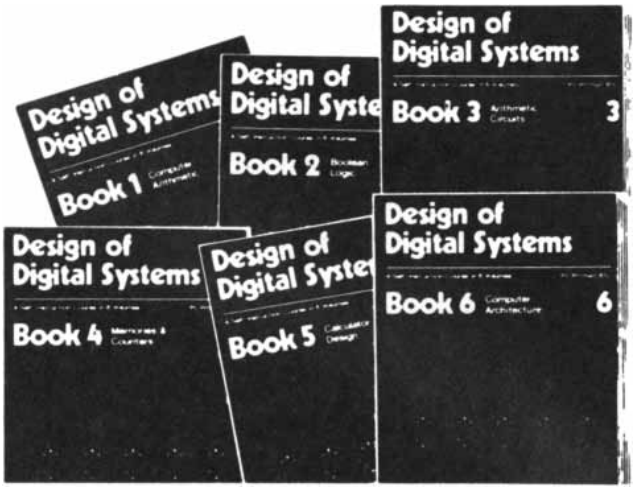
Color can be arranged in the lightness solid with long-, middle- and short-wave axes of lightness. All visible colors reside in this solid independent of flux, each color having a unique position given by the three axial values of lightness. It should be remembered that the reality of color lies in this solid. When the color Mondrian is nonuniformly illuminated, photographed and measured, reflectance in the photograph no longer correlates with the color but the lightness does. The three sets of ratios of integrals at edges and the product of these integrals within a set emerge as the physical determinants in the partnership between the biological system and areas in the external world.



RETINEX RECORDS OF RED-AND-WHITE projections show that red-and-white images produce a triplet of lightnesses for each part of the scene that are consistent with the observed color sensations. The two photographs in the top half of this illustration are reproductions of the long-wave (*left*) and middle-wave (*right*) separation records taken of the original still life. The long-wave record was projected onto a screen with a long-wave (red) filter in the beam of light. The middle-wave record was projected in superposition onto the same screen in the light of a tungsten-filament lamp. Three retinex photographs were then taken of projected images appearing on screen. The retinex records are reproduced in the bottom part of the illustration: long-wave at the left, middle-wave in the middle and short-wave at the right. The colors seen in red-and-white projections are those one would expect from their triplets of lightnesses. The apple is light on the long record and darker in the middle and short records. The orange is lightest on the long record, intermediate on the middle record and darkest on the short. It is impressive that with his own retinex systems the observer can see a blue cup, a brown straw basket and pale yellow bananas with lightness differences so small as to challenge photoengraving process used to reproduce photographs.

Understanding Digital Electronics

New teach-yourself courses



Design of Digital Systems is written for the engineer seeking to learn more about digital electronics. Its six volumes — each 11-1/2" x 8-1/4" are packed with information, diagrams and questions designed to lead you step-by-step through number systems and Boolean algebra to memories, counters and simple arithmetic circuits, and finally to a complete understanding of the design and operation of calculators and computers.

The contents of Design of Digital Systems include:

Book 1 Octal, hexadecimal and binary number systems; conversion between number systems; representation of negative numbers; complementary systems; binary multiplication and division.
Book 2 OR and AND functions; logic gates; NOT, exclusive-OR, NAND, NOR and exclusive-NOR functions; multiple input gates; truth tables; De Morgans Laws; canonical forms; logic conventions; Karnaugh mapping; three-state and wired logic.
Book 3 Half adders and full adders; subtractors; serial and parallel adders; processors and arithmetic logic units (ALUs); multiplication and division systems.
Book 4 Flip flops; shift registers; asynchronous and synchronous counters; ring, Johnson and exclusive-OR feedback counters; random access memories (RAMs) and read only memories (ROMs).
Book 5 Structure of calculators; keyboard encoding; decoding display data; register systems; control unit; program ROM; address decoding; instruction sets; instruction decoding; control program structure.
Book 6 Central processing unit (CPU); memory organization; character representation; program storage; address modes; input/output systems; program interrupts; interrupt priorities; programming; assemblers; computers; executive programs; operating systems and time sharing.

Digital Computer Logic and Electronics is designed for the beginner. No mathematical knowledge other than simple arithmetic is assumed, though the student should have an aptitude for logical thought. It consists of four volumes — each 11-1/2" x 8-1/4" — and serves as an introduction to the subject of digital electronics. Everyone can learn from it — designer, executive, scientist, student, engineer.

Contents include: Binary, octal and decimal number systems; conversion between number systems; AND, OR, NOR and NAND gates and inverters; Boolean algebra and truth tables; De Morgans Laws; design of logic circuits using NOR gates; R-S and J-K flip flops; binary counters, shift registers and half adders.

In the years ahead the products of digital electronics technology will play an important part in your life. Calculators and digital watches are already commonplace. Tomorrow a digital display could show your automobile speed and gas consumption; you could be calling people by entering their name into a telephone which would automatically look up their number and dial it for you.

These courses were written by experts in electronics and learning systems so that you could teach yourself the theory and application of digital logic. Learning by self-instruction has the advantages of being faster and more thorough than classroom learning. You work at your own pace and must respond by answering questions on each new piece of information before proceeding.

After completing these courses you will have broadened your career prospects and increased your fundamental understanding of the rapidly changing technological world around you.

The six volumes of Design of Digital Systems cost only: **$19.88**

And the four volumes of Digital Computer Logic and Electronics cost only: **$14.88**

But if you buy both courses, the total cost is only: **$29.90**

a saving of over: **$5.00**

SEVEN-DAY MONEY-BACK GUARANTEE: If you are not satisfied with your Cambridge course, return it within 7 days for a full refund.

To order your books, complete the order form below and send it together with your check or money order to GFN Industries, Inc., 6 Commercial Street, Hicksville, N.Y. 11801.

To: GFN INDUSTRIES, INC.
6 COMMERCIAL STREET, HICKSVILLE, NY 11801

Please send me:
___ Sets of Design of Digital Systems $19.88 _____
___ Sets of Digital Computer Logic & Electronics $14.88 _____
___ Sets of both courses $29.90 _____
Sales tax (N.Y. residents) _____
Shipping and handling $2.50 per set _____
Enclosed is check/mo (payable to GFN Industries, Inc.)

Total $_____

Name _____

Address _____

City/State/Zip _____

Prices include overseas surface mail postage. SA12B

Disclinations

They are defects in warped or twisted materials. Unlike a related defect, the dislocation, they do not appear in ordinary crystals, but they abound in thin periodic structures such as virus shells

by William F. Harris

A crystalline solid is often described as a perfectly regular array of atoms or molecules. Real crystals, however, are never perfect; in the arrangement of the atoms or molecules there are always flaws. It might seem that most such defects would be minor disruptions that could safely be ignored, but such is not the case. Imperfections in the crystal lattice are responsible for many of the physical and chemical properties of solids. Moreover, the nature of the imperfections themselves is of interest. It turns out that a regular structure can go wrong in only a few, well-defined ways.

A kind of crystal imperfection that has been studied intensively, particularly in metals and semiconductors, is the dislocation. It results from the translation, or linear displacement, of one part of the crystal with respect to another part. In recent years a related defect has come under investigation; it is called a disclination. In a disclinated solid one part of a structure is displaced by a rotation rather than by a translation.

The study of disclinations has its origins in the physics of crystalline solids, but as it happens disclinations are seldom observed in ordinary three-dimensional crystals such as those of metals. They do appear in the arrays of oriented molecules called liquid crystals. What is more, they are important structural elements in many ordered materials other than conventional crystals, such as the protein coats of viruses. Disclinations can even be observed in the pattern of fingerprints, in the pelts of striped animals such as zebras and in basketwork.

The geometry of both dislocations and disclinations can be illustrated with the aid of an imaginary material that has no microscopic structure and has ideal mechanical properties. The first step in creating a suitable defect is to slice halfway through a sphere of the material, so that the cut terminates along a diameter of the sphere. A thin half disk of the same material is then forced into the slit. If the inserted half disk is held in place by an ideal glue, all external forces acting on the sphere can be removed, although internal strains will persist. In this deformed body the two faces of the original slice remain parallel at all points; they change only their relative location, not their orientation, and so the defect created in this way is called a dislocation.

A disclination can be formed by a similar procedure. Again we slice halfway through a sphere, but instead of separating the cut faces by a parallel displacement we pry them apart at the circumference of the sphere and insert a wedge of material. (To be precise, the piece inserted is not a wedge but a sector of a sphere.) Once again the inserted material is glued in place, leaving an internally strained body. In this case the faces of the slice have been rotated, and it is not only their relative position that has changed but also their relative inclination, hence the term disclination. The defect was first called a "disinclination" by Sir Charles Frank of the University of Bristol, but the name was changed at the urging of a professor of English who was disinclined to grant a new meaning to that word.

If the glue employed in these thought experiments is perfect, the surfaces that are cut and rejoined cannot be identified. Indeed, from an examination of the deformed body it is not possible to tell where the slice was made or where the new material was inserted. The stresses and strains caused by the insertion of new material are distributed smoothly throughout the body. With one exception there are no singularities, or points specially distinguished from their surroundings, anywhere in the sphere.

The exception is the diameter defined by the original slice. If a probe measuring internal stress could be inserted from the surface of the sphere, it would reveal a smooth and continuous increase in stress until it reached the center. At the center, and all along the diameter where the cut terminated, the stress becomes infinite. The infinity arises no matter how thin the inserted half disk or wedge is. A larger deformation increases the stress everywhere else in the body, but the stress is infinite along the diameter even with an arbitrarily small insert. The influence of the defect is felt throughout the body, but it is convenient to regard the defect as being localized along the line of infinite stress; the line is called the dislocation line or the disclination line.

Real materials, of course, cannot withstand infinite stress, and so these thought experiments could not be performed in a real substance. Indeed, intuition suggests that a half disk or a wedge could not be forced all the way to the center of the cut sphere without breaking or tearing the object. The way to avoid this difficulty is to drill a hole through the sphere along the diameter, leaving a hollow core. The sphere is then formally described as a torus. Dislocations and disclinations can be created as in the sphere, but now the stress varies smoothly from point to point throughout the structure. There are no points or lines where the stress becomes infinite. The dislocation line or disclination line passes through the core, which can be left empty or filled with material after the defect has been created.

The deformations I have described each result from a simple movement of the cut surfaces. There are many other possible displacements. In a torus that has been sliced from the perimeter through to the core a dislocation is generated by any relative, parallel displacement of the cut faces and a disclination is formed by rotation of the two faces around any axis.

The large variety of possible dislocations and disclinations can be analyzed in terms of three general classes, defined by three mutually perpendicular axes. The kind of dislocation discussed above, in which the cut surfaces are pulled straight apart, results from translation along an axis perpendicular to the cut

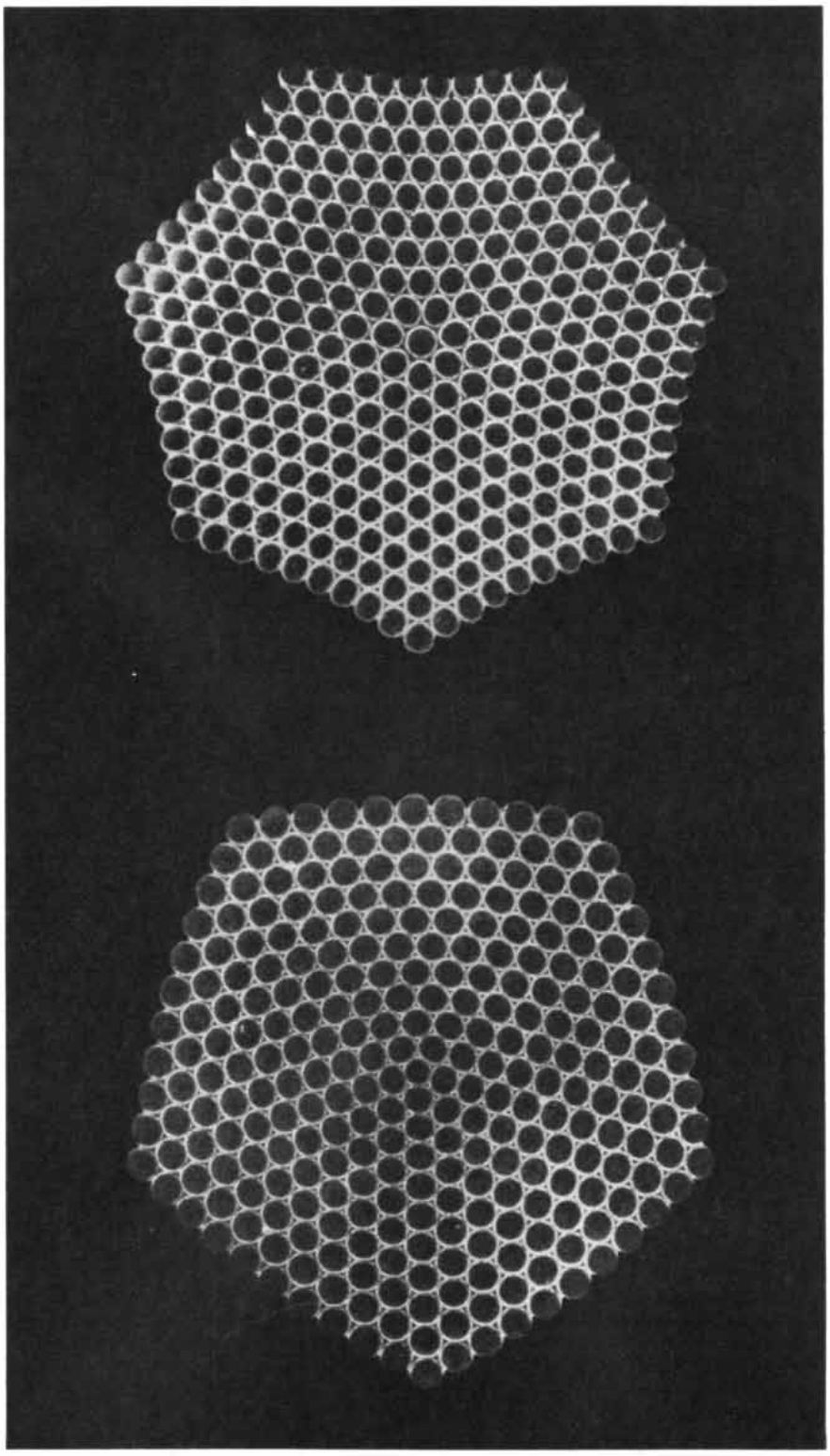
surfaces. It is called an edge dislocation. An edge dislocation also results when the faces are pushed together instead of being pulled apart; material must then be removed instead of added. Another edge dislocation arises from the displacement of one face along an axis that lies in the plane of the cut surfaces and passes through the axis of the torus; again the motion can be in either direction, toward the center of the torus or away from it. The remaining axis is parallel to the cut faces and also parallel to the axis of the torus itself. The result of a displacement along this axis (in either direction) is called a screw dislocation. A torus with a screw dislocation looks much like a common lock washer.

The disclination described above is generated by rotating the cut faces around the axis of the torus or sphere. Because of the shape of the resulting gap it is called a wedge disclination. In this case too the motion can be reversed, provided that a suitable wedge of material is removed.

Rotations around either of the two axes perpendicular to that axis result in deformations called twist disclinations. One of these axes lies in the slice and passes through the center of the torus. Although rotation around this axis is classified as a twist disclination, it opens a gap that must be filled by a wedge of material, which leads to an unfortunate confusion in terminology. The twist character of the remaining disclination is more readily apparent. It is generated by rotation of the faces around an axis perpendicular to them, and in the simplest case the motion is equivalent to that employed in wringing a washcloth. The cut surfaces rotate by merely sliding past each other, and no material need be added or removed.

The orientation of these axes is fixed, but their positions are not. In a wedge disclination, for example, the axis of rotation must be parallel to the axis of the torus or sphere, but it need not be coincident with it. The axis need not even pass through the structure. Moving the axis changes the appearance of the disclination, but not its basic form or its classification. Moreover, defects in solids are by no means confined to motions along or around a single axis. Any relative motion of the two cut faces is possible, as long as the faces themselves are not distorted. All such motions, however, can be decomposed into combinations of these six basic translations and rotations. It should be emphasized that dislocations and disclinations are classified according to their structure and without regard for the process by which the defect is formed. There are means for creating such defects other than the ones described above.

We have considered dislocations and disclinations in a continuous, structure-

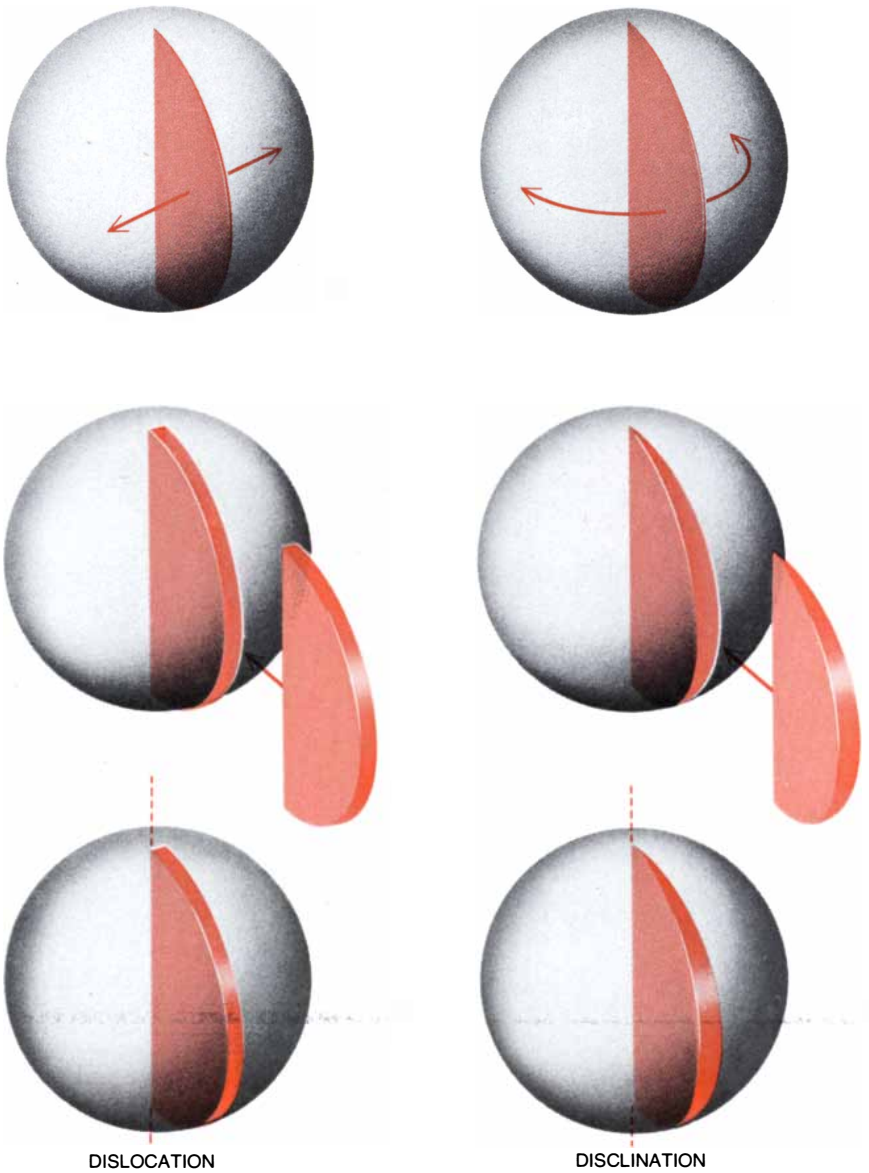


WEDGE DISCLINATIONS alter the rotational symmetry of a lattice of floating bubbles. The lattice is assembled by herding small, uniform bubbles into a close-packed array. Like any close-packing of spheres, the array has sixfold rotational symmetry: each bubble is surrounded by six other bubbles. In the bubble raft at the top a disclination was created by a process equivalent to prying apart two rows of bubbles in order to open a gap extending from the circumference to the central bubble, then inserting a wedge 60 degrees wide (an equilateral triangle of bubbles). The symmetry of the structure is thereby changed to sevenfold: the central bubble has seven neighbors. The raft at the bottom was also created by transforming a hexagonal raft, but in this case a 60-degree wedge was removed and the exposed edges were rejoined. The resulting array has fivefold symmetry. The positions where a sector was inserted or removed cannot be identified. The stresses introduced by the deformation are distributed throughout the structure. Nevertheless, it is convenient to regard the defect as being at the central bubble.

less medium, but in fact there are no such materials. Since all substances are made up of atoms, they all have some internal structure at a microscopic scale; in the majority of solids the arrangement of the atoms is spatially periodic. Such structures are distinguished by their symmetry. For the study of dislocations and disclinations two kinds of symmetry are important: translational and rotational.

Symmetries are defined by "symmetry operations," which are motions or other transformations that leave the form or the appearance of an object unchanged. The nature of such operations can be conveniently explored in a simple cubic lattice in which all the atoms are found at the vertexes of cubes. Translational symmetry is exhibited when such a lattice is shifted in a direction parallel to one set of lattice lines. After moving a distance equivalent to the spacing between two lattice lines the entire structure comes into coincidence with itself. If the boundaries can be ignored, or if the lattice is imagined to be an infinite one, the new configuration cannot be distinguished from the old one. Since the motion leaves the structure unchanged, it is by definition a symmetry operation for that lattice.

Rotational symmetry becomes apparent when the entire structure is turned around, say, any selected lattice line, which is designated a symmetry axis. In the case of a simple cubic lattice the structure comes into coincidence with itself after turning 90 degrees. Such points of equivalence are encountered four times in the course of a complete revolution, and so the crystal is said to have fourfold rotational symmetry. Four- and sixfold symmetries are the commonest ones in crystals, and no crystal has a rotational symmetry of an order higher than sixfold. Onefold symmetry is a trivial case: any object comes into coincidence with itself once at the completion of a revolution.

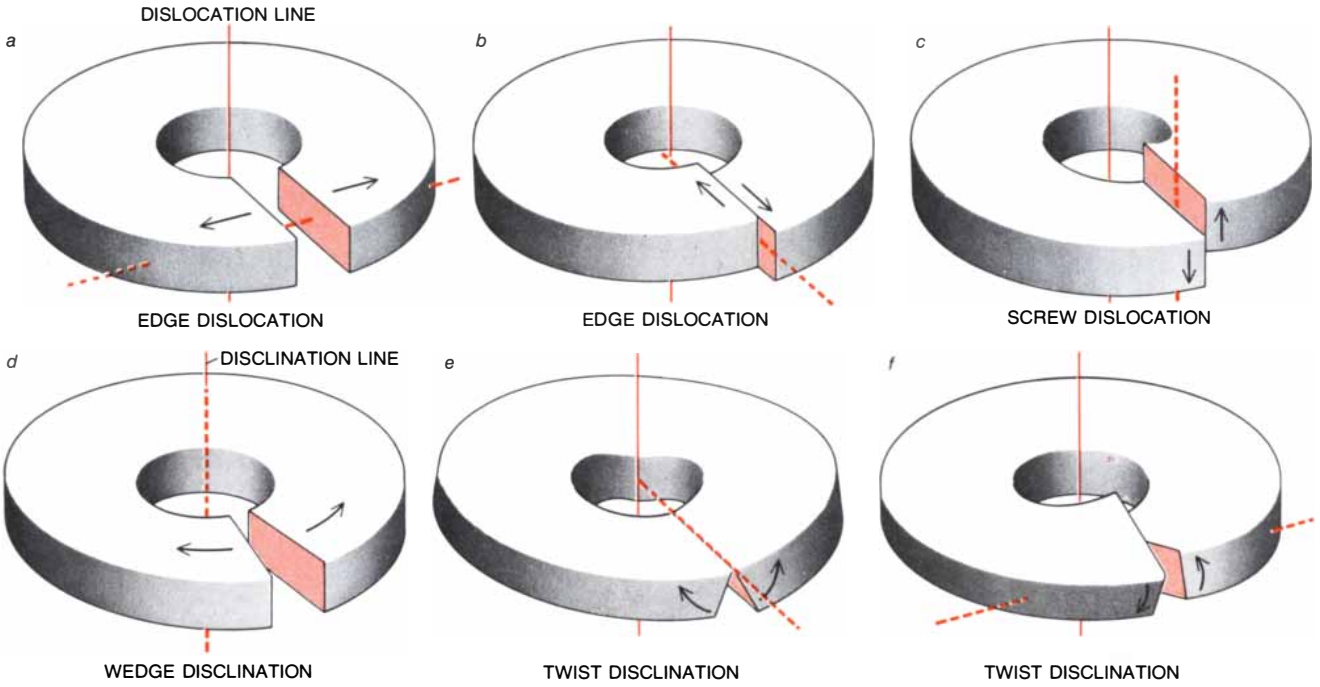


TRANSLATION AND ROTATION in a rigid body give rise respectively to a dislocation and a disclination. Both defects can be created by first cutting partway through a body; here the cut is made halfway through a sphere, so that the line on which the cut terminates is a diameter of the sphere. In the deformed body this line becomes the dislocation line or the disclination line. If the cut surfaces are pulled straight apart, so that they remain parallel, and the gap between them is filled with a half disk of the same material, the result is a dislocation (*left*). Rotating the cut surfaces and inserting a wedge yields a disclination (*right*). If the sphere has no periodic structure, a half disk or a wedge of any size can be inserted. Stress, however, increases toward the diameter, and along the dislocation line or disclination line itself the stress becomes infinite.

The symmetries of structured materials impose significant constraints on dislocations and disclinations. In an unstructured medium a wedge disclination can be formed by inserting an arbitrarily thin sector; no matter what the angular width of the wedge is, the surfaces can be joined without seams. In a crystalline structure with finite symmetry, however, the insertion of such an arbitrarily thin sector would create a conspicuous discontinuity. Along at least one of the surfaces where the new material joins the old the lattices would not be in register. Hence the defect would not qualify as a disclination.

The only instances where the continuity of the crystal lattice can be preserved are those in which the rotation of the disclination is a symmetry rotation of the lattice. Thus in a simple cubic lattice the cut surfaces can be rotated 90 degrees and a 90-degree wedge of unstrained material can be inserted. When the body deformed in this way is subsequently allowed to relax, the lattice is continuous at all points across the join. Some lattice lines may be smoothly bent, but none are broken. The same constraint applies to twist disclinations. In a cubic crystal two faces that are rotated around an axis passing perpendicularly through them (like two wheels on a single shaft) can match only if the rotation is by 90 degrees or a multiple of 90 degrees. In crystals with sixfold symmetry the minimum rotation for all classes of disclinations is a sixth of a circle, or 60 degrees.

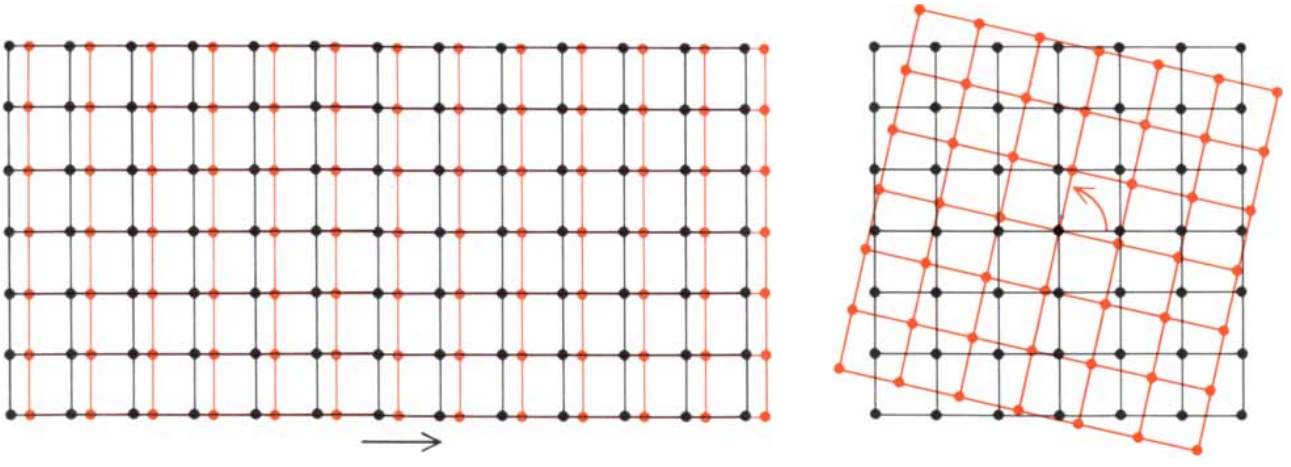
Dislocations in crystals are also constrained by the periodic structure of the lattice, but the governing symmetry is translational rather than rotational. The atoms in the two cut surfaces can be aligned only if the displacement is a translational symmetry operation of the lattice. The minimum displacement, therefore, is in the simplest case the distance between two adjacent atoms. Since that displacement is quite small, it can often be introduced without de-



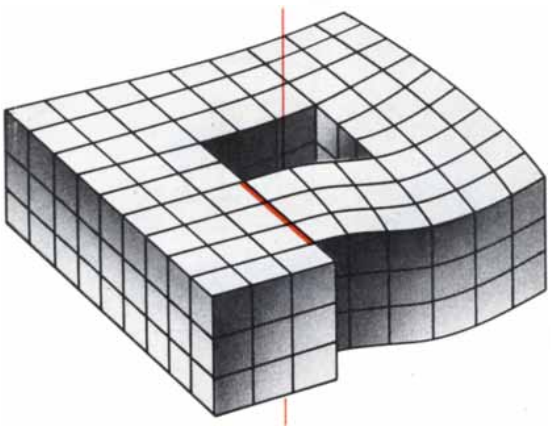
CLASSIFICATION OF DEFECTS is determined by the relative motion of the two cut surfaces. In order to avoid the infinite stresses that develop in a deformed sphere, the defects are defined in a torus that has been cut from the outer edge to the hollow core. Any parallel displacement, or in other words any motion that has no rotational component, yields a dislocation. There are three axes for such motion; two are perpendicular to the axis of the torus and define edge dislocations (*a, b*) and one is parallel to the axis of the torus and defines a screw dislocation (*c*). There are also three axes of rotation associated with disclinations. Rotation around the axis of the torus gives rise to a wedge disclination (*d*); rotations around axes perpendicular to the axis of the torus give rise to defects that are classified as twist disclinations (*e, f*) even though one of them (*e*) opens a wedge-shaped gap in the body. Any relative motion of the cut surfaces can be decomposed into translations along and rotations around these six axes or axes parallel to them. All the motions shown can also be reversed.

stroying a comparatively rigid structure, and dislocations are common defects of many crystals. The minimum rotation for a disclination, on the other hand, is equal to the fundamental symmetry rotation of the crystal. Since no crystal has a symmetry of an order higher than sixfold, the smallest wedge that can be inserted, or the smallest twist that can be applied, is 60 degrees. Such a large deformation would give rise to enormous stress, which ordinary crystals cannot withstand without breaking. As a result disclinations of the kinds I have discussed so far are not observed in three-dimensional crystalline solids. We must look for them instead in other periodic structures; in these materials disclinations that embody rotations by as much as 360 degrees are quite common.

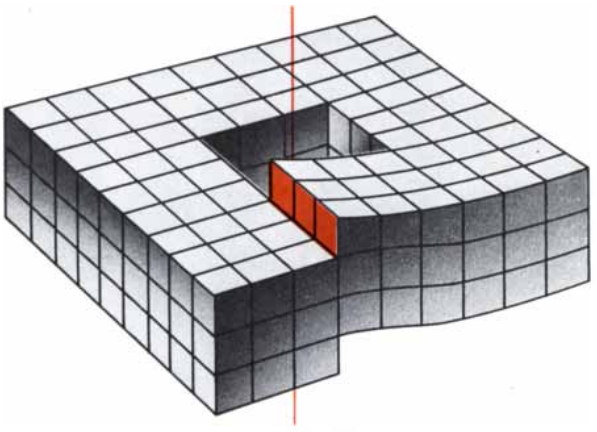
It is not always a simple matter, from the mere examination of a crystal defect, to deduce the nature of the displacement that produced it. In the case



SYMMETRIES of spatially periodic materials confine the relative motions that generate dislocations and disclinations to those that are symmetry operations of the material. Translational symmetry (*left*) is observed in straight-line displacements of the lattice. When the structure is shifted by an amount equal to the distance between two lattice lines, it comes into coincidence with itself; if the boundaries of the lattice are ignored, the new position cannot be distinguished from the old one. Rotational symmetry (*right*) describes a similar equivalence observed when the structure is turned around some axis of symmetry. A cubic lattice comes into coincidence with itself after a rotation of 90 degrees; since four such points of equivalence are encountered in a full revolution, the lattice has fourfold rotational symmetry.

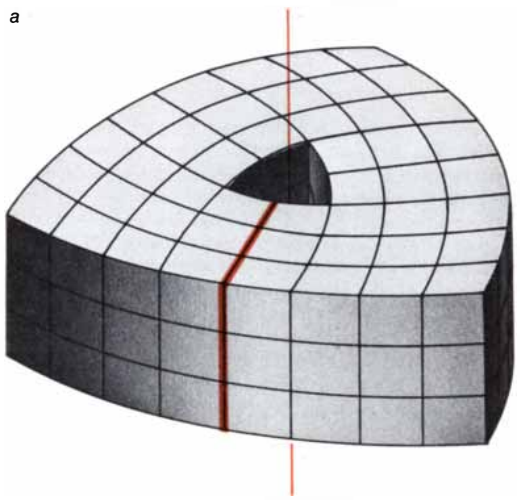


EDGE DISLOCATION



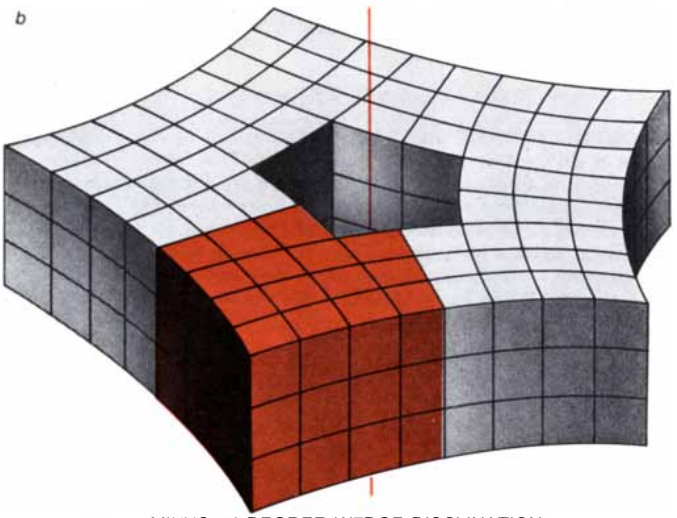
SCREW DISLOCATION

DISLOCATIONS in spatially periodic materials must conform to the translational symmetry of the lattice. Here edge and screw dislocations are created in a hypothetical cubic crystal by cutting along one plane of the lattice (*color*) and displacing the cut surfaces by one unit of the lattice spacing. By this procedure the continuity of the lattice is preserved across the rejoined surfaces. The dislocation core, which encircles the dislocation line (*colored line*), can be hollow as shown, but in real crystals it is usually filled with more or less ordered material.



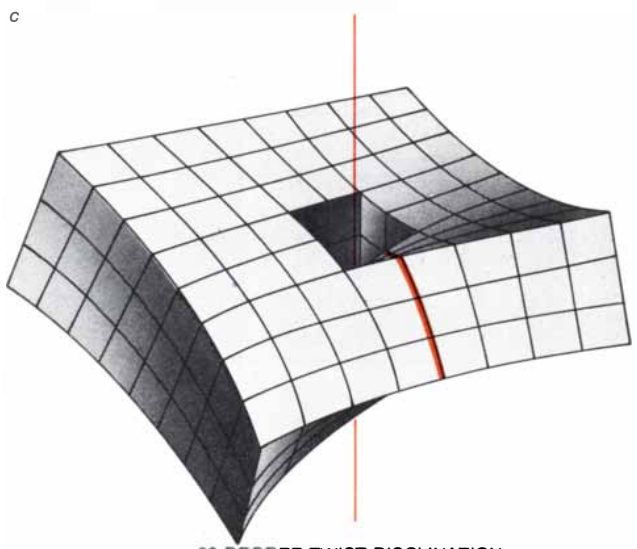
a

PLUS-90-DEGREE WEDGE DISCLINATION



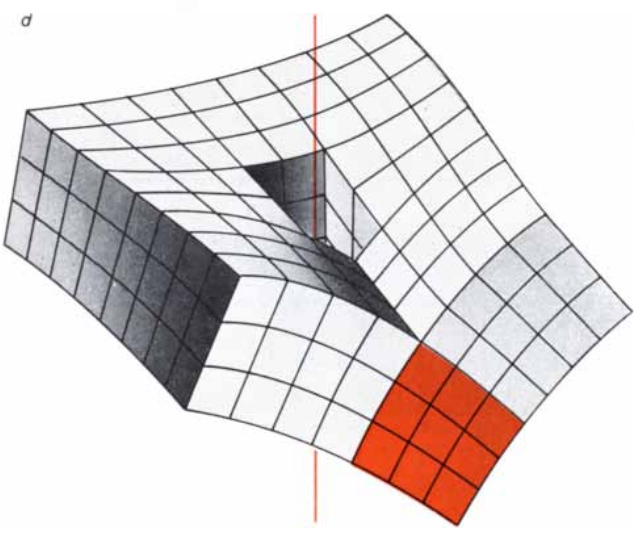
b

MINUS-90-DEGREE WEDGE DISCLINATION



c

90-DEGREE TWIST DISCLINATION



d

90-DEGREE TWIST DISCLINATION

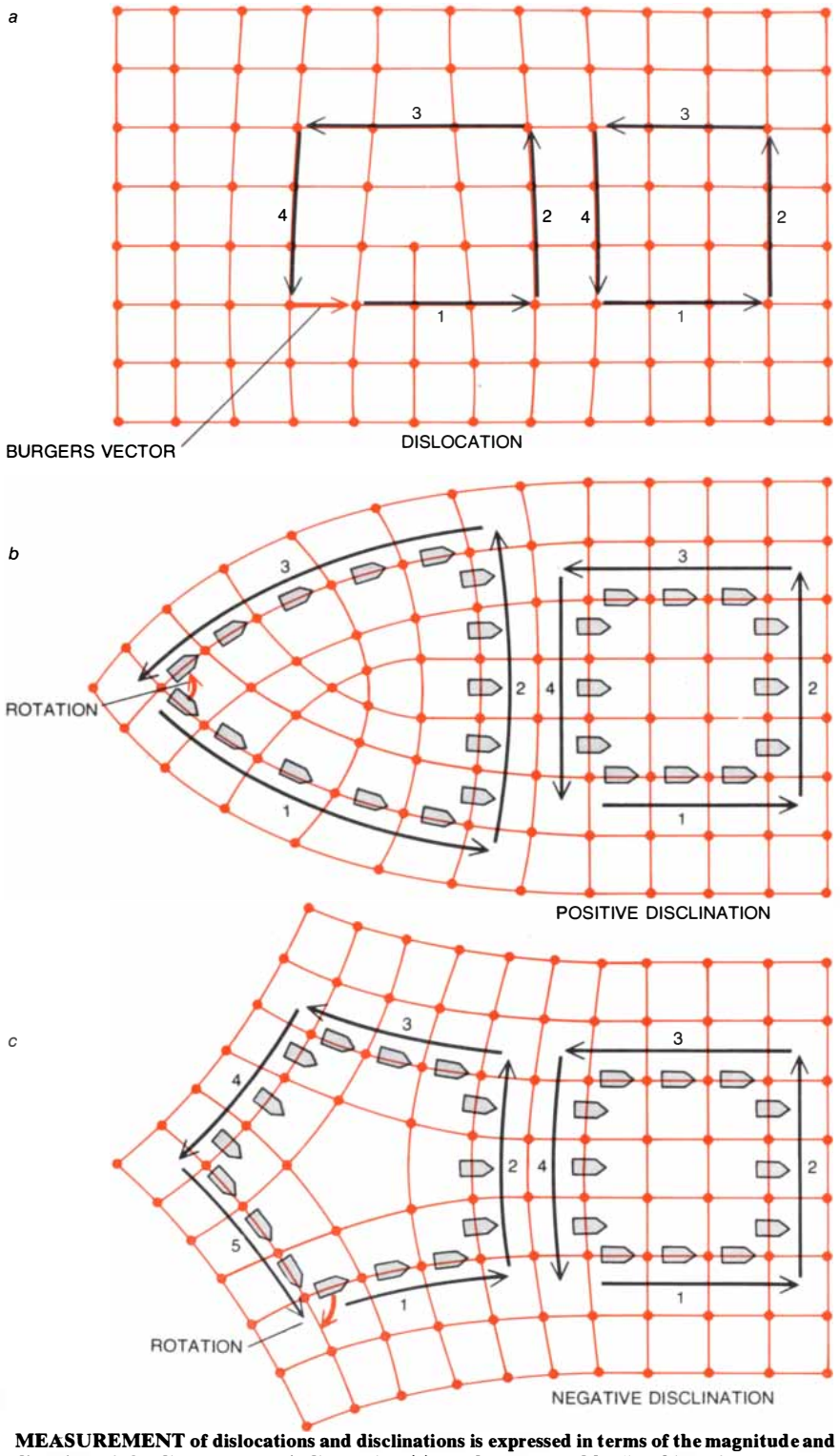
DISCLINATIONS in structured materials are possible only if the rotation required to generate the defect is a symmetry operation of the lattice. In a cubic lattice, with fourfold symmetry, the minimum rotation is 90 degrees. Wedge disclinations can be created by removing a 90-degree section of material (*a*) or by inserting a similar section (*b*). Twist disclinations are formed by a 90-degree rotation around an axis that passes perpendicularly through the cut faces (*c*) or by the same rotation around an axis that lies in the plane of the cut and passes through the axis of the torus at right angles (*d*). Rotations by any angle other than 90 degrees (or multiples of 90 degrees) would result in a discontinuity: the lattices on opposite sides of the cut surface would not match. Because rotations of that magnitude introduce large stresses, disclinations cannot form in ordinary crystals. The disclination core, surrounding the disclination line, can be hollow or filled.

of dislocations a technique for classifying defects was devised by the Dutch physicist J. M. Burgers. One "walks" from atom to atom through the lattice along any path that would be a closed circuit if the lattice were perfect. In a cubic crystal, for example, the simplest path requires just four steps and describes a square. The corresponding "Burgers circuit" constructed around a dislocation fails to close; the last step does not return to the starting point but terminates at some new position. An arrow pointing from this position to the starting point is called the Burgers vector; it gives the magnitude and the direction of the displacement that generated the dislocation.

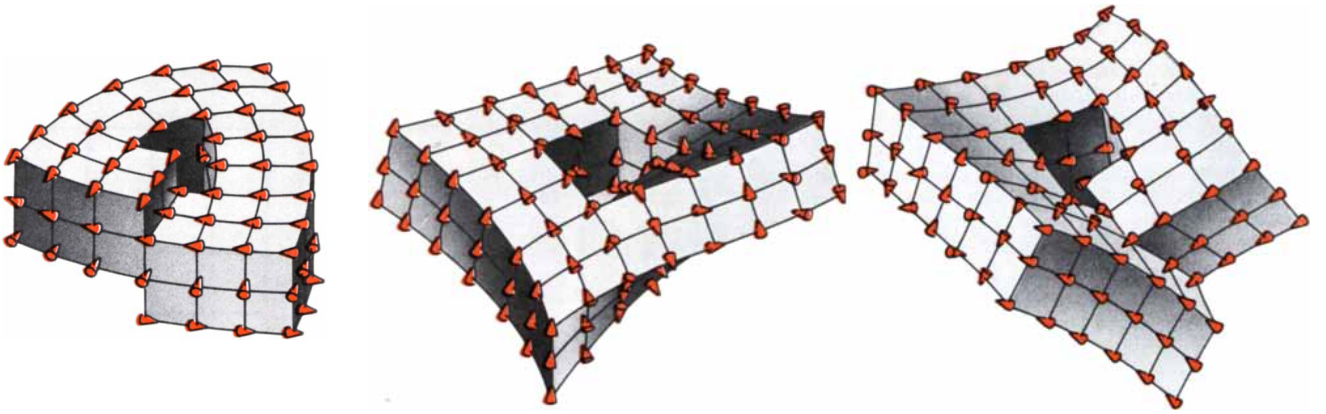
For measuring the rotation of a disclination a similar technique has been developed by Frank R. N. Nabarro of the University of the Witwatersrand. Once again a circuit (the Nabarro circuit) is selected, but instead of the number of steps it is the change in the orientation of the lattice that is significant. A particular lattice orientation is chosen; then the continuous change in orientation is followed as one moves around the circuit and back to the start. The net change in orientation is equal to the rotation of the disclination enclosed by the circuit. (Strictly speaking, the angle should be measured in undistorted material.) If the change in orientation is in the same sense as the circuit (for example, if both are clockwise), then the defect is called a positive wedge disclination; such defects can be created by removing a wedge. If the senses of rotation are opposite, the disclination is negative and can be made by inserting a wedge.

In unstructured materials dislocations and disclinations are the only line defects possible. That is also the case for many periodic structures: the only effect of introducing a lattice is to confine the translations and rotations to symmetry operations of the undeformed material. In a special class of periodic structures, however, a third symmetry-related defect can appear. Just as the dislocation and the disclination are governed by translational and rotational symmetry, so the third defect is a manifestation of screw symmetry, the kind of symmetry characteristic of a threaded shaft. Screw symmetry can be interpreted as a combination of translation and rotation operations, and it might seem that the corresponding defect is merely a combination of a dislocation and a disclination. Actually it is a distinct kind of structural defect, which has been given the name dispiration.

One kind of structure in which dispirations can appear is a crystal made up of units, such as atoms, that do not have spherical symmetry. A crystal exhibiting screw symmetry can readily be constructed of such atoms. It can be made, for example, of planes of atoms with all



MEASUREMENT of dislocations and disclinations is expressed in terms of the magnitude and direction of the displacement. A dislocation (*a*) can be measured by "walking" from atom to atom along any circuit around the flaw. In an undeformed cubic lattice, for example, a circuit of 12 steps might be selected to bring one back to the starting point. If the circuit encloses a dislocation, however, the same 12 steps terminate at a new position. An arrow pointing from that new position to the starting point defines the dislocation; the arrow is called the Burgers vector. The measurement of a disclination involves a similar circuit, but what is significant is not the displacement from the origin but the change in the orientation of the lattice. A particular lattice orientation is represented at the start by an arrow (*pentagonal box*), which is carried around the circuit, always parallel to the local orientation of the lattice. In an undeformed crystal the arrow points in the same direction at the end as it did at the start, but if the circuit encloses a disclination, the arrow undergoes a net rotation, namely the rotation of the disclination itself. The arrow rotates in the same direction as the circuit itself if a wedge has been removed (*b*), but it rotates in the opposite direction if a wedge has been added (*c*). Rotation in the same sense as the circuit defines a positive wedge disclination, in the opposite sense a negative one. The disclinations shown have rotations of plus 90 degrees (*b*) and minus 90 degrees (*c*).



WEDGE DISPIRATION TWIST DISPIRATION TWIST DISPIRATION

DISPIRATIONS constitute a third class of lattice defects; they are based on screw symmetry, which combines translational and rotational components. Screw symmetry can exist in a periodic structure made up of units that are not spherical, such as the small cones here. In a perfect lattice all the cones in a layer point in the same direction, but the direction rotates in passing from one layer to the next. If either a dislocation or a disclination (of minimum displacement) were introduced into the structure, a discontinuity would appear where the orientation of the cones changed abruptly. Dispirations are generated by simultaneous translation and rotation, and they leave the lattice in register at all points. The discontinuities that would be created by either defect separately cancel when the defects are combined.

the atoms in a single plane pointing in the same direction but with successive planes pointing north, east, south, west and so forth. An attempt to introduce a 90-degree disclination into such a body inevitably creates a joint where the pattern changes abruptly. Dislocations with displacements equal to the separation between successive planes also lead to such forbidden transitions. By combining these rotational and translational motions, however, the crystal can be distorted while retaining a continuous pattern of orientations. Both wedge dispirations and twist dispirations are possible. As in the case of disclinations, however, a minimum rotation of 60 degrees is required, and dispirations are not observed in ordinary crystals.

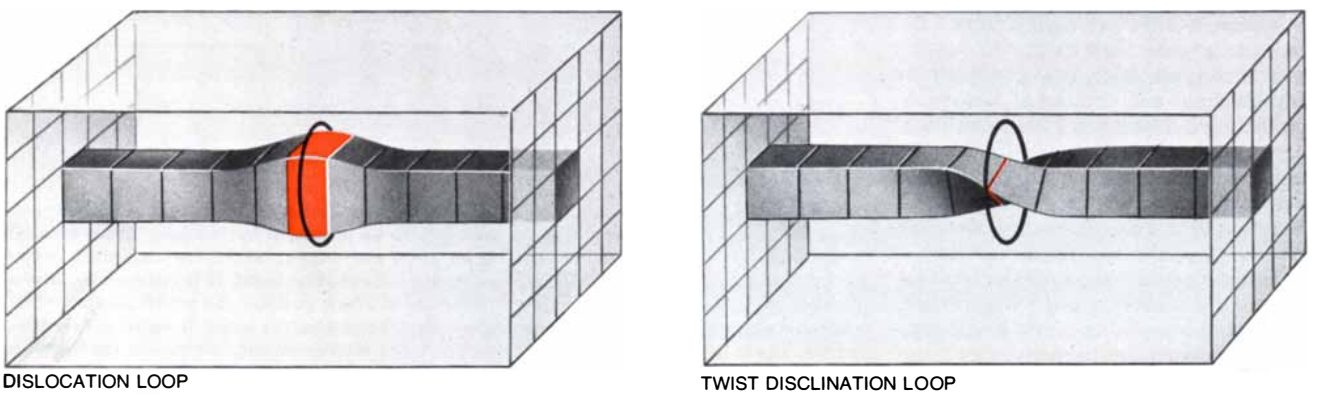
A dispiration could be decomposed, of course, into translational and rotational components. The defect is not a combination of a dislocation and a disclination, however, because neither the translation nor the rotation is by itself a symmetry operation of the lattice. The dispiration is a product of screw symmetry. Indeed, it was the screw symmetry of certain biological structures, such as muscle fibers, that led me to the discovery that a lattice defect could be based on screw symmetry.

All the defects I have discussed so far can be made with cuts that terminate on straight lines, but that is not a necessary condition. In slicing a sphere in order to create a wedge disclination, for example, one could employ a serrated knife. Indeed, the resulting defect line could have any shape as long as it is a continuous line. The only restriction is that the line itself must not terminate within the body except at an intersection with another defect.

The requirement that the defect line must not end inside the body does not exclude the possibility that the line might form a closed loop entirely within the body. It has been found that both dislocation loops and disclination loops of this kind can be created.

Dislocation loops are common in crystals. They can be regarded as having been formed by the insertion of a coin-shaped slug of material into the interior of the crystal, or by the removal of a similar disk. Thus the loop encloses material that is under compression or tension. In either case the displacement of the surface outlined by the loop, and the Burgers vector that defines the displacement, are perpendicular to the plane of the loop. Such a dislocation loop can move through the crystal; the region of compression or tension merely migrates through the structure the way an earthworm moves through its hole.

A similar disclination loop can easily



DISLOCATION LOOP TWIST DISCLINATION LOOP

DEFECT LOOPS are formed when a dislocation line or a disclination line describes a closed curve. The loop is created by making a small circular cut entirely within the body. If the cut faces are then displaced by a translation, the circular boundary becomes a dislocation loop; if the faces are rotated, it becomes a disclination loop. Here the dislocation loop is made by pulling the faces of the cut apart and somehow inserting an additional lattice unit; the loop then surrounds material under compression. If a unit were removed, the dislocation loop would enclose a region under tension. In making the disclination loop shown no material is added or removed; the cut faces are merely twisted 90 degrees with respect to each other. Both dislocation loops and disclination loops can travel through a crystal lattice.

IT UNTYPES AS EASILY AS IT TYPES.

THE TYPER

Typer cartridges come in black, red, blue, green, and brown executive quality film for the sharpest image of any portable; as well as everyday black nylon.

THE UNTYPER

When you make a mistake, just snap out the typer cartridge, and snap in the untyper cartridge.

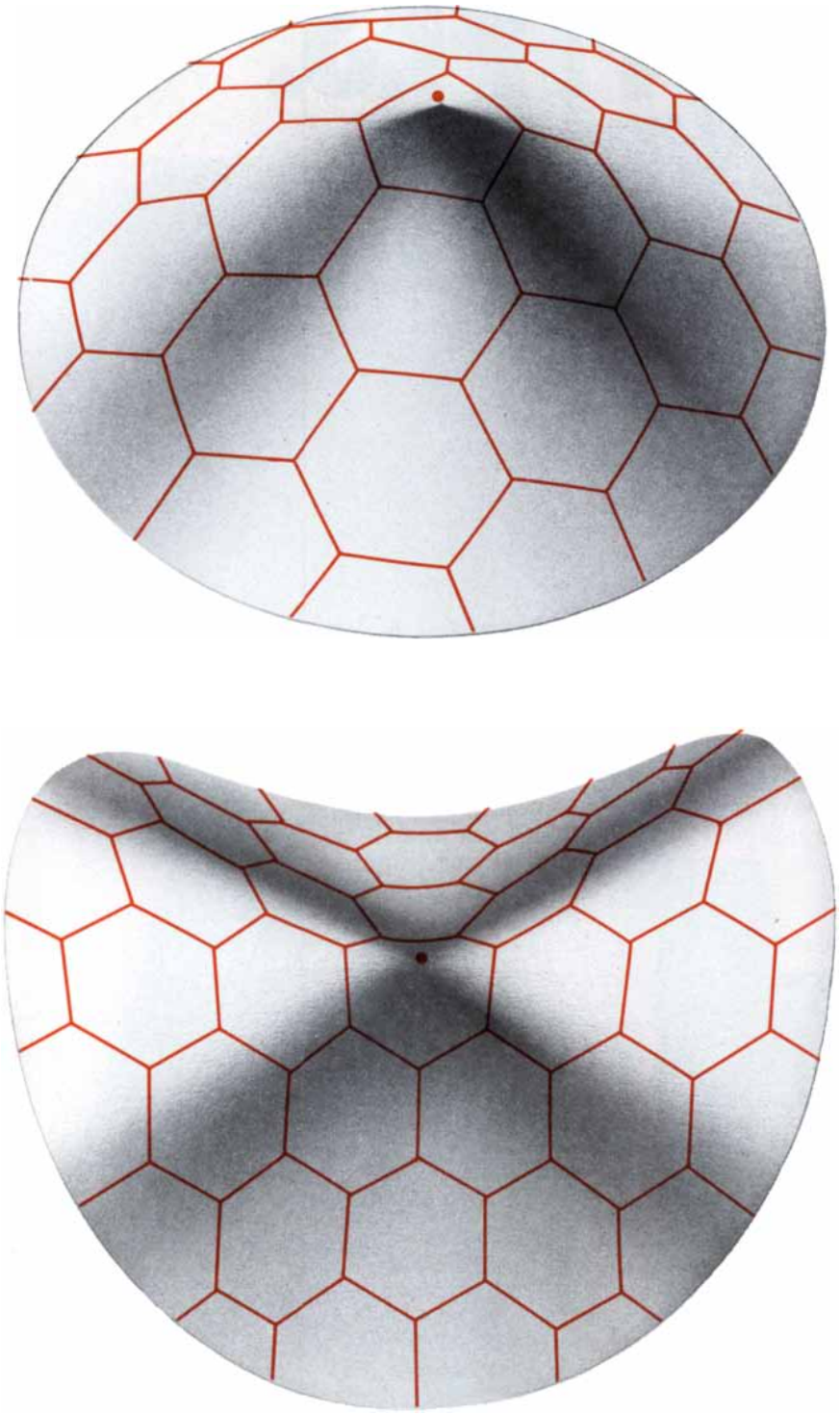
In seconds, good-bye mistake.

SCM SMITH-CORONA
SCM CORPORATION

be imagined, consisting of a twisted cylinder or fiber of material within an otherwise undeformed lattice. Just as the dislocation loop encloses material under compression or tension, so the fiber inside the twist disclination loop is under torsional strain. The axis of rotation in the disclination loop is perpendicular to the plane of the loop and passes through its center. Because such fibers could have rotational symmetry of an order higher than sixfold, disclination loops might exist with rotations of less than 60 degrees. The disclination loop could travel along the fiber it surrounds, accompanied by a local rotation of the fiber. Such wavelike propagation of twist would be far more likely than rigid-body rotation of the fiber as a whole.

The stresses and strains associated with disclination loops have been studied theoretically by several investigators, including Tsu-Wei Chou of the University of Delaware. It seems the loops could exist in crystals, but so far no examples have been found.

A structure that is thought to exist, however, is the twist dispiration loop; it is a loop defect surrounding a short region of a fiber that is both twisted and either compressed or stretched. These loops too are mobile. As they pass a region of the crystal, the fiber they enclose both rotates and is displaced linearly along the axis of symmetry. The net motion is like that of a screw, but it is important to note that the fiber does not turn as a rigid body does; the motion results instead from the passage of localized waves of twist and compression or tension. This mechanism is thought to be important in the deformation of crystalline polyethylene, which is made up of long-chain molecules with twofold screw symmetry. Twist dispiration loops of rotation 180 degrees travel along these molecules. The loops are called Reneker defects, after Darrell H. Reneker of the National Bureau of Standards, who described them in 1962.

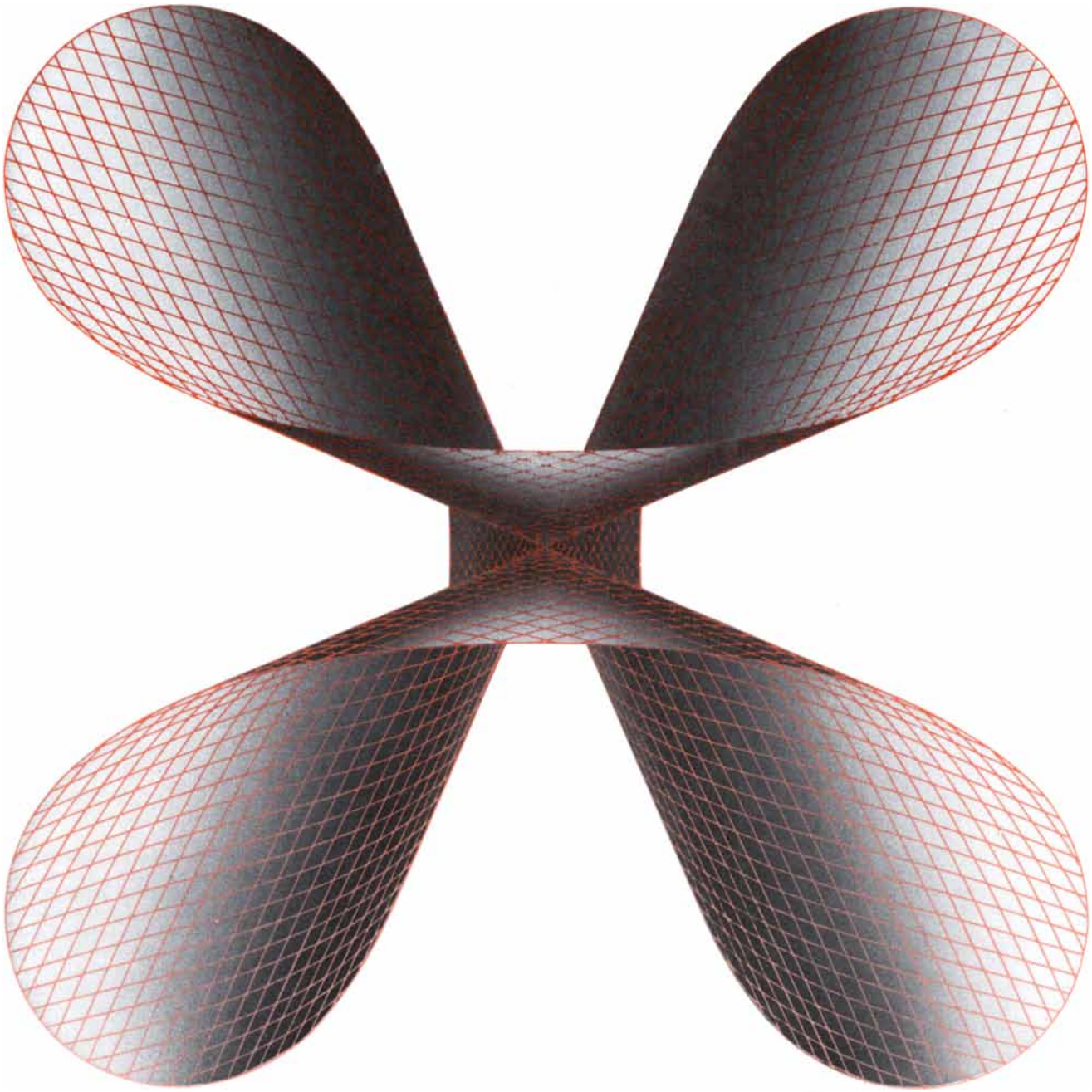


THIN SHEETS can absorb the strain of a disclination by buckling. At the top a positive wedge disclination of 60 degrees has been introduced into a thin disk, that is, a 60-degree sector has been removed. As a result the disk buckles to form a circular cone. At the bottom a negative 60-degree disclination, formed by inserting a sector, creates a saddlelike object whose surface has a more complex curvature, with four radial lines where the direction of curvature changes.

Edge and screw dislocations and the twist dispiration loop are the only symmetry-related defects that are believed to exist in materials spatially periodic in three dimensions. Twist disclination loops seem to be possible, but they have not been found. Wedge disclinations and dispirations, however, give rise to stresses that are prohibitively large. It seems unlikely they will be observed in ordinary crystals.

One kind of structure that can withstand the stress associated with a wedge disclination is a floating raft of bubbles. Such rafts, made up ideally of small, uniform bubbles, have proved to be a particularly convenient medium for the study of wedge disclinations. In a perfect array of bubbles the raft exhibits the sixfold symmetry characteristic of any array of close-packed spheres. Yoichi Ishida of the University of Tokyo has shown that wedge disclinations of minus 60 degrees and plus 60 degrees can be introduced into these rafts without disrupting the cohesive forces that hold the bubbles together. When a 60-degree wedge is inserted, the symmetry of the array becomes sevenfold; when a similar wedge is withdrawn, the bubble raft acquires fivefold symmetry.

Besides bubble rafts there are two classes of materials in which wedge disclinations can be formed. In one class,



DISCLINATION OF 360 DEGREES in a thin sheet gives rise to a severely deformed surface. As in the illustration at the bottom of the opposite page, the object is constructed by cutting a slit along some radius of a disk, rotating the cut edges around the center of the disk so that a gap is opened and then filling the gap with a wedge of material. In this case, however, the "wedge" is itself a full circle. The perimeter of the surface traces a curve somewhat like that described by the seam of a baseball. The grid lines are continuous at all points on the surface and the seam where the edges were joined cannot be identified. Indeed, any pattern could be made continuous on the surface.

exemplified by liquid crystals, much of the stress can be relieved by flow. The other class is made up of bodies that are thin enough to buckle when they are stressed.

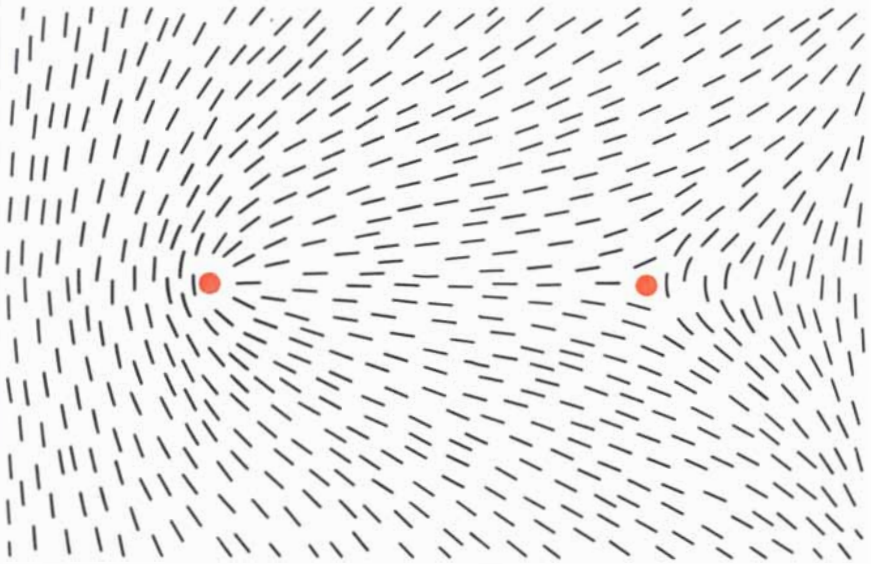
The curious properties of a material that is crystalline but still a liquid favor the creation of disclinations. Among the liquids of greatest interest are those called nematic liquid crystals. They are made up of rod-shaped molecules that tend to line up parallel to one another. Thus the molecules have a defined orientation, but they do not occupy fixed positions in a rigid lattice. In general the orientation of the molecules changes smoothly and slowly with change in position. The exceptions are at line singularities, or in other words disclinations; there the orientation changes abruptly. Under a microscope the disclination lines appear as threads. The liquid crystals were named for these features; "nematic" is from the Greek for thread.

Viewed along an axis perpendicular to the rods, the molecules of a nematic liquid crystal have twofold rotational symmetry, and so the smallest disclinations possible are those of 180 degrees. Removing a 180-degree sector of material changes the liquid crystal to a configuration that has only onefold symmetry; adding a 180-degree wedge results in threefold symmetry. A vast body of knowledge about disclinations in liquid crystals has been acquired during the

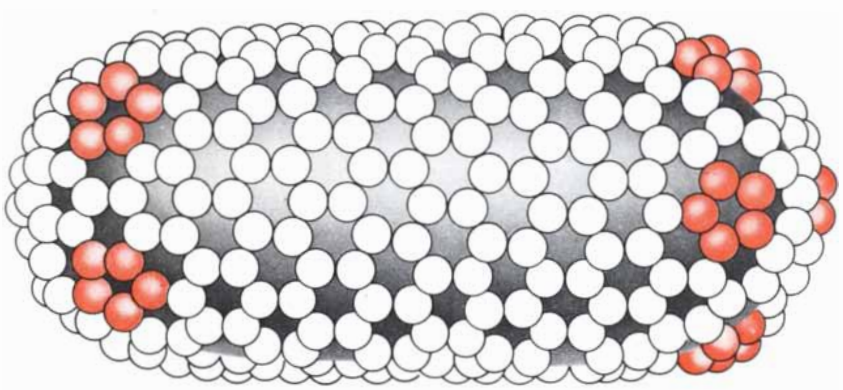
past 10 years, largely as a result of the work of Yves Bouligand of the École Normale Supérieure and that of P. G. de Gennes, M. Kleman, Jacques Rault and others of the University of Paris-Sud.

A related group of materials in which disclinations have only recently been recognized consists of cokes, tars, pitches and graphites. Layers of fused benzene rings are important structural elements in these substances, and James E. Zimmer and Jack L. White of the Aerospace Corporation have shown that disclinations are present in the arrangement of the molecular layers. They have also found that the defects have a significant influence on the physical properties of the materials.

The same twofold symmetries observed in liquid crystals, and therefore the same 180-degree disclinations, are found in many two-dimensional patterns. They appear, for example, in the stripes of zebras. An even more familiar example is the pattern of ridges on the fingers of man and the other primates. In addition to rotations of plus 180 degrees and minus 180 degrees, wedge disclinations of plus 360 degrees are observed in fingerprints. The resulting patterns are



TWO-DIMENSIONAL PATTERNS frequently exhibit wedge disclinations. The underlying rotational symmetry of this pattern, if it had no imperfections, would be twofold, and so the only disclinations possible are those with rotations of 180 degrees or multiples of 180 degrees. The undistorted pattern can be imagined as an array of parallel bars. The dot at the left marks the position of a positive 180-degree wedge disclination: half of a circle has been removed and the lines bend smoothly to fill the void. At the right a 180-degree wedge has been inserted. In fingerprint patterns similar disclinations are called loops and triradii. Positive disclinations of 360 degrees, in which the lattice lines form closed loops, are also seen in fingerprints; they are called whorls. Similar patterns are found in liquid crystals and in the pelts of striped animals.



CLOSED SURFACES have disclinations as an essential structural element. It can be proved that any surface that can be deformed into a sphere (and that is therefore topologically equivalent to a sphere) must possess disclinations with a total rotation of 720 degrees. In this surface the basic symmetry element is hexagonal, and closure requires the insertion of 12 pentagons (*color*). Each pentagon represents a 60-degree disclination. The drawing is based on a model made by Ralph O. Erickson of the University of Pennsylvania of the protein coat of a virus.

known to the dermatologist respectively as loops, triradii and whorls.

Patterns formed by human hair and by the hair of some other mammals can also show wedge disclinations. The hair grows at an oblique angle to the skin, and it therefore has only onefold axes of rotational symmetry perpendicular to the skin. The only allowed disclinations in such a pattern are those with rotations of minus 360 degrees or plus 360 degrees. The corresponding point singularities are respectively called crosses and whorls.

The large strains that prevent disclinations from forming in most solids are much reduced in a thin sheet that is free to deform by buckling. In the absence of buckling the energy of the strain is proportional to the square of the rotation; when buckling is allowed, on the other hand, the energy is proportional to the rotation itself. This relation is actually valid only for vanishingly small rotations; for large rotations Stanley L. Thomas and I have shown that the strain energy of a buckled sheet is still smaller.

A positive wedge disclination (that is, one generated by the removal of a wedge) changes a thin circular sheet to a circular cone. As larger wedges are removed the cone becomes steeper, until when the rotation reaches 360 degrees the apical angle of the cone falls to zero. Then the sector to be removed is the entire sheet, and the cone disappears.

The insertion of a wedge into a flexible circular sheet (equivalent to a negative rotation) yields a body with a quite different geometry. The surface vaguely resembles that of a saddle, at least for comparatively small rotations, but it is not smooth at the center. Formally it is still a cone, since the surface can be generated by moving a straight line fixed at one end. The fixed end corresponds to the disclination point; the other end traces a closed curve on the surface of a sphere, a curve similar to the one described by the seam of a baseball.

A positive wedge disclination of the kind I have described cannot exceed 360 degrees; after the entire circle has been removed there is nothing left. There is no limit, however, to the amount of material that can be added. It is a simple matter to insert a "wedge" of 360 degrees, yielding a wedge disclination with a rotation of minus 360 degrees. A model of such a disclination can be constructed from two circular sheets of paper; each is slit along a radius and the slit edges of the two sheets are smoothly joined. A structure with four symmetrical lobes results. Insertion of still larger wedges brings pairs of lobes closer together, until they touch at a rotation of about minus 403 degrees. What happens at still larger angles depends on whether the surface is allowed to intersect itself;

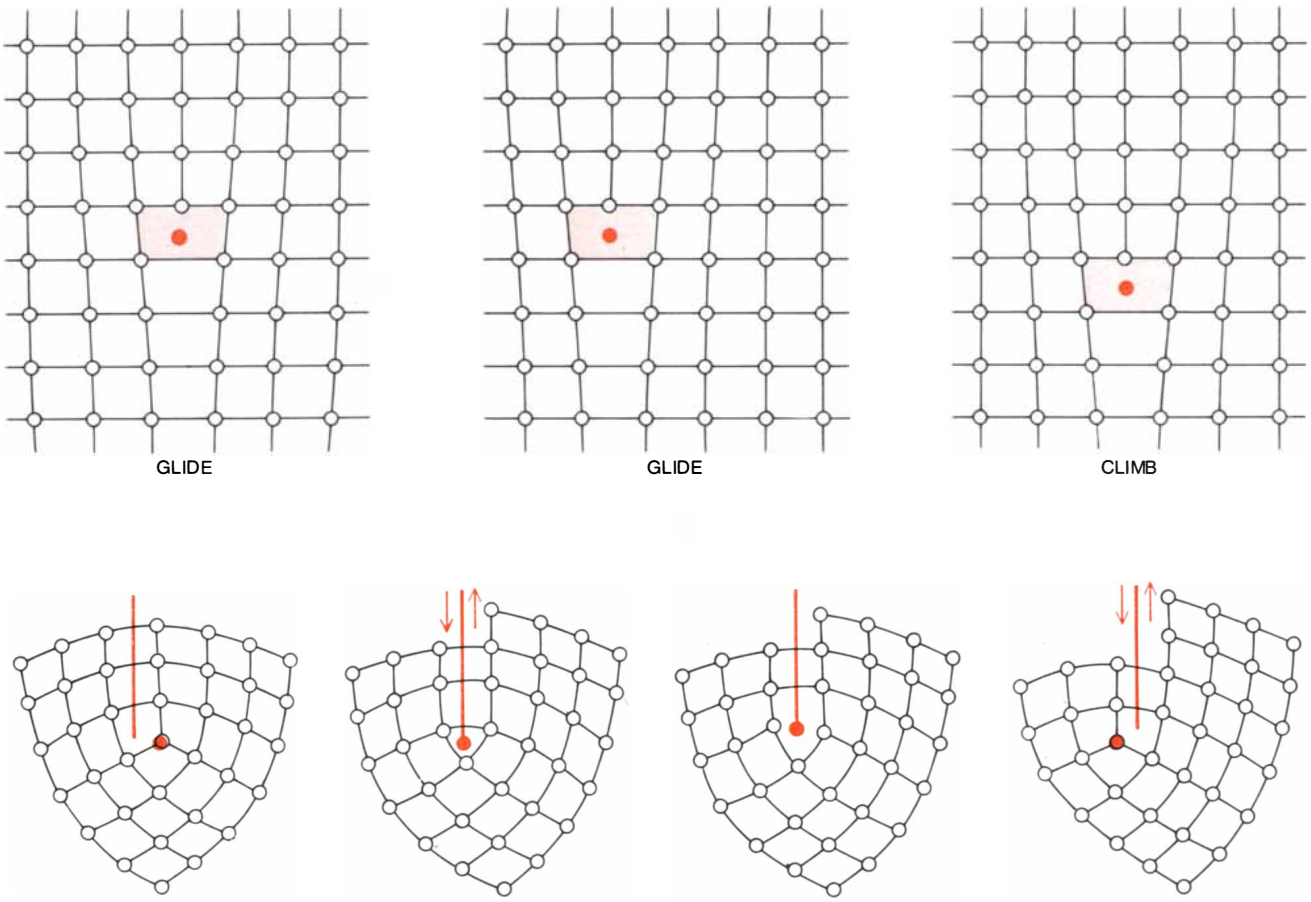
Thomas and I are now investigating that geometry.

The requirement that a positive wedge disclination cannot exceed 360 degrees applies only if the disclination line or point coincides with the axis of rotation. Positive rotations of 360 degrees or more are also possible if the wedge disclination is combined with some other defect. An example of a structure having such a compound defect is a cylindrical crystal: a two-dimensional crystal rolled up to form a cylinder of circular cross section. The bending of the lattice through a full circle around the axis of the cylinder reveals a disclination with a rotation of 360 degrees. There is also another defect in the structure, an edge dislocation, although its presence could be demonstrated only through a detailed analysis of the cylindrical lattice.

An extraordinary feature of the cylindrical crystal is that it has no singular points or lines. In ordinary crystals dislocations and disclinations can always be identified as line imperfections. In surface crystals the line imperfection is ordinarily reduced to a point, such as the apex of a cone. The cylindrical crystal, however, has no identifiable imperfections even though it possesses both a disclination and a dislocation. One can imagine bending a long cylinder and joining the ends to form a toroidal crystal. Again no singular points or lines could be identified, even though the lattice would now include two disclinations and two dislocations. The strains would vary smoothly over the entire surface. I am not aware of any toroidal surface crystals in nature, but there are many biological examples of cylindrical crystals. They include the protein coats of rod-shaped viruses, the flagella of bacteria and the microtubules that are structural elements in many living cells.

Instead of bending a cylindrical crystal to form a torus we could cap the ends to form a closed surface topologically equivalent to a sphere. From a fundamental theorem of geometry, proved by Leonhard Euler, it can be shown that any such sphere must include wedge disclinations with a total rotation of 720 degrees. The 720 degrees of disclination can be distributed in many ways. A particularly common distribution is in 12 disclinations of 60 degrees each. For example, it is not possible to tile a sphere with hexagons, but the tiling can be accomplished if 12 pentagons are included among the hexagons. If the pentagons are regarded as defective hexagons, then each represents a plus-60-degree wedge disclination. In the protein coats of spherical viruses, and in the coats of rod-shaped viruses that are capped at the ends, 12 disclinations of 60 degrees each can be identified.

The singularities at points and lines associated with lattice imperfections are not necessarily fixed in the lattice. We have already seen that movement is possible for the internally looped forms of dislocations, twist disclinations and dispirations. Among defects that terminate along a straight line, dislocations have long been known to be mobile. For the purpose of analyzing the motion we



MIGRATION of edge dislocations and wedge disclinations takes place through the step-by-step rearrangement of the atoms in a lattice. Dislocations have two possible modes of motion. In the glide mode a bond is transferred from the atom at the left of the dislocation to the one that terminates the inserted half plane of atoms; as a result the dislocation steps one lattice unit to the left. In the climb mode an atom is added to extend the half plane by one unit. The movement of disclinations is somewhat more complicated. Here a positive wedge disclination is shifted diagonally to the left across a distorted square, and in the process it generates two edge dislocations. The dislocations are depicted as moving by glide along the colored line. The first dislocation shears the lattice, shifting the disclination point to a triangular space between atoms. The second dislocation shears the lattice again. The disclination point then completes its migration.

The Sony 7800 Receiver. It'll put you on the receiving end of a lot more than compliments.

Even before you switch on the STR-7800SD receiver, it'll be receiving. Receiving oohs and aahs.

After it goes on, the accolades will really come in.

After all, it *is* the finest receiver ever designed by Sony. The 7800 puts you on the receiving end of the most tomorrow-looking technology available today.

You'll receive a feeling of power.

The 7800 brings power to the power-hungry, and can even make the mild-mannered lust for power. Rated at 125 watts per channel, it's powerful enough to drive any speakers—satisfy any need.

The 125 watts, minimum RMS at 8 ohms, is from 20 to 20kHz—with no more than 0.07% Total Harmonic Distortion.

And that's Sony's conservative rating.

How this combination of power and low distortion was achieved, is an example of Sony's engineering muscle.

Let's start with the toroidal coil transformer. A more efficient structure, it fully exploits the high-performance power amp. As do two over-sized capacitors, each 22,000uF.

So the feeling of power throughout the frequency range is unmistakable.

You'll receive tuning that'll leave you swooning.

FM circuitry found usually in separates appears in the 7800.

Pardon our initials, but MOS FET's are used in the FM RF amplification. The result: good linearity, low noise and high sensitivity.

For you FM Dolby listeners, a complete FM Dolby noise reduction system, to minimize noise *and* overload distortion.

And there's a new local oscillator circuit. Plus our uni-phase IF filters are so advanced, a computer designed them.

A Multipath switch and meter indicate optimum antenna orientation,

thereby reducing distortion. An LED dial pointer doubles in length when an FM signal is received for easy tuning.

You'll receive a pre-amp that's pre-eminent.

High marks for our low emitter concentration (LEC) transistor. Designed, made by, and exclusive to Sony. It guarantees low noise, and a wide dynamic range. It also keeps RIAA equalization to within ±0.5dB

You'll receive power. And the means to control it.

The pre-amp section also gives you the control you need to keep all that power in line.

A presence switch is a special present: it lets you equalize the mid-range.

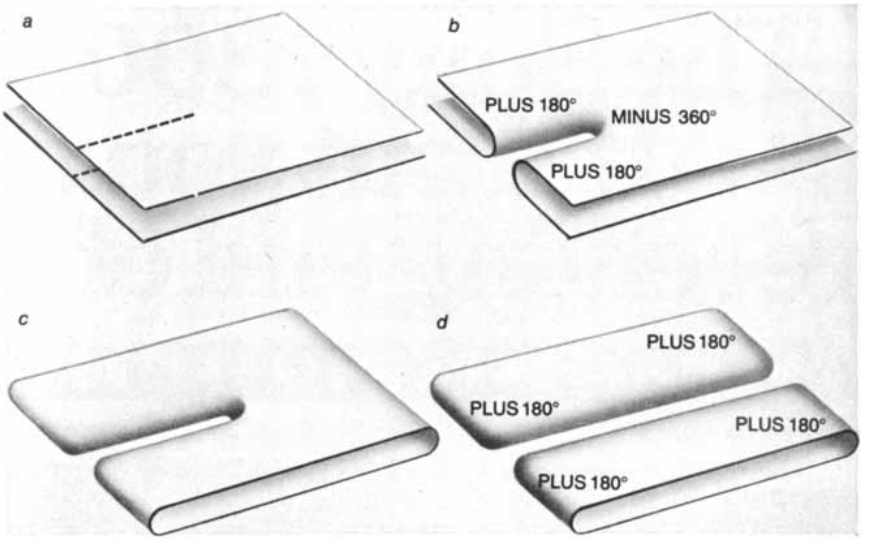
Importantly, the 7800 was built with a Professional Attenuator Main Volume Control. It eliminates gang error between channels.

Nor have we overlooked a special loudness network, or an audio muting switch.

Some input on the inputs: Phono 1, Phono 2, External Adaptor, Auxilliary, two tape decks—and tape-to-tape dubbing facilities.

Certain pieces of machinery simply ooze quality and power. Such is true of the 7800SD. It'll put you on the receiving end of the living end.

SONY®



FISSIONING OF LIVING CELLS can be represented by a model in which the essential process is the movement of a minus-360-degree disclination. The "wedge" of 360 degrees is inserted by slitting two sheets (*a*) and joining the cut edges (*b*). The disclination point is at the end of the resulting fissure, and it can be moved by extending the cuts and simultaneously resealing the edges. If the two sheets are connected to form a flattened cylinder, the disclination can be started at a fold (*c*) without cutting open the surface. The "corners" created in this way, however, are special points that can be interpreted as flattened cones and hence as positive wedge disclinations; in the simplest case each has a rotation of plus 180 degrees. By moving all the way across the surface the disclination can divide the body in two (*d*). At the moment of separation the minus-360-degree disclination disappears and two plus-180-degree disclinations are created; the total rotation of the disclinations increases by 720 degrees. If the ends of the original cylinder are closed, the result is two closed surfaces. By this mechanism a closed body can be cut in half without breaking the surface. The flagellate *Euglena* divides this way, although the 360-degree disclination forms and disappears not at folds but at other points on the surface.

can regard the dislocation as a point defect where an extra half row of atoms terminates in the interior of a square lattice. Two modes of motion are possible: glide and climb. Glide refers to motion of the defect parallel to the Burgers vector by the rearrangement of the atoms surrounding it. The part of the crystal with the extra half row of atoms steps like a caterpillar across the rest of the crystal. Climb is a motion perpendicular to the Burgers vector and requires the addition of atoms to the lattice. Each atom is inserted at the end of the half row, perfecting the lattice at the former site of the dislocation but creating a new defect in the next row of atoms. Climb can also be reversed, of course, by removing atoms from the end of the half row. The movement of dislocations is the primary mechanism for the plastic deformation of crystalline materials.

An interesting question is whether or not wedge disclinations can move. It was raised at a conference on the theory of dislocations in 1969 and was answered soon after by me and independently by Roland de Wit of the Bureau of Standards. It turns out that they can move, but only through a complicated sequence of events in which dislocations are involved. A 90-degree wedge disclination, for example, can move diagonally across a distorted square in the lattice; in the process it generates two edge dislocations. The dislocations move away by glide or climb. The same process can operate in reverse: the disclination can move by absorbing dislocations that reach it from elsewhere in the crystal.

Three years ago I was intrigued to discover a simple but unexpected mode of motion for disclinations with a rotation of minus 360 degrees. For this special case I found that the Burgers vectors of the dislocations created by the movement of the defect seem to vanish. An obvious interpretation of that finding is that minus-360-degree disclinations can move without generating dislocations.

The mechanism is best visualized in a minus-360-degree disclination made from two sheets, one lying on the other. We slit both sheets and reconnect the cut edges of one sheet to those of the other. The disclination point then lies at the end of the cut. The disclination can move across the material by a tearing and simultaneous rejoining of opposite edges. The same thing could happen with just one sheet, rolled into a cylinder and then flattened. The minus-360-degree disclination could travel across the cylinder from one folded edge to the other by the same process of simultaneous tearing and rejoining. Where the moving disclination first nicks the folded edge it leaves behind two special points: the "corners" where the slit begins. Each corner can be regarded as a flattened cone or, in other words, a wedge disclination. In the simplest case

each has a rotation of plus 180 degrees. A similar pair of disclinations is created where the minus-360-degree disclination vanishes at the end of its journey. At the instant the minus-360-degree disclination disappears the original cylinder is split in half.

If the flattened cylinder is given closed ends, Euler's theorem requires that the surface have disclinations whose rotations add to 720 degrees. This property is not affected by the introduction of the minus-360-degree disclination, since two plus-180-degree disclinations are necessarily created at the same time. When the minus-360-degree disclination reaches the far edge, the structure divides into two closed surfaces, each of which has the topology of a sphere. As the disclination reaches the edge it disappears and is replaced by two more plus-180-degree disclinations. As is required by Euler's theorem, each of the new bodies has disclinations with a total rotation of 720 degrees.

This mechanism is an appealing model of a fundamental process in biology: cell division. It provides an organized means for a cell to split into daughter cells without spilling any of the contents. It can be generalized by allowing the minus-360-degree disclination to form and to disappear at any defect on the cell surface, rather than only at folds in a flattened envelope.

In 1975 I published an account of this idea, but in the literature of physics rather than that of biology. (I suspected that biologists would find it naïve.) A year later, however, Cecilia Hofmann and G. Benjamin Bouck of the University of Illinois described the same process in the division of the flagellate *Euglena*. In *Euglena* a wedge disclination of rotation minus 360 degrees forms at a defect in the surface near one end of the cell. It travels by a localized tearing and rejoining and ultimately disappears at another defect at the opposite end. The presence of disclinations with rotations of plus 360 and minus 360 degrees is revealed by stripes in the cell surface.

It should be emphasized that most living cells probably do not divide in this way. Their surface membrane is fluid and does not exhibit any recognizable disclinations. Cells with a structured surface that is maintained throughout division, however, probably do employ this mechanism. Indeed, if no rearrangement of the cell surface is allowed other than local tearing and fusing, then it is the only way a cell can divide.

Until quite recently disclinations, particularly those of large rotation, have been regarded as little more than academic curiosities. As I have tried to show, however, imperfections such as disclinations can be expected to appear wherever there is order. The imperfections do not destroy that order; on the contrary, they enhance it.

Weaver Ants

These social insects use their own larvae as shuttles to weave leaves into large nests in the rain forests of Africa and Asia. Their behavior is coordinated by complex chemical stimuli

by Berthold K. Hölldobler and Edward O. Wilson

Some insect species live in advanced social orders characterized by cooperation, caste specialization and individual altruism. Among the thousands of species of social insects a few deserve to be called classic, because certain remarkable features in their behavior have prompted unusually careful and thorough studies. The honeybees, the bumblebees, the driver ants, the army ants, the leafcutter ants, the slavemaker ants and the fungus-growing termites are all examples of classic social insects. The latest candidates for this select group are the weaver ants of the genus *Oecophylla* of Africa and tropical Asia. These ants devote a large part of their behavioral repertory to communication. The communication is further enriched by variations based on the weaver ants' caste system. As a result a weaver ant colony can perform feats far beyond the capabilities of single ants.

Weaver ants are extremely abundant, aggressive and territorial. They have achieved a position of exceptional ecological importance in the rain forests, cacao plantations and similar wooded environments they inhabit. For this reason the weaver ants have been the object of an increasing number of field studies. Over the past two years, however, we have succeeded in cultivating colonies of the African species *Oecophylla longinoda* in the laboratory. We induced the ants to live in potted trees and glass tubes. Under these conditions it has been possible for the first time to study the full range of the social life of the weaver ant.

These slender yellow insects are exquisitely adapted for life in the leafy canopies of tropical forests. Their main social unit is the colony, which consists of as many as 500,000 female workers, the progeny of a single enormous queen. The caste system within each colony consists of three forms of adult female: the heavy-bodied queen, a large population of "major" workers and a smaller population of "minor" workers. The weaver ant males, like those of other ant species, participate relatively little in the social life of the colony. They leave soon after maturing to participate in wedding flights with the virgin queens, after which they die without returning to the nest.

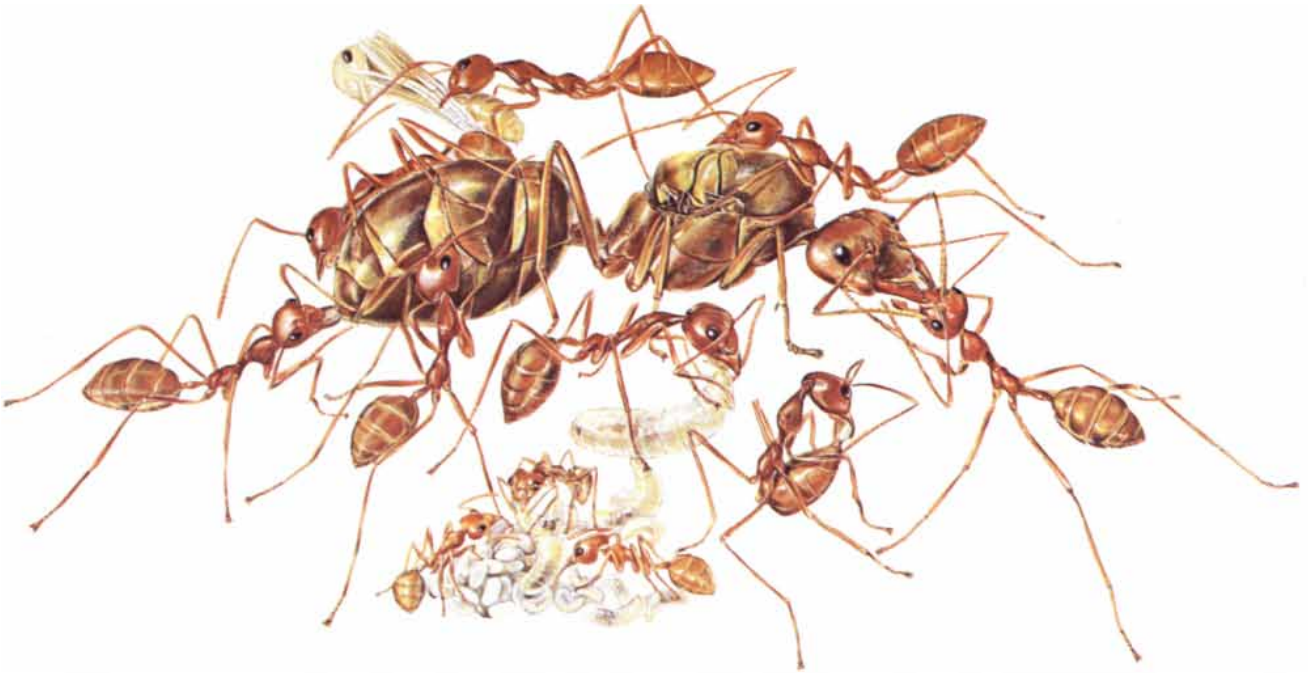
The major workers are fairly large, averaging six millimeters (about a quarter of an inch) in length. They are the general laborers, responsible for most of the foraging and nest construction. The more aggressive of the two worker castes, they rush from the nest at the slightest disturbance to bite an intruder and release formic acid from their poison gland. Major workers also form a dense retinue around the queen. They grasp her with their powerful legs so tightly that at times she is held in midair in the center of the nest cavity. About once a minute one of the major workers regurgitates a liquid meal into the mouth of the queen. At somewhat less frequent intervals a member of the queen's retinue lays a special trophic egg—a flaccid object without the ability to survive—that is immediately fed to the queen. This virtually continuous flow of nutrients enables the queen to manufacture hundreds of eggs a day. As the eggs are extruded from the queen's oviduct major workers carry them to special brood piles. There the smaller minor workers care for the eggs and feed and wash the tiny larvae that hatch from them. When the larvae are close to their maximum size, the major workers and minor workers share about equally in their care.

Weaver ants are named for their method of nest construction. The nests are made of leaves folded or fastened together to form tight, tentlike compartments. The leaves are held in place by seams of silk spun by the larvae, which the major workers employ like shuttles for weaving the nests. This nest building is one of the most remarkable instances of social cooperation among lower animals.

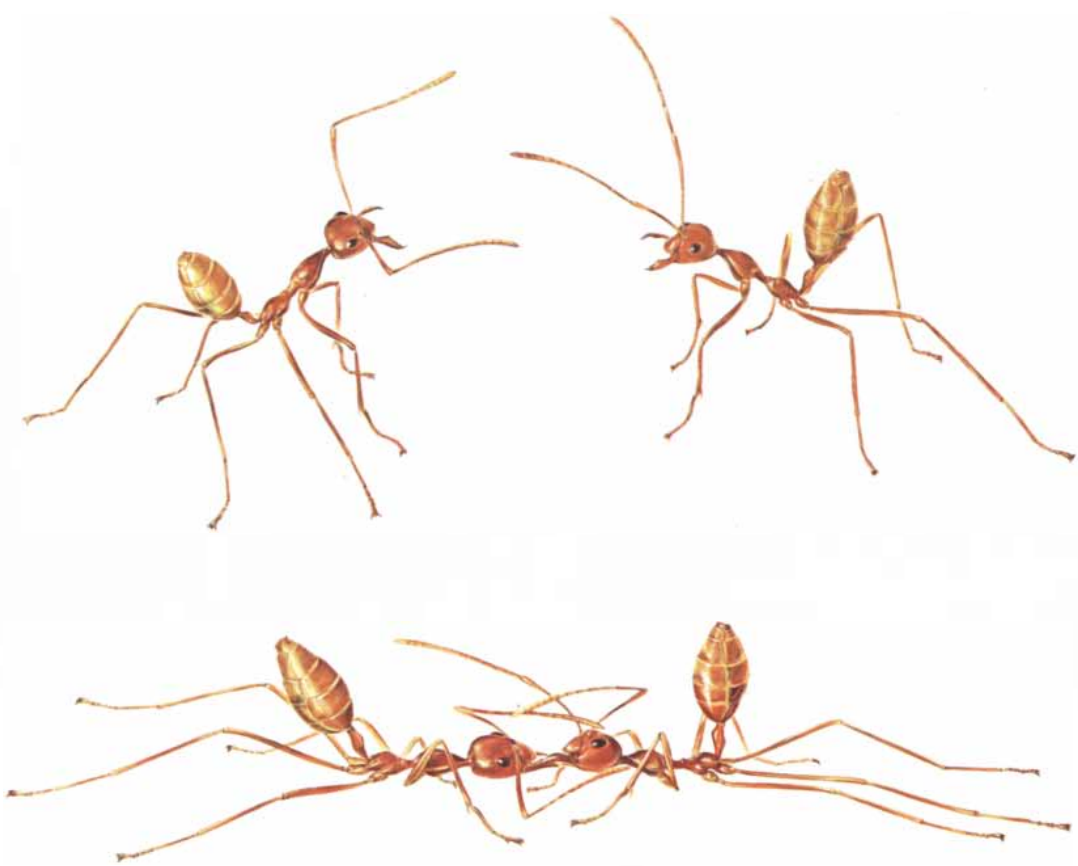
Once the weaver ants have chosen a tree branch suitable for a nest they spread out on the leaves of the branch and begin to pull on the tips and edges. When an ant succeeds in turning up a segment of a leaf, nearby workers are attracted to that part of the leaf, and soon there is a small group of ants pulling in unison. When a leaf is broader than the length of an ant's body, or when two leaves must be pulled together across a wide space, the workers form living bridges between the points to be joined. Then some of the ants in the chain climb onto the backs of their neighbors and pull backward, thus shortening the chain and bringing the leaf edges together. When the leaves have been maneuvered into shape, some of the ants remain on them, employing their legs and mandibles to hold the leaves in place. Other ants go back to already established nests and return to the new site carrying partly grown larvae. The workers wave the larvae back and forth across the leaf seams. This causes the larvae to release threads of silk from gland openings located just below their mouth. Thousands of these threads woven into sheets are strong enough to hold the leaves in place. Sheets of silk are also spun to make circular entrances and outer galleries leading to the interior of the new nest.

A single weaver ant colony can occupy a substantial volume in the canopy of a forest. The colony can fill an entire tree or even several adjacent trees without breaking the lines of communication that are vital to social insects. From the leaves of the trees the weaver ants construct hundreds of nests to serve as retreats, nurseries and outposts. During the day foragers patrol every square centimeter of leaf and bark within their territory. They rout enemies, capture insect prey and gather the sweet "honeydew" excrement of the swarms of scale insects and other sap-feeding homopterous insects that the ants guard as though they were dairy cattle.

The species of *Oecophylla* are not the only insects that weave. A few other



CASTE SYSTEM of the African weaver ant consists of three forms of adult female: a single queen, a population of "major" (large) workers that forage for food and perform a variety of other tasks, and a lesser number of "minor" (small) workers that care for the eggs and younger larvae. The queen (*center*) is occupied solely with receiving food from the major workers, chiefly by regurgitation, and laying eggs. In the foreground one major worker regurgitates to a larva; a second lays a "trophic" egg that lacks survival ability and will be fed to the queen. Minor workers are clustered around a brood pile. Scene was painted by Turid Hölldobler. Ants are enlarged about six times.



TERRITORIAL COMBAT of two weaver ants is initiated by a dance-like maneuver in which the combatants raise themselves by extending their legs and circle each other with stiff, jerky movements (*top*). Then one ant attempts to seize the other with its mandibles (*bottom*).

tropical species use larval silk to construct nests in trees and shrubs. The weaver ants are distinguished, however, by their close control of their environment. Indeed, man has employed weaver ants to control the arboreal environment for him.

Records from the Canton area of China show that weaver ant nests were gathered, sold and placed in selected citrus trees to combat insect pests in about A.D. 300. The same technique was noted in the 12th century and was still practiced in southern China well into the 20th century. The weaver ant used for this purpose is the Asian species *Oecophylla smaragdina*. This utilization of weaver ants is the oldest-known instance of the biological control of insects in the history of agriculture. Recently Dennis Leston, formerly of the University of Ghana, and other entomologists have recommended employing the African species of weaver ant to control pests of tree crops such as cacao. Studies in Ghana have shown that the presence of weaver ants reduces the incidence of two of the most serious diseases of cacao, one caused by a virus and the other by a fungus. In both cases the pathogen is transmitted by mirid leaf bugs. The weaver ants evidently combat the diseases by attacking the bugs. The *Oecophylla* workers are also particularly effective in hunting insects that feed on the tissue and sap of trees.

The weaver ants' exceptional control of their environment has been achieved through the evolution of advanced forms of social behavior. The communication system we have observed in our studies of the African species of weaver ant is one of the most complex and advanced systems known among the social insects. The great strength of the weaver ants lies in their ability, demonstrated in their nest building, to cooperate in group activities. They employ five different recruitment systems, consisting of distinct combinations of chemical and tactile signals, to initiate other group endeavors. These recruitment systems are employed in the main occupations of the weaver ants outside the nest: penetrating new territory, defending it and extracting food from it.

Weaver ants have an impressive sense of place that helps them to secure new territory. Their large eyes give them vision that is unusually acute for ants. Moreover, they are able to remember the appearance of many details of the nest area. If an object is simply shifted from one side of a nest tree to the other, workers come out of the nest to explore the object as if it were fresh terrain. In fact, when a conspicuous object such as a potted plant or a box is moved close to a weaver ant nest, the alert workers crawl out over the branches and leaves of their own tree in an effort to reach it. If the ants fail to get onto the new surface by reaching with their legs, they begin to climb on top of one another, constructing pyramids or chains with their bodies until the chasm is bridged. Then workers rush onto the new territory and begin to explore it.

The first explorers return to the nest to recruit other workers to help in securing the new territory. They mark the route from the territory to the nest with odor trails, that is, trails of a pheromone, or message-bearing chemical, that will guide their nestmates to the new area. The odor trails are created in an unusual way. The ants extrude a glandular segment of the hindgut through the anus. This organ, which we call the rectal gland and which is known only in weaver ants, is used in four of the recruitment systems. When the gland is extruded, it rests on a tiny sled consisting of two bristles that project from the tip of the ant's abdomen. As an ant runs back to the nest a secretion from the rectal gland is brushed on the ground, thereby creating the odor trail. When an ant laying one of these trails encounters nestmates, it jerks its body in their direction while touching them on the head with its antennae. The greeting stimulates the nestmates, with the result that they follow the trail to the new territory and begin to explore it. We call this process the system of recruitment to new terrain.

LIVING CHAIN of weaver ants folds over the tip of a leaf during the construction of the arboreal nest, a remarkable example of cooperative behavior. The workers shorten the chains by climbing up on the backs of nestmates while hauling edge of the leaf behind them.

THE NEW CUTLASS SALON. A CARFULL OF NEW IDEAS.

Space efficient trunk.

All new bi-level instrument panel.

Plenty of headroom and legroom.

Rugged new suspension.

New functional design

Easy to park.

They add up to a shining example of what a restyled car can do:

Lots of Headroom and Legroom—Good news for six footers. You can stretch out your legs, sit up straight. Salon has your kind of room inside.

New Instrument Panel—Salon's dash is neatly arranged so everything falls within your frame of vision.

Impressive Feel-For-The-Road—Salon's trim body and sophisticated suspension add up to a responsive, exciting road machine. Try it on your favorite road and see what we mean.

Good Gas Mileage—EPA estimates are 29 mpg highway, 20 mpg city, 23 mpg combined, with available 260 V8 and standard manual transmission. Your mileage will depend on how and where you drive, your car's condition and equipment. Power train not available in California. Cutlass Salon is equipped with GM-built engines produced by various divisions. See your dealer for details.

Easy Parking—When you test drive the Salon, try parking it. Spaces you used to pass up now seem larger, thanks to its new, tight turn diameter.

One more idea—Go down to your Olds dealer and take the Salon for a test drive, and discover that great Cutlass feeling for yourself. Spending a few minutes with so many great new ideas just might spark a few of your own.

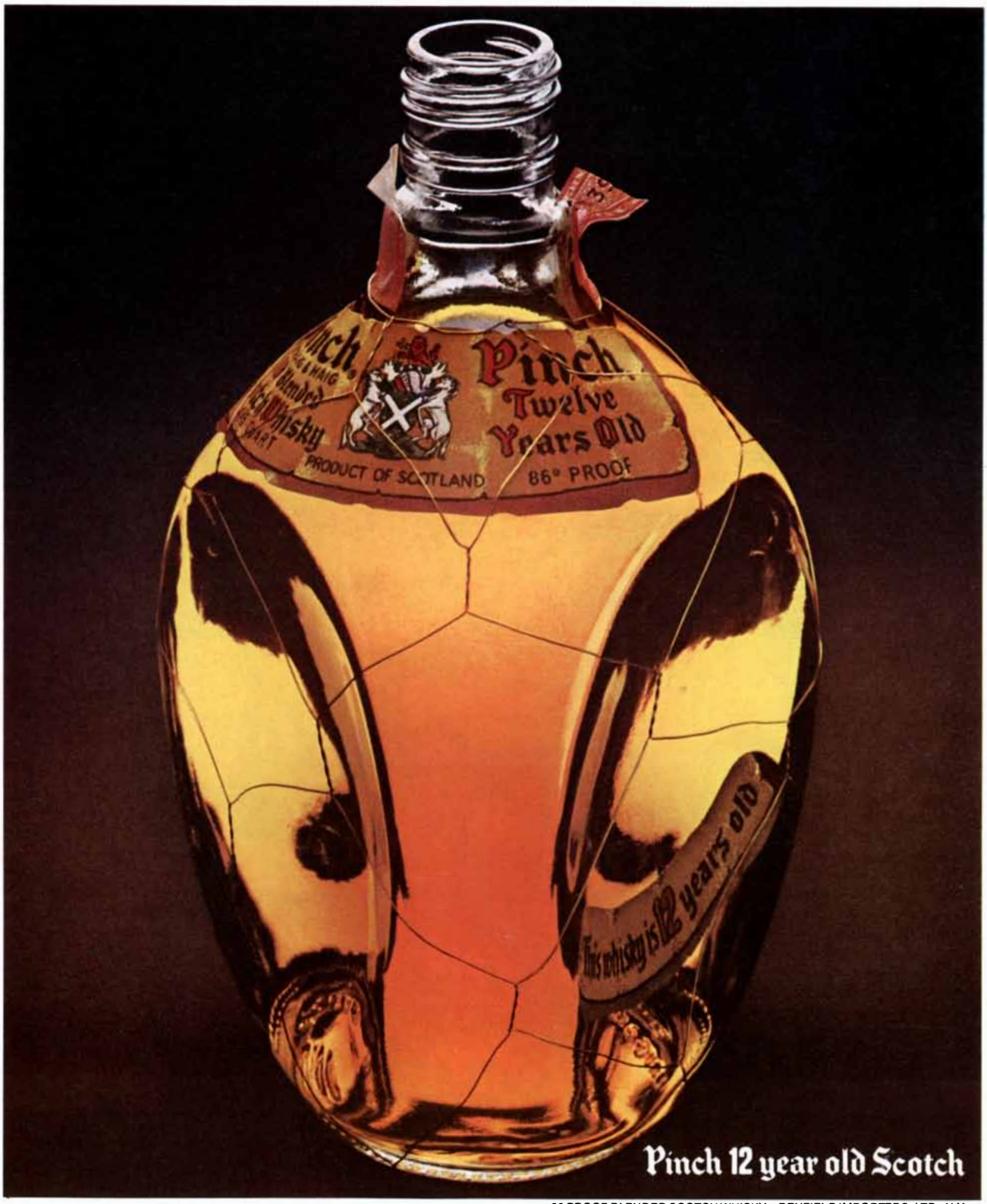
CUTLASS SALON

Oldsmobile

Can we build one for you?

GM

Look at it this way:
Your mother's in from Cleveland for the holidays.
So for two weeks you get to be "the baby" again.
This is no time for ordinary scotch.



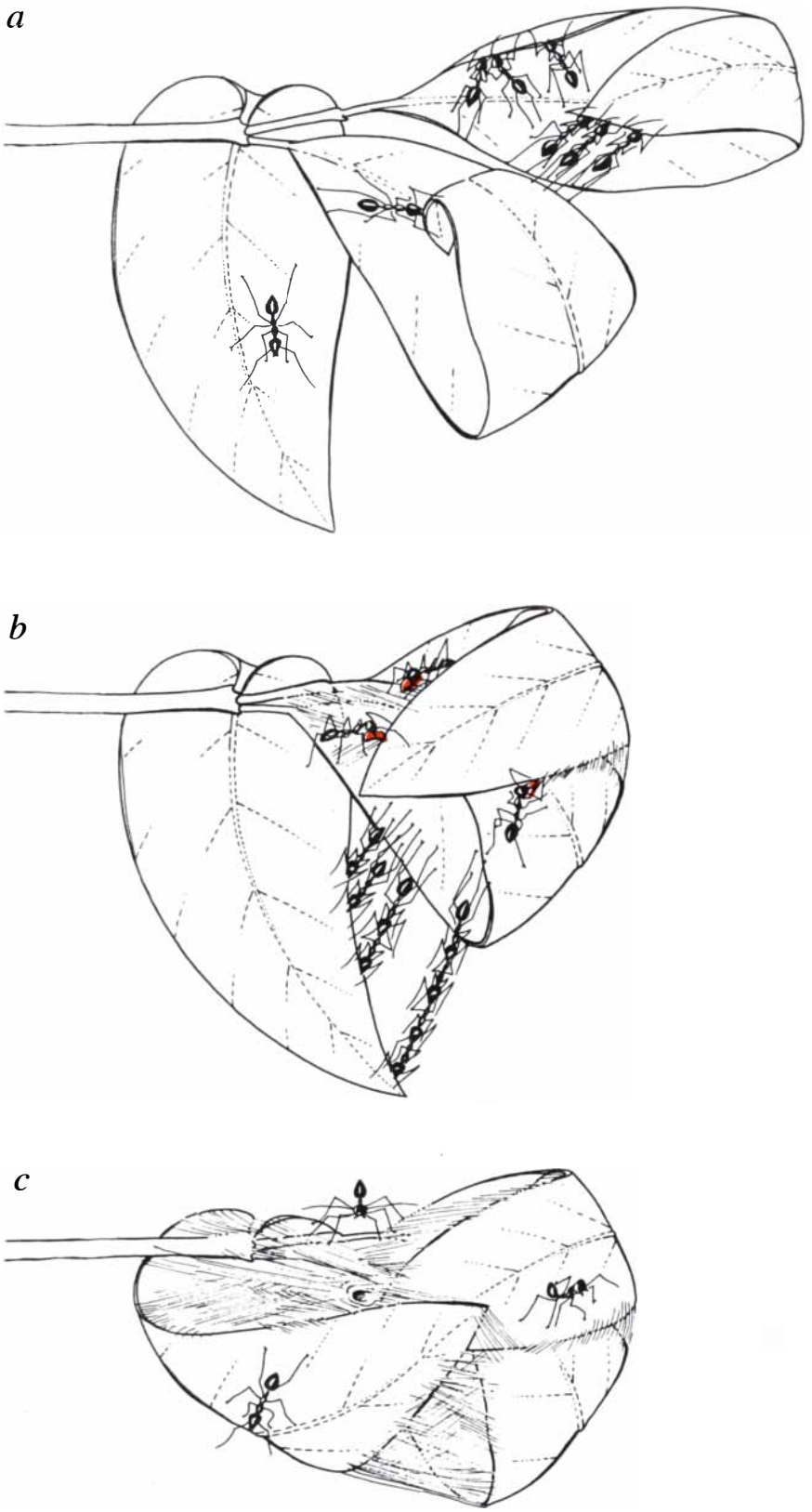
86 PROOF BLENDED SCOTCH WHISKY – RENFIELD IMPORTERS, LTD., N.Y.

After an extension of a weaver ant colony's territory has been secured, the workers sweep back and forth across it in search of food. When a worker encounters a sugary secretion (usually from scale insects), it returns to the nest to recruit nestmates. Once again the rectal gland is employed to lay an odor trail. In this recruitment system the worker stimulates its nestmates by stroking them with its antennae while offering them regurgitated food from the find. Emigration to nests built in the new territory is effected by still another recruitment system, one that includes a rectal-gland odor trail, stroking with antennae and the physical transport of nestmates.

Two more recruitment systems are employed in defending the colony's territory. Weaver ants are particularly aggressive toward members of other weaver ant colonies. In fact, Leston has found that the territories of different colonies are separated by "no-ant's-lands," that is, narrow zones into which few ants venture. A contact between colonies of weaver ants usually results in immediate, spectacular warfare, numerous casualties and eventually the retreat of one of the colonies from part or all of its territory. In the natural habitat of the African species such battles can last for days as the massed opponents struggle along slowly shifting lines of defense.

Foraging workers that encounter enemy weaver ants react with a series of swift, precise movements. Individual combat is initiated with a dancelike maneuver in which the combatants raise themselves on extended legs and circle each other with stiff, jerky movements. Then they thrust and snap at each other with their mandibles. A defeated ant is pinned spread-eagled to the ground. Its legs and antennae are then clipped off and scattered, and its abdomen is often sliced open as well. Throughout the melee ants crumple dead and dying. Some of the workers rush back to the nest laying rectal-gland odor trails. When the trail layers encounter nestmates, they jerk their bodies in what appears to be a ritualized version of the preliminary combat dance. The nestmates respond, however, not by fighting but by running out along the trail to the battle site.

At the same time the fighting workers employ a shorter-range recruitment system to organize group attacks. When a forager encounters an enemy ant but fails to engage it in combat, the forager often runs short looping patterns while dragging its abdomen over the ground. In this recruitment system the ant rotates the terminal segment of its abdomen to expose its sternal gland. This newly discovered organ is also known only in weaver ants. The ants are attract-



COOPERATIVE NEST BUILDING in the weaver ant is illustrated in this sequence. At first the workers labor independently in their attempts to pull down or roll up leaves. When success is achieved by one or more of them at any part of the leaf, other workers in the vicinity abandon their own efforts and join in (*a*). When the leaves have finally been shaped into tentlike configurations, some of the ants continue to hold them in place with their legs and mandibles while others carry partially grown larvae (*color*) from preexisting nests and bind leaves together with sticky larval silk (*b*). Sheets of silk are then added to create circular entrances and galleries (*c*).

SILK TO BIND NESTS is spun by partially grown larvae, which the adult workers hold in their mandibles and move back and forth over the leaf seam like shuttles. The threads of silk, released from gland openings just below the larva's mouth, are woven into sheets.

COMPLETED NEST was constructed by a colony of the African weaver ant (*Oecophylla longinoda*) in a potted grapefruit tree grown in authors' laboratory at Harvard University. In the wild a colony of some half-million workers (the progeny of a single queen) may construct hundreds of nests from the leaves of one or more trees. The ants patrol this domain by day and withdraw into the nests at night.

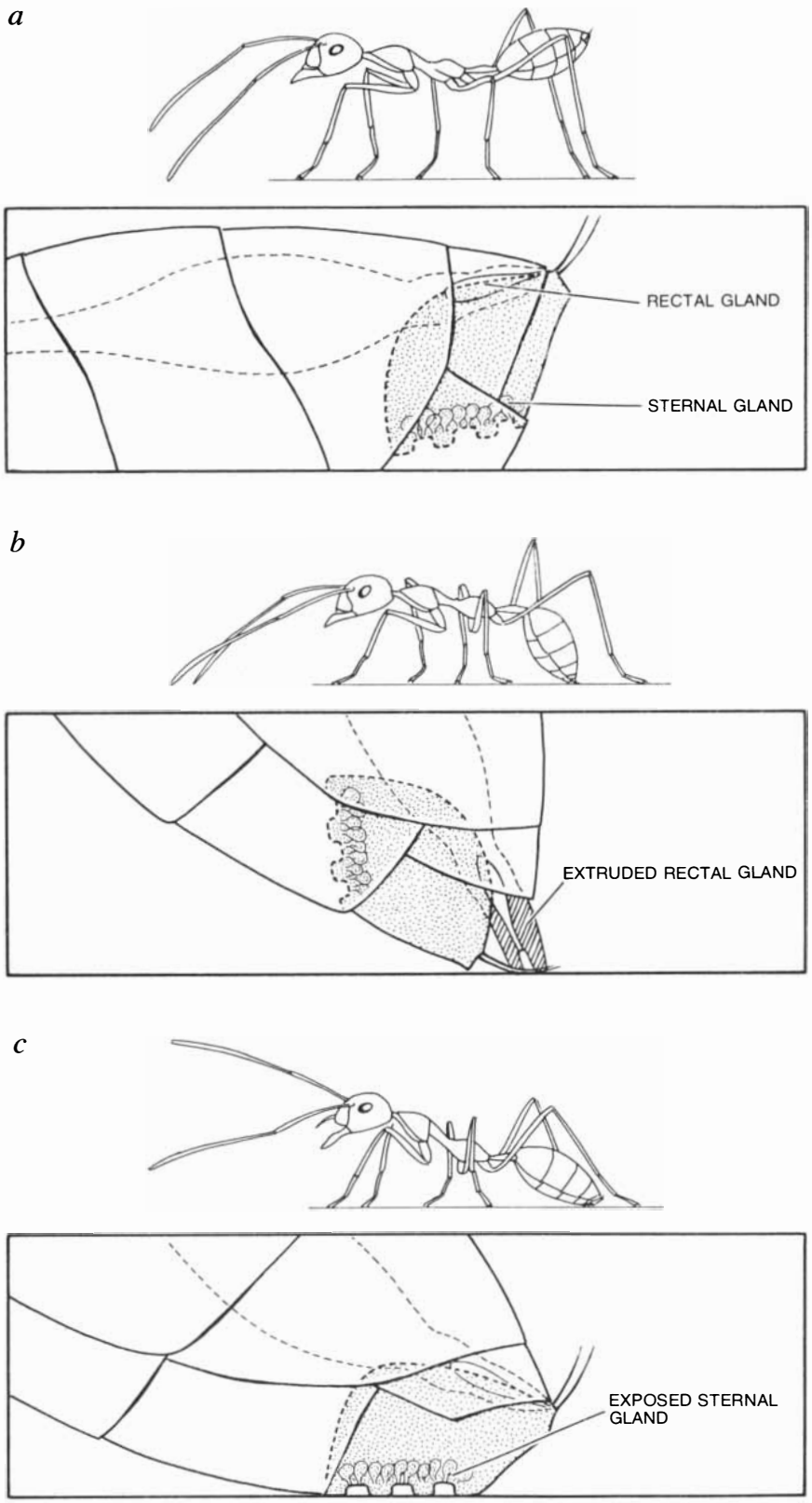
ed by the scent of the sternal-gland secretion from as far away as 10 centimeters. As a result small clusters of ants gather at the places where the enemy was first encountered.

J. W. S. Bradshaw and his colleagues at the University of Southampton have shown that this effect is enhanced by the release of various alarm substances from glands located at the base of the weaver ant's mandibles. Some of these substances attract nestmates to the scene. Others increase the ants' level of excitement and aggressiveness. We have observed repeatedly that these groups of recruited ants are far more effective in combat than individual ants. For example, workers of the large black tree ant *Polyrhachis militaris* can easily knock aside single weaver ants. When three or more weaver ants form a tight group, however, they can seize a tree ant in concert and pin it to the ground. Other workers quickly converge on the spot and assist in the kill.

The weaver ant is a sophisticated judge of odor cues. We noticed that after foragers occupied a new territory in the laboratory they began to deposit large droplets of fecal matter over the surface of the territory. This behavior differs from that displayed by most kinds of ants, which concentrate their excrement in refuse areas and other restricted locations. When the weaver ants patrol their territory, they inspect the fecal droplets. If a weaver ant encounters a droplet left by a member of another colony, it reacts momentarily with aversion, assumes a hostile posture and then inspects the droplet more closely. We have been able to induce the same set of reactions with fluid taken from the hindgut of alien ants.

The fecal substances give weaver ants an advantage when they are the defenders in territorial combat. We arranged a series of eight "wars" between colonies in areas previously marked out with fecal droplets by one or another of the colonies. In each case the members of the colony that had deposited the droplets were less hesitant to forage over the terrain and quicker to recruit nestmates when they encountered alien ants. As a result they gained the initial advantage and secured more ground during the initial fighting.

It appears that the distinguishing characteristics of the weaver ants evolved a long time ago. The African and Asian species are survivors of one of the most distinctive and ancient lineages of ants. Many insect fossils from over the past 100 million years are preserved in amber, which is fossilized resin. Fossils of two extinct species of weaver ant, *Oecophylla brevinodis* and *Oecophylla brischkei*, are found in amber that was deposited in the area of the Baltic Sea some 30



ODOR TRAILS are laid down by weaver ants to coordinate social activity. Normally the foraging worker ant walks with its abdomen elevated (*a*). When the ant encounters a new terrain or food source, it lowers its abdomen, extrudes the rectal gland through the anus and deposits a pheromone, or message-bearing substance, in a line along the ground (*b*). While the rectal gland is extruded it rests on a tiny sled consisting of two bristles. Nestmates that have been tactilely stimulated by the trail-laying ant will follow the trail to a new terrain. Short-range recruitment of major workers to fight invading ants and other intruders is achieved by exposing the sternal gland of the abdomen and dragging it over the ground to deposit a short, looping odor trail (*c*).

million years ago in the Oligocene epoch. During that period northern Europe contained both tropical and Temperate Zone forests. The fossil record shows that many of the insects of those forests resembled the insects of similar environments in the Europe and tropical Asia of today. In particular the extinct species of *Oecophylla* are related more closely to the modern Asian species than to any ants that are now found in Europe.

Some 15 years ago a first glimpse was obtained of the social organization of the extinct *Oecophylla*. In 1963 Mary Leakey was searching for fossils on Mfwangano Island in Lake Victoria in Kenya. She uncovered an assemblage of ant fossils: 366 tiny crystalline ants clustered in a single spot. The ants apparently had been living in a leaf nest that had fallen into a freshwater pool, where the nest and its inhabitants were quickly covered with sediment. Under those conditions an unusual amount of surface detail was preserved. One of us (Wilson) identified the assemblage as a portion of a colony of extinct weaver ants. It is the first and so far the only insect society that has been found in the fossil state. The assemblage includes clusters of larvae and pupae. Some of the fossil ants are still attached to fragments of leaves. Since there was a small population rather than the usual single fossil specimens, it was possible to make a statistical study of the caste system of the fossil species. The anatomical characteristics of these ants and the relative abundances of the two worker castes turned out to be quite similar to those that, among modern ants, are unique to the *Oecophylla* species.

The most distinctive element of the caste system of modern weaver ants is the minor worker: the caste of smaller and less numerous ants specializing in the care of eggs and small larvae. In most ant species that have more than one caste it is the major workers that are anatomically deviant and less numerous. The fossil population discovered by Mary Leakey was in a lower Miocene deposit and is therefore at least 15 million years old. The unusual anatomy and size distribution shared by the Miocene and modern weaver ants suggest that the peculiar division of labor among the living species of *Oecophylla* is of great age.

As the evidence from the fossil weaver ants implies, advanced social organization confers an evolutionary stability on the insect societies. This stability has advantages and disadvantages. On the one hand our studies suggest that individual workers perform no more than 50 distinct behavioral acts, most of which are for the purpose of communication. The result of this almost exclusively social orientation is that the weaver ant colony is an extremely successful working unit. On the other hand, the effectiveness of the colony is obtained by the rigid programming of the relatively simple components of individual behavior, which secures a complex but lockstep pattern of cooperation during group activities. It appears that the colony can flourish only at the expense of any semblance of independent action on the part of the individual. In other words, the weaver ants, like the honeybees, seem to have reached the extreme of one spoke of adaptive radiation among the social insects.

SYSTEM	FUNCTION	CHEMICAL SIGNALS	TACTILE SIGNALS	PATTERN OF MOVEMENT
Recruitment to food	Recruitment of major workers to immobile food source, particularly sugary materials	Odor trail from rectal gland and regurgitation of liquid from crop	Touching with antennae, head-waving and mandible-opening associated with the offering of food	Looping trails laid around food source, with main trail leading directly to nest
Recruitment to new terrain	Recruitment of major workers to new terrain	Odor trail from rectal gland	Touching with antennae and occasionally jerking body back and forth	Broad, looping trails laid around new terrain with deposition of hindgut material containing territorial pheremone; main trail leads directly to nest
Recruitment during emigration	Emigration of members of colony to a new nest site	Odor trail from rectal gland	By touching with antennae, one ant indicates its readiness to carry another to new nest site	Main trail leads directly to nest site without additional looping trails. Workers carry first mostly larvae and pupae, then other workers to nest site
Short-range recruitment to enemies	Short-range recruitment of nestmates for assembly and more rapid capture of invaders and prey	Short, looping odor trails from sternal gland; exposure of gland surface when abdomen is lifted in air	None	Short, looping trails limited to the vicinity of contact with enemy
Long-range recruitment to enemies	Long-range recruitment of major workers to fight invaders. Particularly intense during territorial wars with members of same species	Odor trail from rectal gland	Touching with antennae. During periods of greater excitement, body is jerked back and forth	Main trail leads directly to nest

RECRUITMENT SYSTEMS of the weaver ant are summarized in this table. The two pheromones secreted by the rectal gland and the sternal gland, when combined with tactile signals and the spatial configuration of the odor trail, can communicate five different messages.

Silent and Sound reasons to give a Kodak Moviedeck projector.

Kodak movie projectors make showing movies easy. Count the Kodak benefits:

1 Watching movies without setting up a screen or turning down room lights. Our unique pull-out viewing screen is ideal for small audiences. There's also the option of projecting movies onto a traditional screen.

2 Bright images.

3 Automatic threading and automatic rewind.

4 Showing super 8 and 8 mm movies at the flip of a single switch on all models.

5 Replaying previous scenes in just seconds, and a choice of fast or slow motion on some models.

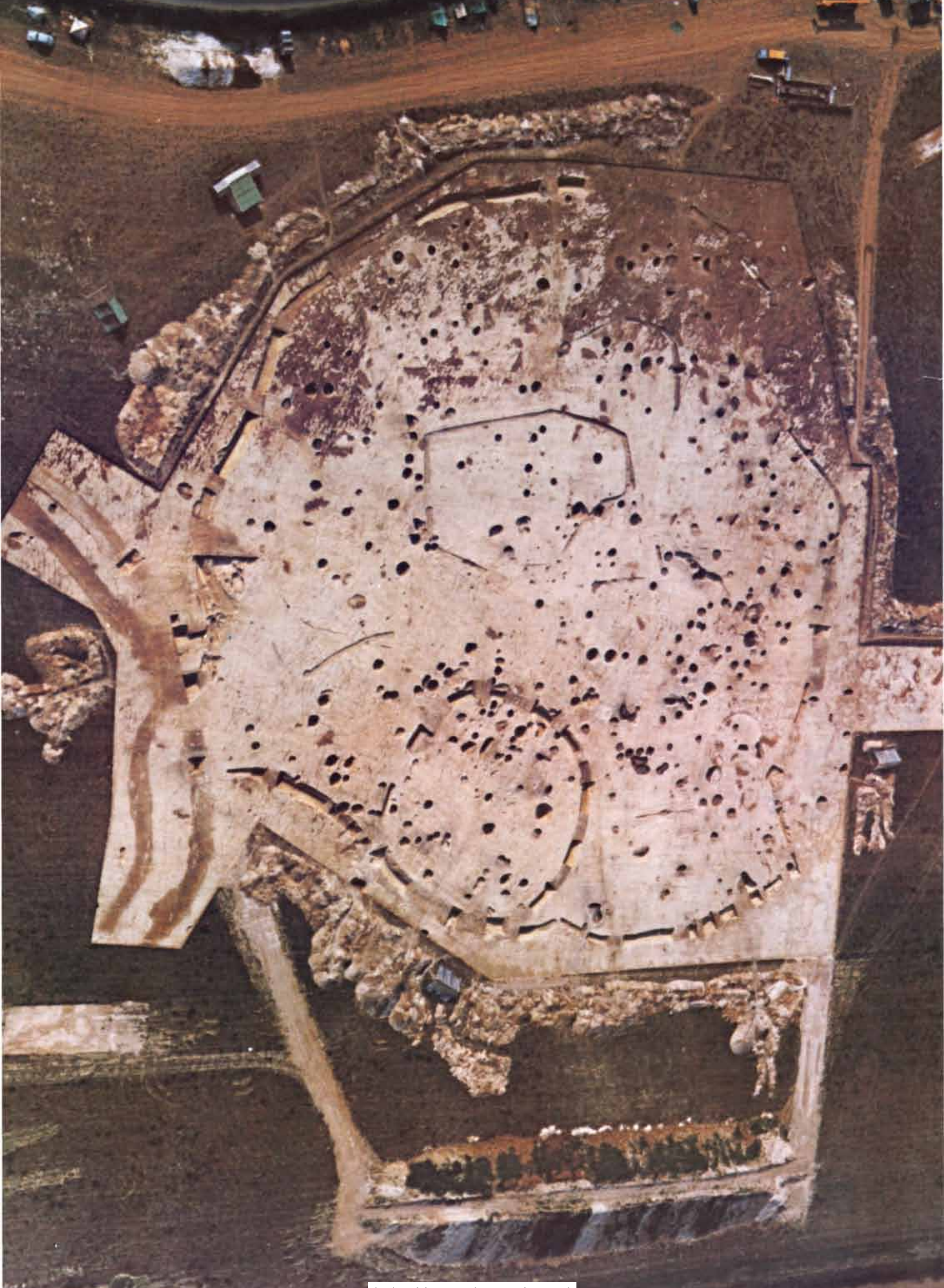
6 Low-profile design and smoke-tinted dust covers that make these projectors easy to look at.

Visit your photo dealer's soon. Ask to see a silent or sound model. And see all the silent and sound reasons to give your family a Kodak Moviedeck projector.

Kodak

Kodak Moviedeck® projectors

Kodak gifts say: "Open me first"
...to save Christmas in pictures.



A Celtic Farmstead in Southern Britain

From the sixth century B.C. until the decades following the Roman conquest of Britain three successive settlements of Celtic farmers and herdsmen occupied a ditch-surrounded enclosure in Dorset

by Geoffrey Wainwright

In the last centuries before the Christian Era, Celtic tribes controlled most of Europe between the Danube and the Atlantic. Indeed, Celts once even sacked Rome in its heyday. In the same period successive waves of Celts came to populate the greater part of Britain. The immigrant tribes ruled the island until the Roman conquest, and afterward, as Romanized Celts, they retained much of their authority up to Anglo-Saxon times. Even those who are students of ancient history, however, know little about rural life in Celtic Britain. For this state of ignorance British archaeologists must bear part of the blame.

Until five years ago the only sizable systematically excavated Celtic rural settlement in England that had yielded significant socioeconomic information about pre-Roman Britain was Little Woodbury, an enclosed settlement in Wiltshire near Salisbury. Gerhardt Bersu directed the work there in 1938 and 1939 on behalf of the Prehistoric Society, but the investigation was interrupted when only a third of the enclosure had been uncovered. Even partial excavation allowed some useful deductions about the culture, society and economy of Celtic Britain, but the incompleteness of the work consistently blocked firm conclusions on many points.

Subsequent fieldwork and aerial reconnaissance in southern England revealed a number of settlements comparable to Little Woodbury, including a good-sized Celtic farmstead in Dorset some 15 miles north of the English Channel port of Bournemouth. The original earthworks of the settlement had long since disappeared under the plow, but the outlines of a three-acre enclosure were visible on aerial photographs made by J. K. S. St. Joseph. In size and configuration the enclosure closely resembled Little Woodbury. Thus the new site offered an opportunity to extend Bersu's investigation, this time by means of a total excavation. Work was begun at the Gussage enclosure, so named because of its proximity to the present-day village of Gussage All Saints, in 1972 on behalf of the British Department of the Environment. Because of continuing plow erosion the excavation was classified as an archaeological rescue operation.

The Gussage site lies in the heart of territory occupied in the first century B.C. by a Celtic tribe, the Durotriges. The tribe controlled Dorset and parts of adjacent Wiltshire, Somerset and Devon. Their territory and that of neighboring Celtic tribes is marked by hilltop fortifications such as Hod Hill and a great 45-acre enclosure known as Maiden Castle. The hilltop defenses were occupied by various Celtic groups for some 600 years before this part of Britain was pacified in A.D. 43–44 by the Second (Augustan) Legion, commanded by the future emperor Vespasian.

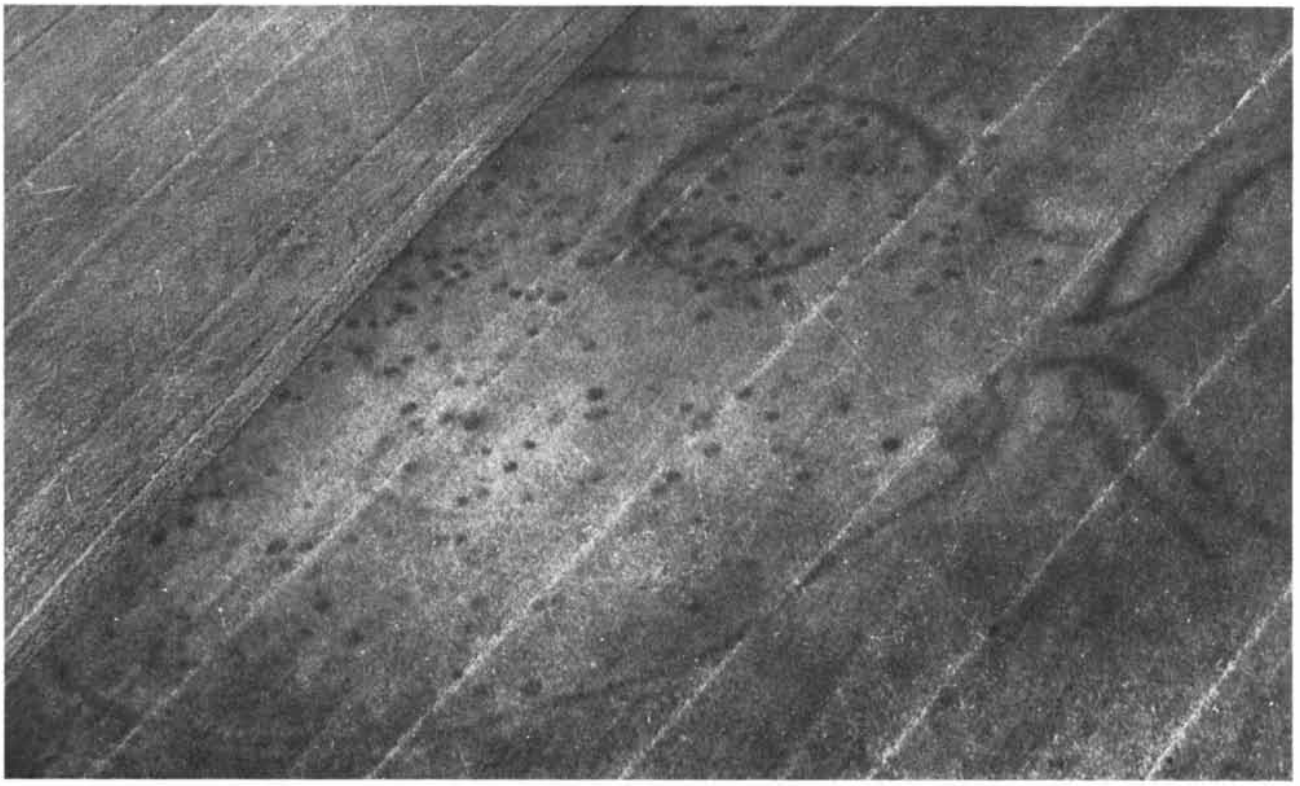
Our excavation showed that the Gussage farmstead had been occupied for roughly the same period of time. It had been founded in about 550 B.C. and was finally abandoned in the third quarter of the first century A.D., not long after the arrival of the Romans. Structurally the development of the settlement took place in three phases.

The first settlement at Gussage, established early in the British Iron Age, spanned some 250 years, from about 550 B.C. to about 300. Its perimeter was marked by a shallow ditch a little more than a meter wide and not quite a meter deep; the excavated soil was used to raise a bank on the outer edge of the ditch. The main entrance to the enclosure was on the east. It was flanked by a pair of subsidiary "antenna" ditches that funneled approaching traffic toward an eight-meter-wide causeway that crossed the main ditch and terminated at a timber gate. The combination of the antenna ditches and the gateway was evidently a defensive feature.

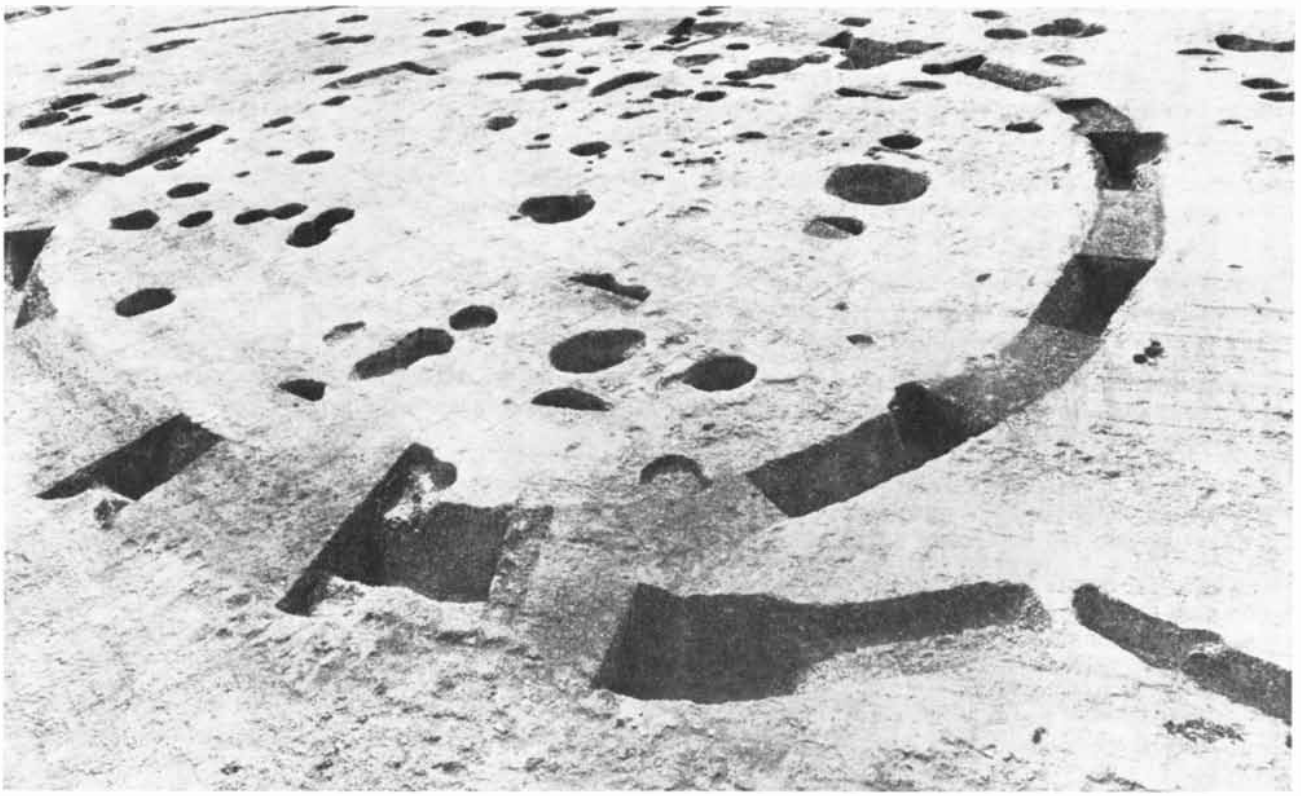
The structures built inside the first Gussage enclosure are known to us only from the postholes that once supported their timber uprights. By the same token the activities of the farmstead are known only from the contents of various rubbish-filled pits associated with the first structures and from the debris found in what are called "working hollows" within the enclosure. Most of the pits and hollows were located some distance away from the perimeter ditch. The pattern of their distribution is distinctive: they are adjacent to groups of postholes many of which form squares, with a post at each corner. The sides of the squares range in length from 1.7 to 2.5 meters.

The square posthole arrays constitute the only evidence from the first-phase settlement at Gussage of what may have been dwellings. It should be kept in mind, however, that centuries of plow erosion have obliterated all evidence of any but quite large structures. Between 15 and 20 square arrays can be identified, concentrated in an area some 50 meters long and 40 meters wide at the center of the enclosure. The pits associated with them number 125; in volume the average pit is a little less than two cubic meters, so that the total space available in the pits of the first phase at Gussage is 229 cubic meters. Some of

AERIAL VIEW of the excavated Celtic farmstead near the village of Gussage All Saints, its six centuries of occupation recorded by a succession of ditches, postholes and rubbish pits, is seen in the photograph on the opposite page. Adopting a policy of total excavation, workers under the author's direction cleared the site to save its remains from further plowing damage.



CROP MARKINGS, the dark spots and lines visible in the tilled fields seen in this oblique aerial photograph, led to recognition of the Gussage settlement. North is at the top; most of the perimeter ditch and both pairs of "antenna" ditches are plainly seen, as is the ring ditch dug late in the final phase. Most of the dark spots indicate rubbish pits. Crop markings appear because the individual plants that grow in disturbed rather than undisturbed soil root themselves deeper, have better access to ground water and become a darker green.



RING DITCH of the final phase at Gussage formed a defensive enclosure within the main enclosure of the settlement. It was built in the middle of the first century A.D. Inside the ring a circular timber structure was further defended by a stout timber gateway. The ring ditch rapidly filled with silt and no effort was made to keep it in repair, suggesting that the need for such defenses was short-lived.

the pits may have been used for grain storage, but in the end all were filled with rubbish and the remains of both animals and human beings.

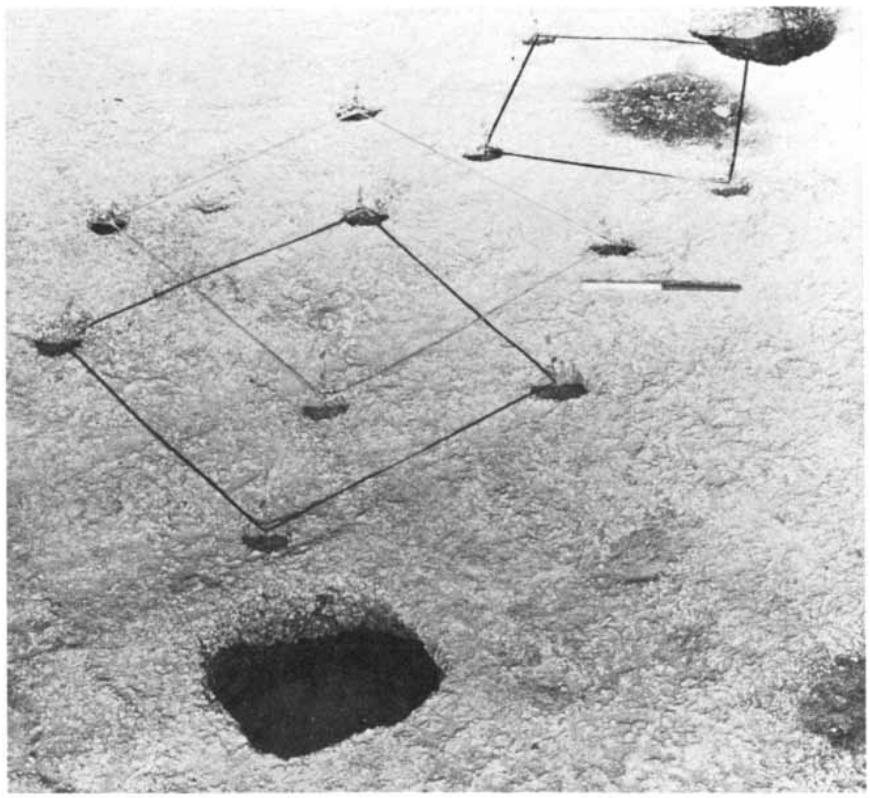
The second phase at Gussage occupied a period from about 300 B.C. to about 100. Some evidence suggests a continuity between the two phases. For example, although the entrance to the enclosure was reconstructed, it was not relocated, and the main ditch, although it was enlarged, followed the original alignment. Moreover, the pits that were dug during the second phase were almost all located at a distance from those of the first phase. The pits are, however, substantially larger than the earlier ones. Only 69 in number, they have a combined capacity of nearly 210 cubic meters.

Architectural evidence from the second phase is meager. The only certain residential structure was circular in plan. As is indicated by the traces of a gully around it, it was some nine meters in diameter. Faint traces of a second such gully suggest that another circular structure existed during this phase, but there is no sign of postholes. As for the ditch around the entire settlement, enlargement in the second phase made it more than two meters wide and nearly a meter and a half deep. Remodeling of the enclosure entrance included the digging of a second pair of antenna ditches. The width of the causeway was narrowed to five meters and a more elaborate timber gateway was built, its main uprights set in a palisade trench. The holes for the uprights are large enough to suggest that the gateway may have been topped by a tower, but the entranceway could not have been more than two and a half meters wide.

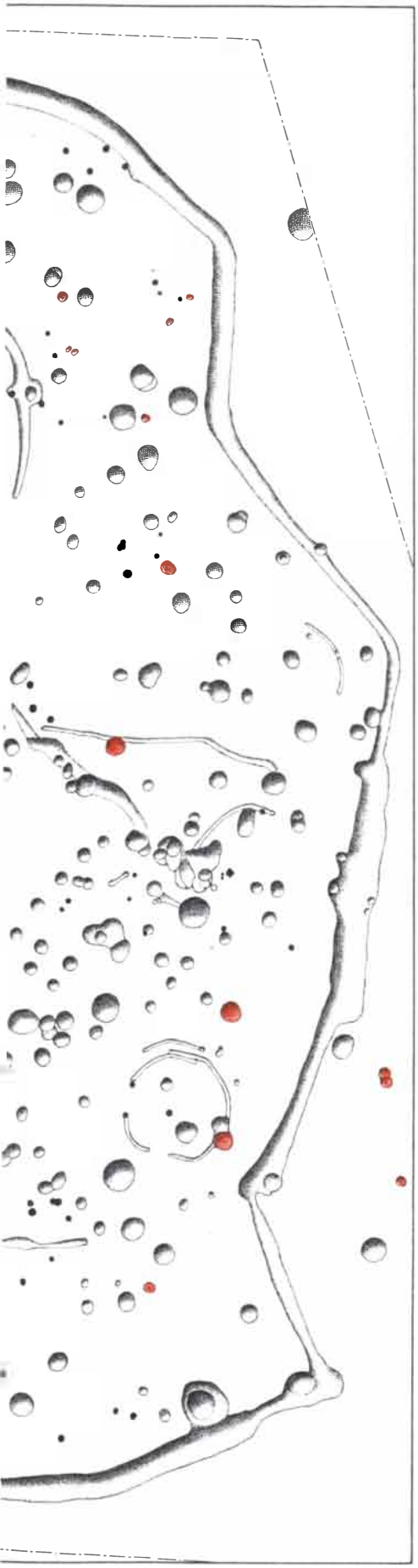
The third and final phase at Gussage lasted from about 100 B.C. until the settlement was deserted in about A.D. 80. Again some degree of continuity between successive phases is suggested by the fact that neither the east entrance nor the perimeter ditch was relocated. The 180 pits that were dug during the final phase at Gussage, however, were distributed around the enclosure without regard to the location of earlier pits. Only slightly smaller than the pits of the second phase, the new pits have a combined volume approaching 500 cubic meters.

During the final phase three subsidiary enclosures were built inside the perimeter ditch. The two most important are respectively a trapezoid that may have served as a livestock pen and a substantial circular earthwork that was probably intended to provide greater security for the structures located at its center. Whatever the inspiration of this ring ditch, it was built late in the history of the settlement: in about the middle of the first century A.D. The ditch silted up

JUMBLE OF BONES was unearthed near the top of rubbish pit No. 293, one of the pits associated with the final phase at Gussage. Included are the broken limb bones and lower jaws of cattle and the skulls of an ox and a horse. The inhabitants of Gussage bred cattle, sheep and goats but obtained horses by rounding up wild ponies and training mature three-year-olds.



POSTHOLE ARRAYS of the first phase at Gussage are marked with string in this photograph to suggest that uprights arranged in squares may have formed the principal supports of residential structures. The squares range from 1.7 to 2.5 meters on a side. It seems improbable that uprights were corner posts because structures so small would have been extremely cramped.



rapidly and was not repaired, suggesting that the need for such a subsidiary defense was short-lived.

So far I have given dates to the successive phases at Gussage without citing any evidence to support them. We did recover a number of charcoal samples suitable for carbon-14 dating. Nine of these were sent to V. R. Switsur of the Radiocarbon Dating Research Laboratory at the University of Cambridge. Three samples from first-phase contexts yielded calibrated age determinations that ranged from 790 to 420 B.C. Two samples from second-phase contexts yielded calibrated readings ranging from 410 to 140 B.C. and two more samples from third-phase contexts ranged from 50 B.C. to A.D. 150.

Such chronological ranges are very broad, and the true age of any event can fall anywhere between the two extremes. For the abandonment of the Gussage farmstead, however, somewhat more precise data are available. A type of red pottery much admired by the Romans and rather loosely classified as "Samian" ware underwent such continuous stylistic changes that different Samian wares may safely be considered indicative of different imperial periods. Five of the rubbish pits at Gussage yielded shards of Samian ware associated with the periods of the Roman emperors Claudius (41–54 A.D.) and Flavius (69–79 A.D.). Their presence demonstrates a continuation of the settlement beyond the second half of the first century A.D. and probably into the last quarter of the century. The oldest of the carbon-14 determinations, in turn, indicate an initial settlement of Gussage sometime before the middle of the first millennium B.C.

Except for the pottery the artifacts recovered from the rubbish pits and working hollows at Gussage are not particularly numerous. Nevertheless, together with the organic debris associated with them they constitute evidence of both a social and an economic nature concerning the lives of the Iron Age Celts of rural Britain. As an example, in the area of technology the saddle quern, a fairly primitive device for milling grain, was still in use early in the first phase at Gussage. Not long after the settlement was founded, however, a much more efficient milling device, the rotary quern, was introduced. As another example, narrow combs made of bone are artifacts indicative of some degree of sophistication in the manufacture of cloth. They are virtually absent from the remains of the first-phase settlement but are relatively numerous in the second phase.

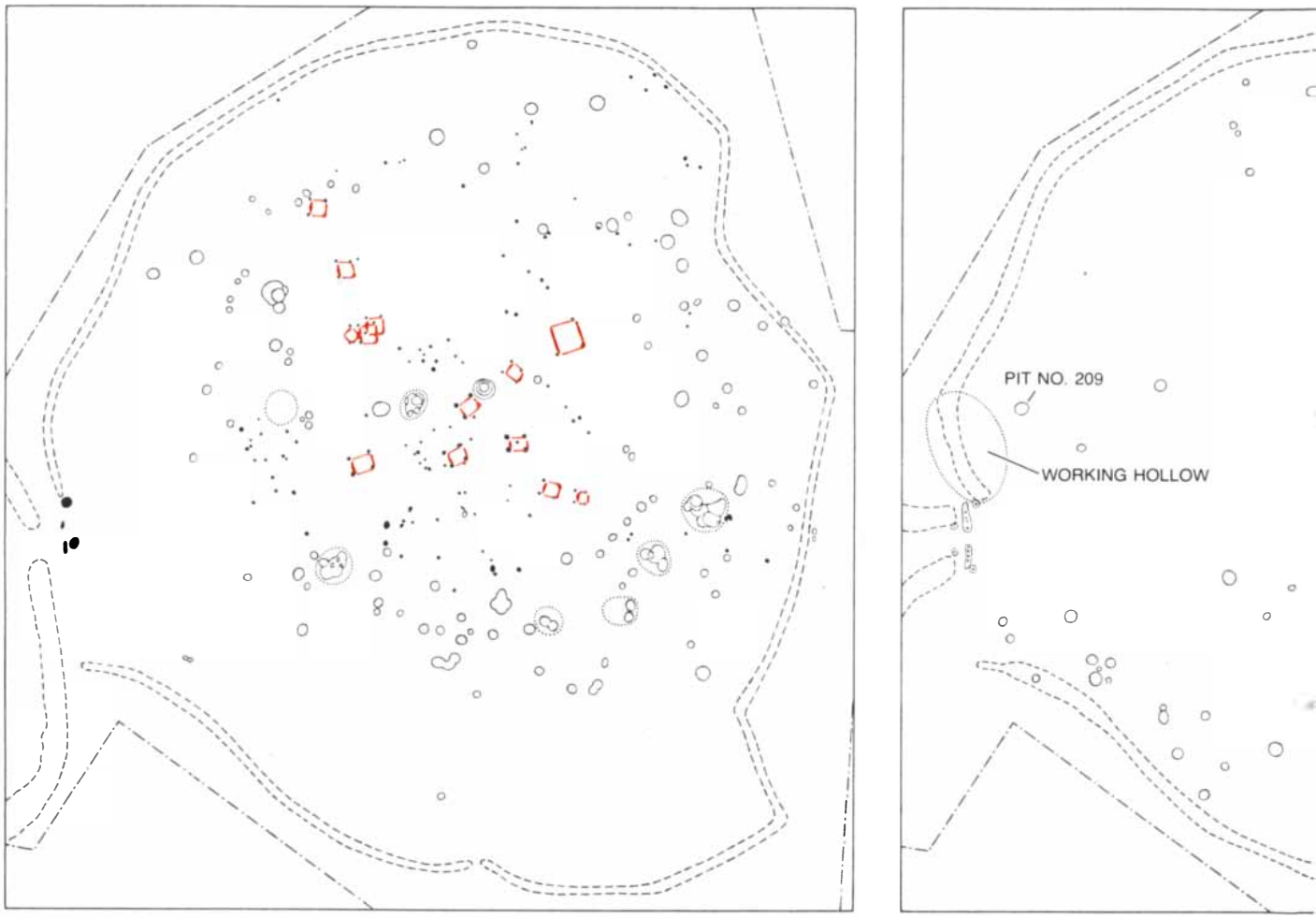
Still another technological advance is indicated by the pottery at Gussage. The wares of the first and second phases are conventional, hand-formed ones typical of this period in southern Britain, for example globular and shouldered jars, fine bowls coated with hematite, saucepan pots with vertical sides and jars with thick pedestal bases. In the final phase at Gussage, coincident with the introduction of the potter's wheel, pottery styles evolved rapidly. The wheel-flung wares include tankards, large jars with countersunk lugs and bowls with foot rings, to mention only a few.

Personal articles are not numerous in any of the three phases. Typical of those found in the first-phase rubbish deposits at Gussage are ring-headed iron pins, brooches made of iron and of bronze (including one bronze brooch typical of the Celtic period recognized across the Channel as La Tène 1-*a*), bangles made of bronze and of slate, glass beads and bone toggles. The second phase is impoverished even in such ordinary articles as these. In the final phase brooches, bone pendants and a variety of toilet articles appear in increasing numbers.

The homelier artifacts include iron tips for plowshares and iron knives with distinctive offset tangs. An equal-arm balance made of bronze bespeaks the practice of weighing, as does an iron weight for a steelyard scale. Objects that might be taken to connote social status, such as scabbard fragments indicative of sword-bearing, are found in the rubbish of all three phases. The possibility exists, however, that they are accumulations of scrap metal; both iron and bronze metallurgy were practiced at the settlement throughout its history.

Carbonized seeds and similar evidence of agriculture at Gussage indicate that the people of the farmstead raised wheat and barley, and probably more of the former than of the latter. The wheat was of a variety now rare: spelt (*Triticum spelta*). Oats were also cultivated, and peas and beans were grown from the first phase at Gussage onward. Indeed, the increasing abundance of legumes over the centuries suggests that they came to be increasingly important as foodstuffs. Traces of the fields where these crops were sown are still to be seen to the north, west and south of the farmstead. After the grain was harvested some or all of it was evidently dried before threshing and stor-

MAIN FEATURES of the three successive phases of the Gussage settlement are represented in the plan at the left. The orientation of the plan is the same as that of the photograph on page 156; east is to the left. The enclosure, some 110 meters across from east to west and more than 120 meters from north to south, covers three acres. The antenna ditches outside the east entrance served to funnel traffic into the gateway. The color indicates a pit of unknown date.



THREE OCCUPATION PHASES at Gussage are dated from 550 B.C. to 300 (*left*), from 300 B.C. to 100 (*middle*) and from 100 B.C. to A.D. 80 (*right*). The ditch surrounding the first settlement was *V*-shaped, 80 centimeters deep and 1.2 meters wide. Two large postholes locate the gateway on the east side of the enclosure; the antenna ditches outside the gate were three meters wide. Many postholes within the enclosure form square arrays; they may have contained the main uprights of residential structures. Thirteen of them appear in color on the plan. Seven working hollows (*dotted lines*) are associated with the first phase at Gussage. The rubbish pits of this phase have a combined storage capacity of 229 cubic meters. Continuity between the first and second phases is suggested by the fact that although the en-

age. Clay ovens, many of them decorated, were used for this purpose; the fragmentary remains of these drying hearths are most numerous in the final phase. Some of the pits dug at Gussage may have served for storing the winnowed grain.

On the basis of the 15,500 specimens of animal bone recovered at Gussage the animal husbandry of the inhabitants can be reconstructed with some certainty. The species represented by this abundance of bone include horses, cattle, sheep, goats, pigs, dogs and cats. The bones more or less equally represent all parts of the body of the seven species. All age groups from fetal and newborn to advanced old age, with the exception of those of the horses, are represented in proportions that suggest dynamic "natural" populations.

The animal remains were analyzed by R. A. Harcourt. He found that they displayed a pattern of peak mortality in the age range between four and six years; the pattern remained the same throughout the history of the settlement. Among cattle such a mortality pattern indicates the systematic culling of those steers that were not required as draft animals; they were presumably slaughtered for their meat. Cows were also culled, perhaps because of infertility or for poor milking performance or other defects.

Among the sheep the mortality pattern indicates a 5 percent natural death rate soon after birth and a 16 to 26 percent selective kill of lambs not required for the upkeep of a flock. That is a pattern one would expect if the flock were maintained primarily to yield wool and milk, with meat a secondary product.

The age range exhibited by the numerous horse bones places these animals apart from all the other domestic stock at Gussage; neither newborn colts nor yearlings are represented. The finding allows for only one interpretation: horse breeding was not practiced. Instead wild ponies must have been rounded up periodically and mature three-year-olds must have been selected from the herd for breaking to harness. Most of the Gussage horses fall within the size range of the British Iron Age breed: from a little less than 10 hands (110 centimeters) to 14 hands (145 cm.) at the shoulder. (A typical present-day saddle horse stands about 17 hands.)

Some of the inhabitants of the Gussage farmstead must have been occupied in "cottage" industries. For example, a number of the pots from the final phase have a perforated base, a feature that suggests their use as whey strainers in cheesemaking. The Gussage herds and flocks may well have provided a milk supply in excess of domestic requirements, and such a surplus could most usefully have been converted into cheese. Thus it seems a safe assumption that cheesemaking was a Gussage in-

trance was reconstructed, it was not relocated, and although the perimeter ditch was deepened and doubled in width, it followed the same alignment. Of the dwellings within the enclosure during the second phase only some postholes associated with one circular structure escaped destruction by plow erosion. The capacity of the 69 pits dug during the second phase is nearly 210 cubic meters. One pit, No. 209, adjacent to a large working hollow, contained the debris from the casting of bronze fittings for pony harness and chariots during the second century B.C. The main features of the final phase at Gussage are the remains of enclosures, possibly pens for livestock, and a ring ditch enclosing numerous postholes near the north boundary. One skeleton from late in final phase, found in pit No. 285, shows signs of violence.

dustry. As another example, among the Gussage artifacts are bone knives that are best interpreted as skinning tools. Perhaps the regular culling of the herds and flocks provided the raw material for an industry in hides. Finally, weaving as a cottage industry is indicated by the presence of the bone weaving combs and numerous loom weights and spindle whorls.

The human skeletal remains unearthed at Gussage represent a minimum of 53 individuals: 15 adults and 38 infants. Only one skeleton was found in a first-phase context; it was that of an infant buried in a pit specially dug for the purpose. Most of the human remains, representing 14 adults and 31 infants, were found in third-phase contexts. They had been casually interred in rubbish pits and ditches with little or no indication of formal burial, such as uniform posture, consistent orientation of the body or the presence of grave goods.

Perhaps formal cemeteries existed outside the settlement boundaries, but none has been located thus far. Scattered fragments of human bone are comparatively numerous, however, from the second phase of the settlement, when one adult and six infants were casually interred. This suggests the possibility that some funerary practice such as preliminary exposure of the body was customary and that the bare bones were later unceremoniously disposed of.

Carole Keepax of the Ancient Monuments Laboratory has analyzed the human remains from Gussage. She found a fairly high incidence of a degenerative bone disease (osteoarthritis) and an extremely high rate of tooth decay. The prevalence of tooth decay may be due to some as yet unidentified dietary factor.

The number of healed fractures is also rather high. Of particular interest is the skeleton of a robust young male found near the top of pit No. 285, from the final phase at Gussage. The youth's skull and left arm bones show the marks of blows with a sharp instrument. The blows were struck immediately before death or possibly just after it. The injuries are consistent with face-to-face combat, so that it is likely the young Celt died violently.

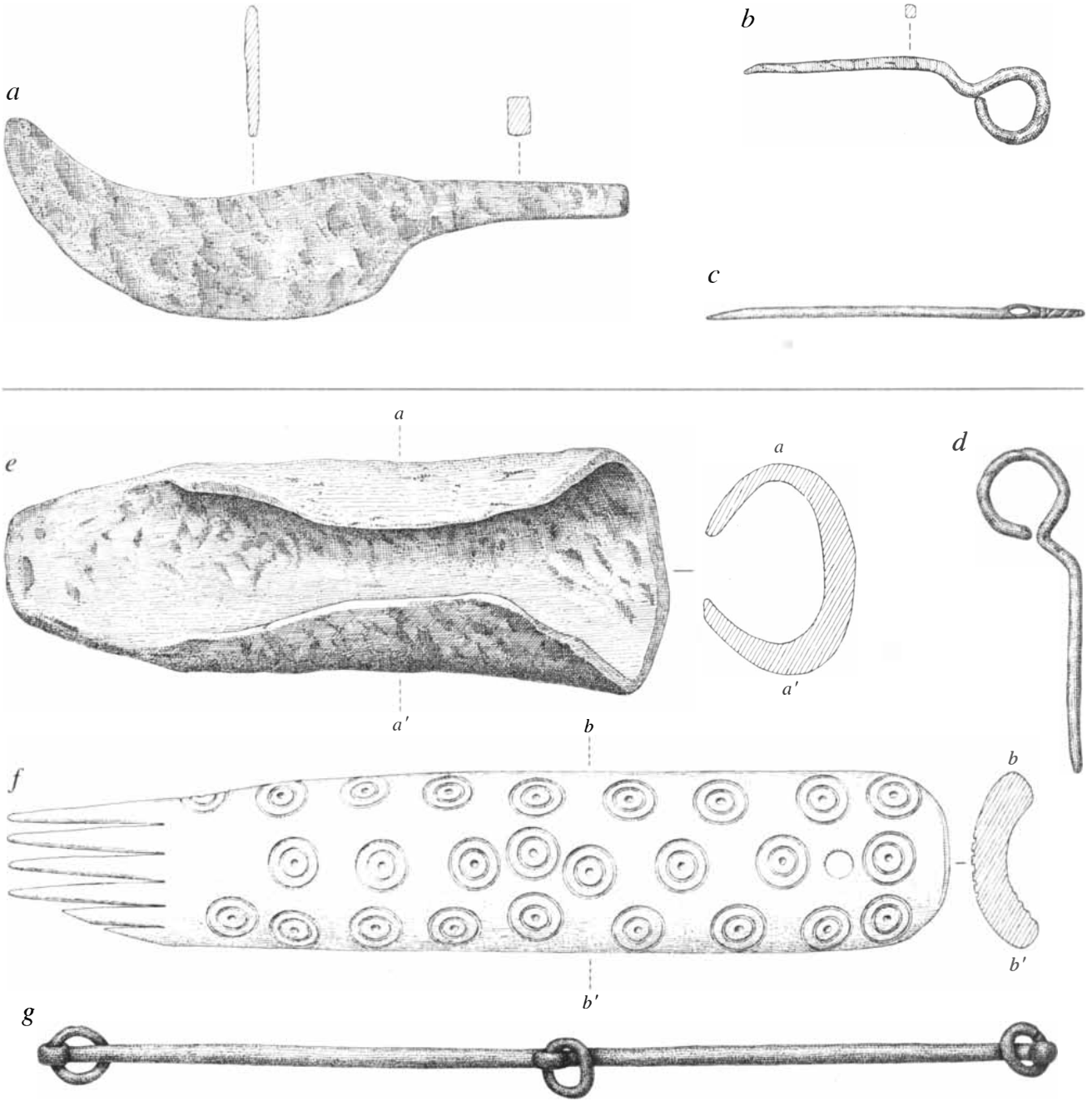
The single most striking discovery at Gussage was the debris from a bronze smithy in a second-phase rubbish deposit: pit No. 209. The find was painstakingly excavated by Karen Stanley, and its contents have been analyzed by an authority on Celtic metallurgy, Mansel G. Spratling of the British Museum. The smith who set up shop at Gussage sometime during the second century B.C. produced some 50 sets of bronze fittings for pony harness and chariots, using the lost-wax method of casting. Spratling estimates that the smith used up, in addition to a supply of

charcoal for his furnace and clay for his molds, some 10 liters of beeswax in the process of smelting and casting a total of 38 kilograms of bronze.

Perhaps the bronzes were simply made for barter or sale. If chariots were also being made at Gussage and Gussage ponies were perhaps being harnessed to draw them, however, seasoned timber and iron fittings would have been needed for the construction of the vehicles and leather would have been needed for harness. Thus it may be that the farmstead population included individuals capable of organizing the logistics behind such an operation.

No trace of the smithy or of the smith's castings was found. The work area must have been near pit No. 209, however, because the smithy debris shows no sign of weathering; it was evidently thrown into the rubbish pit soon after being discarded. The accumulated debris weighed more than 200 kilograms and occupied a volume of about three cubic meters. The major component is fragments of "investment" molds numbering in the thousands.

In lost-wax casting each carefully shaped wax model of the bronze casting is "invested" in a clay mold that surrounds the model completely except for a pouring hole to admit the metal. The clay investment is then baked. The wax melts and evaporates, and the interior of the mold preserves an exact negative impression of the vanished wax model. Molten bronze is then poured into the mold and allowed to cool; after that the mold is knocked apart to release the cold casting. It was these mold fragments that were dumped into the pit;



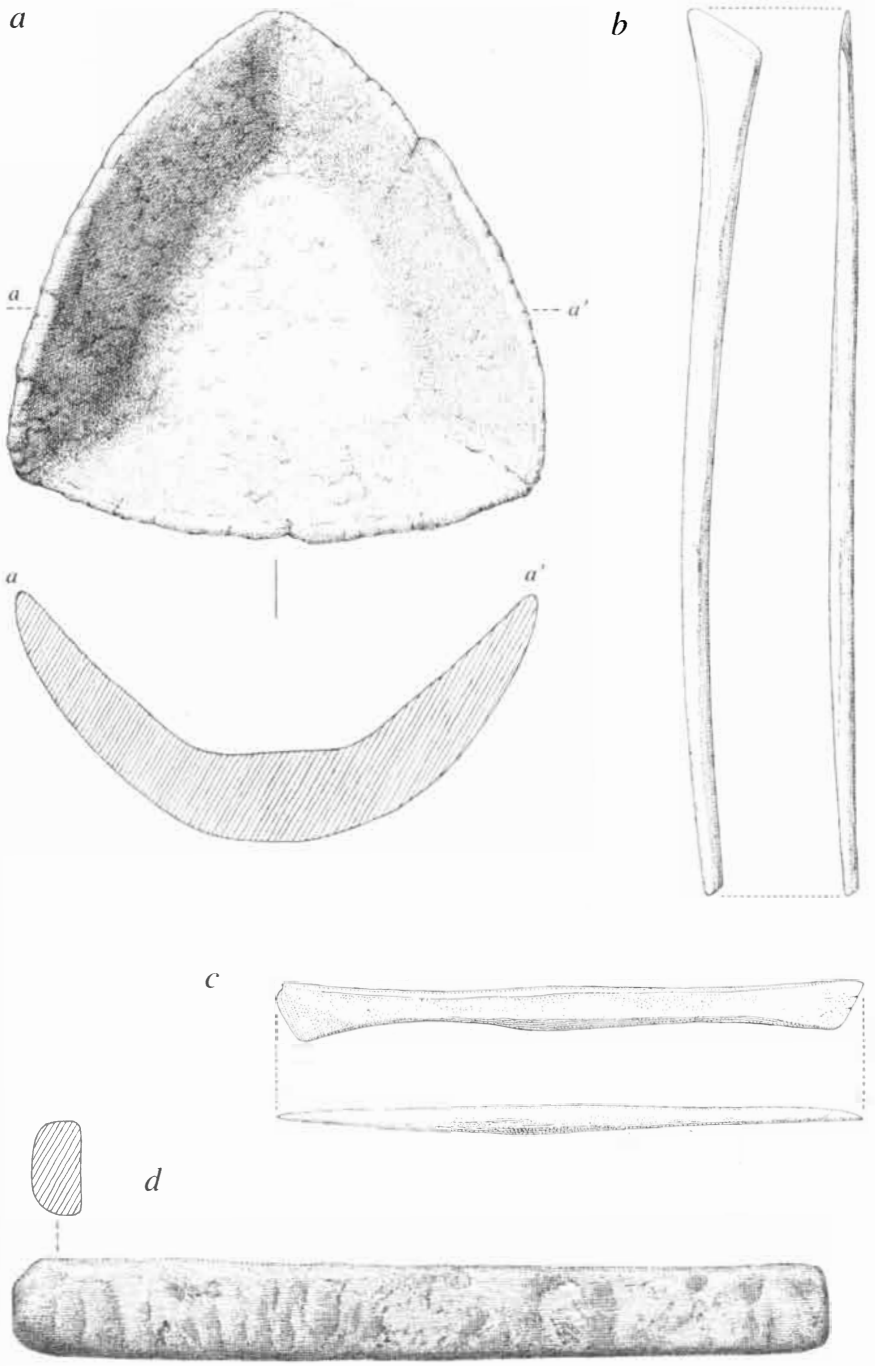
CELTIC ARTIFACTS recovered from the rubbish pits at Gussage include an iron knife (*a*), a roundheaded iron pin (*b*) and a bronze needle decorated above the eye (*c*); all are from the first phase. Another roundheaded pin, this one made of bronze (*d*), an iron tip for a plowshare (*e*), a bone weaving comb (*f*) and a bronze beam balance (*g*) are representative of the artifacts found in second-phase rubbish.

Never buy a Bordeaux by the bottle. Buy a great, velvety Bordeaux by the label.

The B&G label.
B&G. 250 years of winemaking history from all the great wine regions of France: Bordeaux, the Loire Valley, Beaujolais, Côtes du Rhône and Burgundy.
Now B&G brings you Pontet-Latour, a great velvety wine from Bordeaux.
Superb. But affordable.
Just buy it by the label. **B&G.**

Letters of recommendation from France.

IMPORTED BY BROWNE VINTNERS, NEW YORK AND SAN FRANCISCO AD PREPARED BY TT&P, INC.



BRONZE-FOUNDRY RUBBISH, all from pit No. 209, included a clay crucible (*a*), bone tools for modeling wax replicas (*b*, *c*) and an ingot of tin-alloyed bronze, clipped at one end (*d*).

their stratigraphic position showed that the smith had confined himself to making several duplicates of one particular bronze fitting at a time. He then disposed of the broken molds and turned to the duplication of a different kind of fitting.

The three principal fittings cast at Gussage were chariot linchpins (ornamented spikes that fit into sockets at the end of the chariot's axle to keep the wheel in place), four different kinds of harness rings known technically as terrets and bridle bits in the form of three interlocked links. Each finished bit required two consecutive castings. The side links were first cast separately and then combined with a wax model of the center link and invested in a second mold. After this combination was baked to remove the wax the center link was cast in place around the two side links.

In addition to the broken molds, pit No. 209 contained slag, broken crucibles, bits of furnace matrix, many fragments of clay furnace-draft pipes, vari-

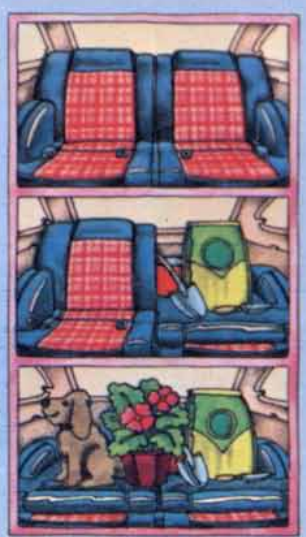
You're gonna love our Great Little Car. Mazda GLC 5-Door Hatchback.

Great little piston engine. Great little goodies. Great little price.

Introducing the Greater Little Car. The GLC Deluxe 5-door Hatchback. $3594**! A roomy new Deluxe GLC in a whole line of roomy GLCs for 1978.

The GLC 5-door is for those who believe the only thing better than a GLC hatchback with three doors is a GLC hatchback with five.

It has GLC's split rear seat, standard. Which lets you take a back-seat passenger along, along with two weeks' groceries or golf clubs or luggage or what have you. Both seats folded down reveals a plushy carpeted deck. Very nice.



SPLIT REAR SEAT

There's a nifty little button under the dash that electrically unlatches the hatch. (You may open the hatch from the outside, too.)

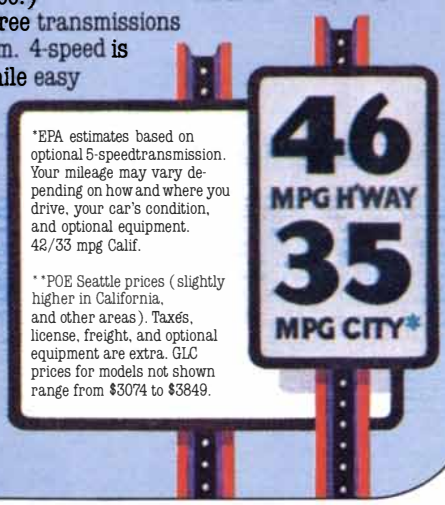
CHOICE OF TRANSMISSIONS

It has three transmissions to choose from. 4-speed is standard, while easy automatic and frugal 5-speed are available.

There's simply much more than you'd ever expect in a car its size. But, for not much money. And that's the case with every GLC model: the GLC Standard 3-door Hatchback, GLC Deluxe 3-door Hatchback, and GLC Sport.

Mazda's Great Little Car is a great little car. You're gonna love it.

REMOTE HATCH RELEASE



*EPA estimates based on optional 5-speedtransmission. Your mileage may vary depending on how and where you drive, your car's condition, and optional equipment. 42/33 mpg Calif.

**POE Seattle prices (slightly higher in California, and other areas). Taxes, license, freight, and optional equipment are extra. GLC prices for models not shown range from $3074 to $3849.

mazda

GLC. Now, four great little cars in all, from $3074.**

ous smithy tools made of iron and bone, lumps of unused mold clay and a single unused ingot of tin bronze weighing almost exactly a tenth of a "heavy" Celtic pound (638 grams). In all, the metallurgical debris from Gussage is unrivaled for quantity, variety and state of preservation anywhere in pre- or protohistoric Europe.

Analysis of the mold fragments shows that the smith also cast button fasteners, loop fasteners and strap unions. Of the four different kinds of terret one kind was produced in at least 11 different variations. Of the 50-odd sets of bronzes, those with simple ornamentation outnumber the more elaborate sets roughly 20 to one. Such an emphasis on simplicity is in marked contrast to the usual concept of elaborate Celtic horse trappings that has grown up over the years on the basis of chance finds.

To judge from findings at other Celtic settlements, the periodic roundup of wild ponies that is indicated by the age structure of the horse remains at Gussage seems to have been something of a general practice in Celtic Britain. It is likely that the practice involved the cooperative effort of several settlements. It may thereby have engendered commercial and social intercourse among tribal neighbors. An additional piece of zoological evidence is similarly suggestive of commercial interchange. Polled, or hornless, cattle are thought to have arisen as a dominant mutation among domestic horned herds; the offspring of the polled mutants would themselves have been polled. At Gussage and elsewhere

BROKEN MOLDS, used in the lost-wax casting of horse and chariot bronzes during the second phase at Gussage, preserve negative impressions of the evaporated wax replicas. At the top is part of the mold for a side link of a three-link bit. Below, at the left, is another fragment of side-link mold with a decorative whorl and, at right, a triskelion decoration for a chariot linchpin.

in Durotrigian territory the skulls of polled cattle are found together with horned skulls. If the polled cattle were indeed of mutant descent, their presence amid horned stock is a strong indication of the practice of exchanging cattle.

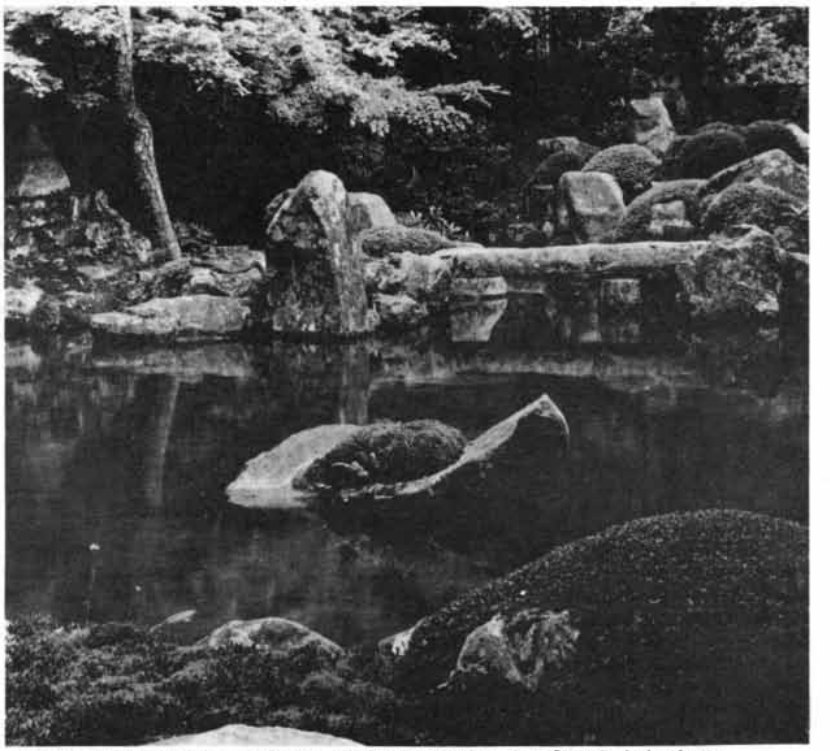
More concrete evidence of trade is furnished by the querns used for grinding grain. These were not made at Gussage but were obtained from stoneworking centers located at least eight miles away. Similarly, the handmade pots of the earlier Gussage settlements were produced from local clays either by farmstead residents or by itinerant specialists working with local materials. The wheel-flung wares of the final phase at Gussage, however, were obtained from specialized production centers elsewhere in southern Britain.

It is not hard to conceive of professional quern-cutters and potters who were able to flourish in southern Britain by trading their products to settlements such as Gussage. The Celtic farmsteads could have offered in return not only raw wool, honey, beeswax, livestock and grain but also cheeses, dressed hides and, at Gussage, even metalwork. Assuming that such surplus was indeed available for trade at Gussage, it may seem surprising that our excavation has given rise to no articles of prestige value: swords, lance blades, shields, helmets, vessels of precious metal or some of the famous Celtic jewelry. It must be remembered, however, that the occupation levels at Gussage, particularly the later levels, have been subjected to centuries of plow erosion.

It would in any event be unwise to assign the Gussage settlement a lowly position in contemporary Celtic society even if no prestige items ever existed there. Consider the farmstead's stout timber gateway with its flanking earthworks, the inhabitants' full granaries, their herds, flocks and trained ponies, their weavers, metalworkers and other craftsmen. Even in complex societies there is no certain correlation between status and the ownership of prestige articles. Celtic society, to the extent that it can be reconstructed, exhibited such disparate traits as frankness and personal vanity, bravery and boastfulness, high spirits and short temper, gluttonous feasting and a love of epic verse. In such a milieu status derived not so much from a man's property as from his personal example and his contributions to his tradition-imbued community. Attributes such as these, of course, are seldom apparent in the archaeological record. It is nonetheless not impossible to imagine that the levy from Gussage held an honorable place in the ranks when a Durotrigian war party assembled and that the levy may have included chariot ponies. Even the young Gussage man who met a violent death may have died defending a Celtic way of life already on the wane in Britain.

Freeman Books on Land and Landscape

Sierra Nevada, photograph by Shelton from *Geology Illustrated*



Rocky Mountain vegetation, and Garden of Rengeji Temple, Kyoto from *Plants in the Landscape*

GEOLOGY ILLUSTRATED

John S. Shelton, Drawings by Hal Shelton

Geologist/pilot Dr. John Shelton has been taking aerial photographs since 1939, and this is his master work. *Geology Illustrated* is a breathtaking view of our home planet: its rocks, minerals, volcanoes, glaciers, lakes, rivers, earthquake faults, mountain ranges, and other features. The nearly 400 photographs and crystal-clear text provide an enjoyable way to understand geology. Shelton addresses the book to "thoughtful people who enjoy the outdoors"—to hikers, rockhounds, pilots (and passengers), environmentalists, and many others. "An artist of aerial photography, Shelton uses some 400 of his finest photos to illuminate a discussion of the whole-earth system. Not a traditional textbook, but a fascinating exploration of problems posed by asking, 'How did that come about?' Worth buying for the photos and book design alone, but you'll probably find yourself becoming interested in geology regardless of your original intentions."

—*Whole Earth Catalog,* Summer 1971

1966, 434 pages, 382 illustrations, $14.50

PLANTS IN THE LANDSCAPE

by Philip L. Carpenter, Theodore D. Walker, and Frederick O. Lanphear, all of Purdue University

Plants in the Landscape is a unique introduction to the principles and practices of ornamental horticulture in landscape architecture. No other single volume covers such a broad range of topics. Here is a wealth of information on subjects ranging from the landscape practices of ancient civilizations to no-nonsense, dollars-and-cents descriptions of the present-day landscaping industry. The focus throughout is on landscape plants: how to select and maintain them for healthy growth in a wide range of settings.

Plants in the Landscape is a beautiful book. The text is closely integrated with more than 400 illustrations (about 300 photographs). Whether you are building a new home or remodeling an old one, whether you are a professional landscaper or a weekend gardener, you'll find *Plants in the Landscape* a valuable addition to your library.

1975, 481 pages, 534 illustrations, 32 tables, $18.00

NATURAL REGIONS OF THE UNITED STATES AND CANADA

Charles B. Hunt, The Johns Hopkins University

With its generous use of photographs, line drawings, and maps (558 illustrations in all), *Natural Regions of the United States and Canada* provides a fascinating environmental tour of North America. Charles B. Hunt, a widely experienced and respected scholar, describes and explains the physical environment of the North American continent, Puerto Rico, and Hawaii. Natural regions are discussed in terms of their topography, geology, climate, soil, water, plants, and the interrelationship among all these factors.

"Just open this book to an area you already know. Within pages, the landforms, soils, colors, plants, and importance of your watershed will intensify. An encyclopedia of physiography written in academic style. Lots of maps. Lots of geologic history. The best book to practice mapping North America into your brain."—*Whole Earth Epilog,* 1974

1974, 725 pages, 558 illustrations, 22 tables, $17.00

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

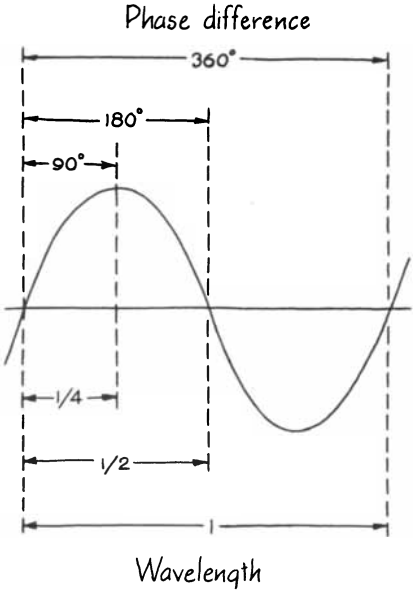
THE AMATEUR SCIENTIST

Studying polarized light with quarter-wave and half-wave plates of one's own making

by Jearl Walker

The light that enters your eyes is a wave pattern of oscillating electric and magnetic fields that are perpendicular to the direction of travel of the light. Most of the light by which you see is unpolarized, that is, the fields oscillate along all possible directions perpendicular to the direction of travel of the light. Some of the light, however, is linearly polarized and thus has electric fields that oscillate along a single axis perpendicular to the direction of travel. For example, light directly from the sun is unpolarized whereas the sunlight reflected from a road surface can be polarized (with the electric field oscillating primarily horizontally).

Probably the most convenient modern way to convert unpolarized light from the sun or some other light source is with a polarizing filter such as a Polaroid filter. If initially unpolarized light passes through such a filter, it becomes linearly polarized. In other words, the filter transmits light with electric fields that oscillate along a single axis perpendicular to the direction of travel of the light. This selective transmission can be helpful. If a motorist wants to eliminate the glare from a road surface, he can wear polarized sunglasses that pass light with electric fields that oscillate vertically, thereby stopping the glare from rays with electric fields that oscillate primarily horizontally.



Phase differences

Standard polarizing filters are made by stretching sheets of the plastic polyvinyl alcohol. Initially the long-chain molecules in the sheets are in random orientations and show no preference for one polarization or another. When the sheets are warmed and then stretched, however, the long molecules tend to align themselves in the direction of stretch. When unpolarized light passes through such a sheet of oriented molecules, the component of the light's electric field that lies along the molecules is absorbed by them. The electric field perpendicular to the length of the molecules is not absorbed. Hence when the light emerges, its electric field is along a single axis and is linearly polarized. The polarizing sheet is called dichroic because its absorption is different for different polarizations of the incident light. Two polarizing filters aligned in the same way will transmit light whereas two filters aligned with perpendicular senses of polarization will block it.

In other materials, such as certain crystals, both of the different senses of polarization for the incident light will be transmitted, but they will have different effective speeds. Such materials, which are called birefringent or doubly refracting, have two perpendicular axes; the speed of light is higher along one axis than along the other.

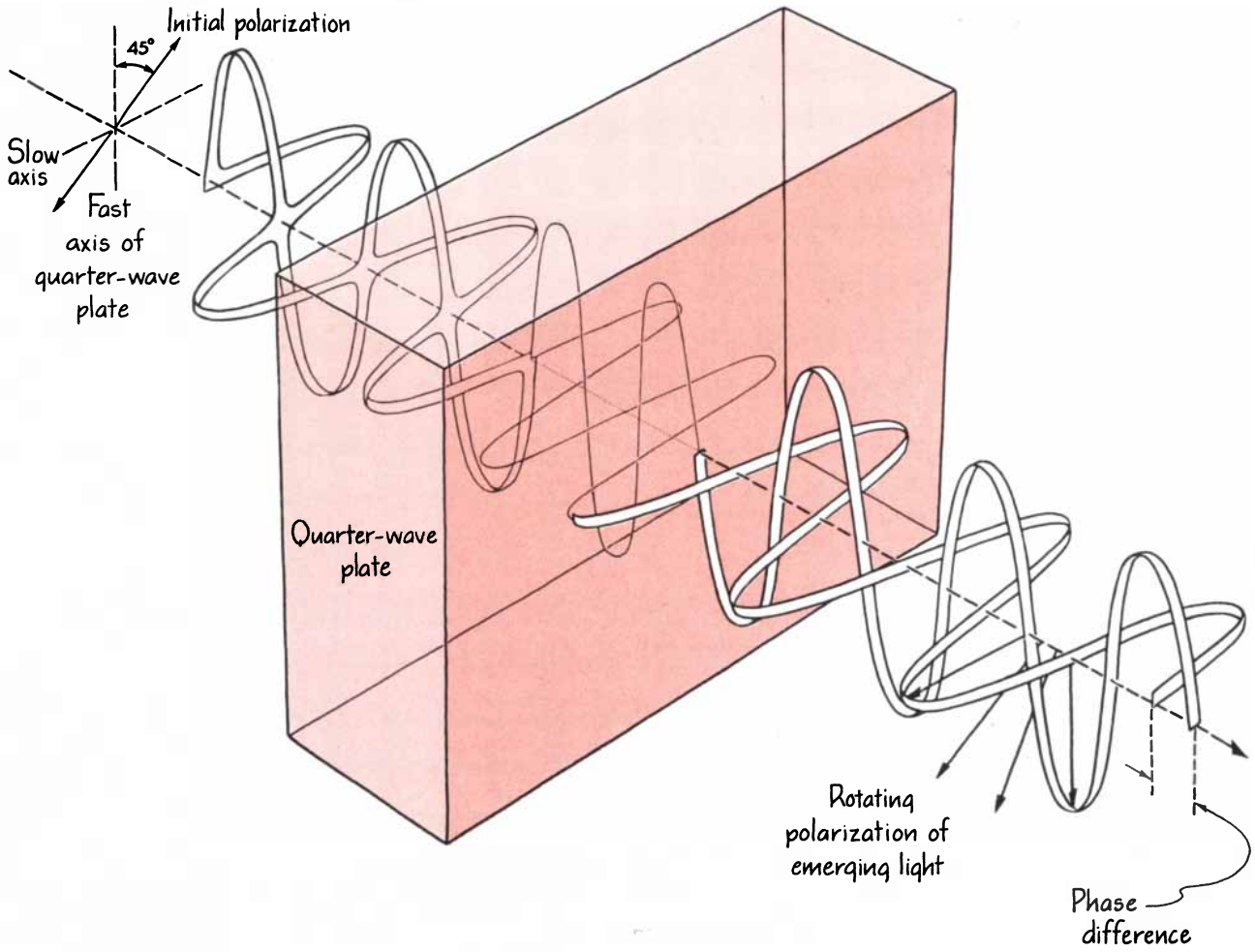
Suppose linearly polarized light is incident on a birefringent material in such a way that the sense of polarization of the light is at an angle to the fast and slow axes of the material. The polarization of the incident light has two components of interest, one component parallel to the fast axis and one parallel to the slow axis. Initially these two parts are in phase, that is, the oscillations in the electric field along the two axes are exactly in step. When the light passes through the birefringent material, however, the difference in the propagation speed of the two components causes one component to differ in phase from the other.

For example, suppose the width of the material is such that the light emerging along the slow axis lags behind the light emerging along the fast axis by a quarter of a wavelength, or 90 degrees (since a full wavelength along a wave is represented by 360 degrees). Then the phase difference between the two emerging components is said to be a quarter of a wavelength, or 90 degrees. In this case the material would be called either a quarter-wave plate or a 90-degree retarding plate, because it took two incident components of light that were in phase and left them 90 degrees out of phase. The emerging light is no longer linearly polarized because of this phase difference.

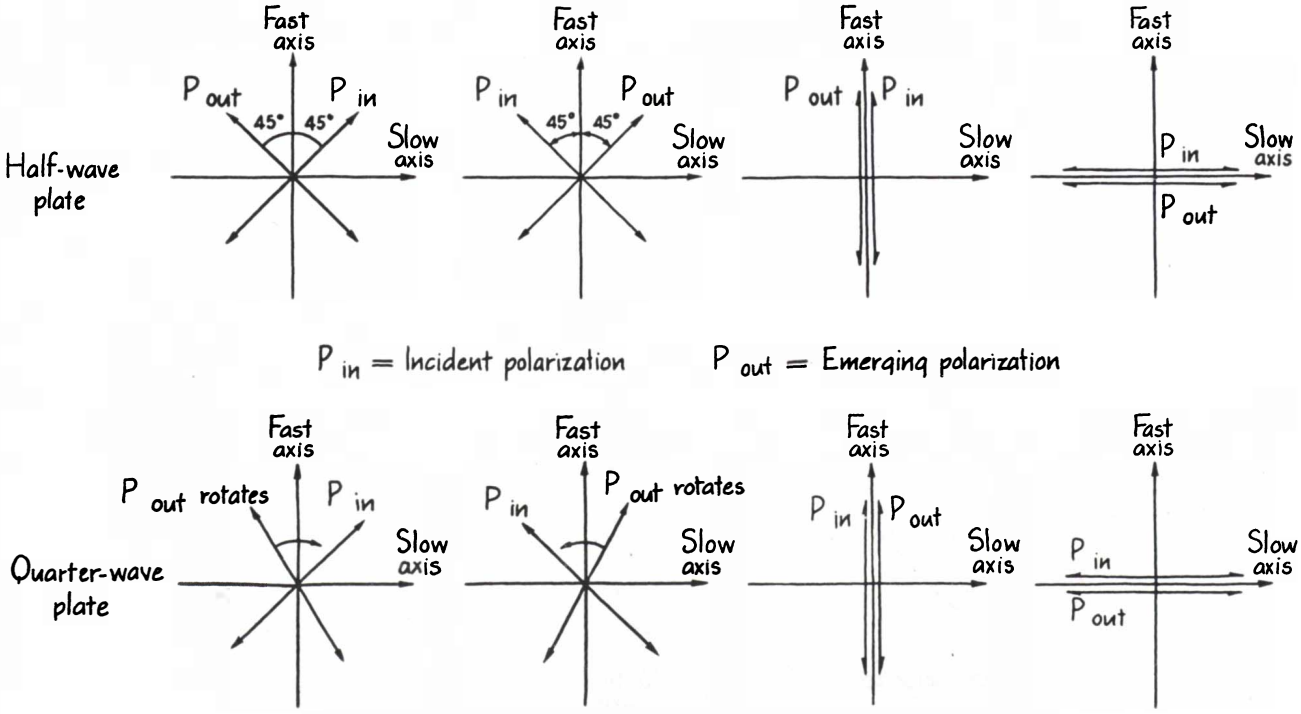
The combination of the two emerging components differing in phase by 90 degrees is elliptically polarized light in which the sense of polarization (the axis along which the net electric field oscillates) constantly rotates around the axis along which the light propagates. You can imagine that the tip of the arrow representing the sense of polarization maps out an ellipse around the propagation axis. If the two emerging components are of equal amplitude (if the maximum electric-field strength is the same along the fast axis and the slow one), the tip maps out a circle and the emerging light is circularly polarized. This result is obtained if the sense of polarization of the incident light lies at an angle of 45 degrees to the fast and slow axes of the birefringent material.

The material could also be of a thickness that would cause the emerging light components to be 180 degrees out of phase with each other, which is a shift of half a wavelength between the two. Then the material would be called a half-wave plate or a 180-degree retarding plate. The net electric field of the emerging components would oscillate along a single axis, and thus the light would again be linearly polarized. Now, however, the axis of polarization is not the same as it is for the incident light. The new axis of polarization will be flipped around either the slow axis or the fast one.

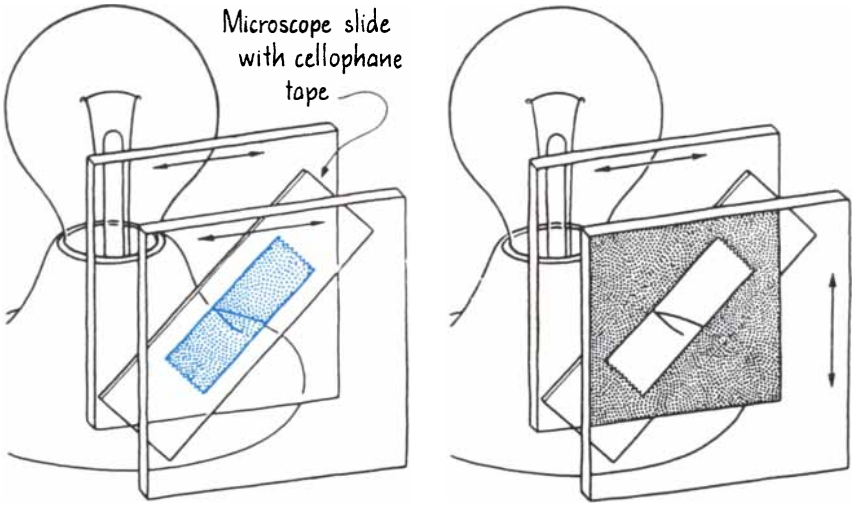
In his book *Waves* Frank S. Crawford, Jr., of the University of California at Berkeley describes several experiments you can do with common birefringent materials. He also provides information on how you can make your own quarter- and half-wave plates. My topic this month involves making these optical devices, investigating their effects on polarized light and using the plates to demonstrate the birefringence in the cornea of your eye. Next month I shall describe



Change of phase in light passing through a quarter-wave plate



Some polarizations of light leaving plates when linearly polarized light enters them



A half-wave plate with cellophane tape

how to employ the devices to make circularly polarized light.

Between two polarizing filters crossed to eliminate any transmitted light place a piece of clear plastic wrapping from a candy bar or a package of cigarettes. (The filters are available from the Edmund Scientific Company, 7778 Edscorp Building, Barrington, N.J. 08007, for $2.50 for a package of 20, or you can dismantle a pair of polarized sunglasses.) Then look at an incandescent lamp that has a clear glass bulb and a straight filament. With the plastic in place some light is transmitted through the system of filters because the plastic reorients the polarization of light incident on it and thereby provides the second polarizing filter with a component of light polarized in the same sense as the second filter. By rotating the plastic you can maximize the transmitted light.

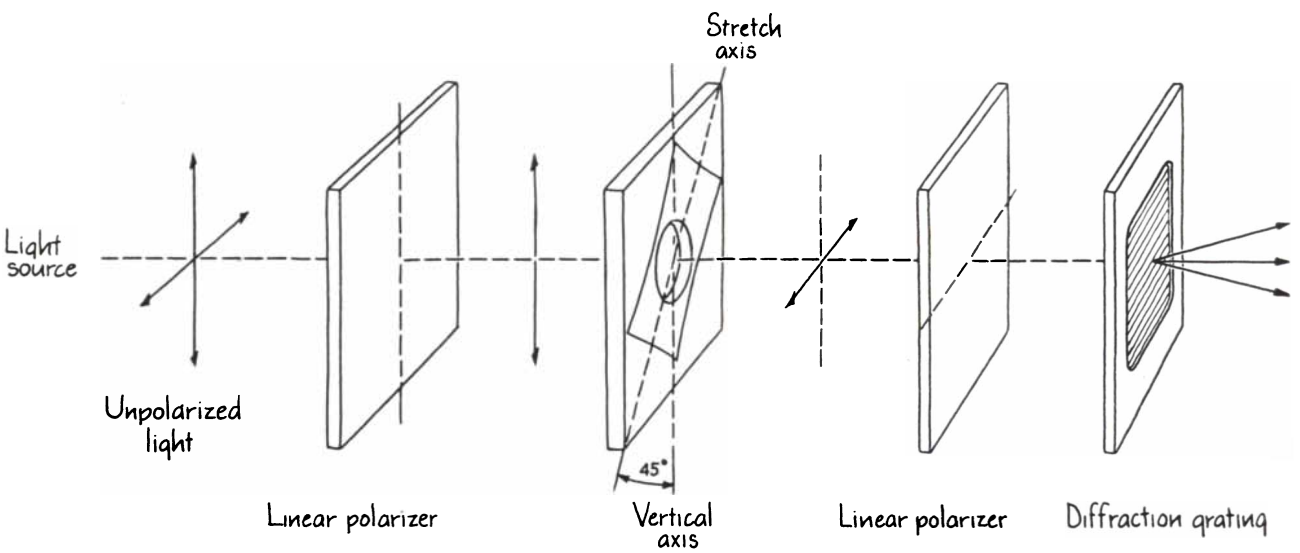
A length of food wrapping (available in the U.S. under such names as Handi-Wrap and Saran Wrap) initially does not act in the same way. By stretching the wrapping, however, you can orient its long molecules as the molecules are oriented in the making of polarizing filters; you can thereby instill fast and slow axes in the material. Between two crossed polarizing filters place a stretched length of food wrapping so that the stretch direction is at 45 degrees to the polarization senses of the filters. You will then get a lot of transmitted light through the filter system.

From several layers of stretched food wrapping you can make your own quarter- or half-wave plate. Over a hole in a sturdy piece of metal, wood or cardboard tape from 10 to 14 layers of food wrapping, with each layer stretched in the same direction. Tape to one side of the hole a polarizing filter with its sense of polarization at an angle of 45 degrees to the stretch direction of the wrapping. The filter provides linearly polarized light to the layers, with equal components along the stretch axis and perpendicular to it.

With the proper number of layers taped in place the wrapping acts as a half-wave plate for a particular wavelength in the incident white light. For that wavelength the two components of the emerging light differ in phase by 180 degrees, or half a wavelength. You can detect this phase shift by placing another polarizing filter in the path of the emerging light and testing for its sense of polarization.

For example, suppose you have taped down enough layers to make a half-wave plate at a wavelength of 500 nanometers, which is green light. Then the green in the incident light undergoes a 180-degree phase shift between the two components in the emerging light. The emerging components add to give a net sense of polarization that is linear and perpendicular to the initial polarization. If the second polarizing filter is oriented to pass this polarization, you will see mostly green light. If the second filter is oriented parallel to the polarization sense of the first filter, the green light is blocked by the second filter and you see instead the visible spectrum with green absent.

For any orientation of the second filter you see colors (perhaps dimly) for which the food wrapping is not a half-wave plate, so that the colors emerge elliptically polarized. You can "tune" the half-wave plate by adding or subtracting layers of food wrapping and then testing for the color transmitted when the second filter is placed perpendicular to the first one. If you halve the



Checking the wavelength passed by a half-wave plate

W. H. Freeman and Company
Publishers in Science

Modes of Speciation

Michael White

The well-known evolutionary biologist Michael White presents a unique synthesis of the processes that enable new species to originate. Unlike previous books on speciation, which deal exclusively with plants or concentrate on phyletic evolution, this book is the first *general* treatment of genetic and geographic models of speciation in both animals and plants. White assesses the mounting evidence for the critical role of chromosomal rearrangements and attacks the view that all speciation is allopatric. A *comprehensive* bibliography will be of interest to both advanced students and research workers.

January 1978, 464 pages, 31 illustrations, cloth $27.50

Evolution

Theodosius Dobzhansky, Francisco J. Ayala, G. Ledyard Stebbins, and James W. Valentine

Evolution is a contemporary synthesis of biological evolution jointly written by four distinguished authorities in the field. The authors have combined their expertise and insight to give you the most complete view of evolutionary biology available in a single edition. It includes presentations from biology, zoology, botany, anthropology, ecology, molecular biology, and philosophy.

Evolution focuses on the *processes* that spark biological evolution. It developes such topics as hereditary variation in natural population, fundamental processes of genetic change. In-depth coverage is also given to the fossil record, cellular organization, animal phylogenies, human evolution and philosophical problems in evolution.

1977, 572 pages, 141 illustrations, cloth $16.95

Psyclosis

The Circularity of Experience

Ralph Berger

Psyclosis: the Circularity of Experience describes how human experience is relative to the biological evolutionary process. "My aim in writing this book was to analyze scientific and humanistic modes of experience in such a way as to integrate them. The key to achieving that aim appeared to me to lie in a consideration of the evolutionary functions of the brain, which form the necessary conditions for all experience . . ."

— From the Preface

Taking a phenomenological approach, Berger examines several closely related concepts: the Second Law of Thermodynamics, organic evolution, ontogenesis, learning, reinforcement, homeostasis, perception, will, consciousness, dreaming, and the Collective Unconscious.

1977, 221 pages, cloth $11.00

Organismic Evolution

Verne Grant

Foreword by George Gaylord Simpson

Organismic Evolution is a lucidly written, broadly based treatment of the fundamentals of evolutionary biology. Author Verne Grant, a major contributor to the modern synthetic theory of evolution, concentrates on "principles rather than details, on fundamentals rather than current topics, and on well-chosen examples rather than catalogs of facts." The emphasis throughout the text is on evolution in organisms, chiefly plants and animals, rather than on molecular evolution or mathematical modeling.

This text is the outgrowth of an upper level undergraduate course. The chapters are brief in order to introduce a wide range of topics. *Organismic Evolution* is an excellent text for college and university courses in evolution. This book will also be of interest and use to professional workers interested in organismic evolution.

1977, 418 pages, 66 illustrations, cloth $15.95

Payment must accompany orders over $25.00

W. H. Freeman and Company
660 Market Street, San Francisco, CA 94104

The Ultimate Tax Shelter

by
TED NICHOLAS

Tax experts are now referring to a small, privately owned corporation as "The Ultimate Tax Shelter." This is especially true since the passage of the Tax Reform Act of 1976. This law makes most former tax shelters either obsolete, or of little advantage. Investments affected include real estate, oil and gas drilling, cattle feeding, movies, etc. These former tax shelters have lost their attractiveness. Aside from that, these tax shelters required a large investment. Only a small segment of the population could benefit from them.

I've written a book showing how you can form your own corporation. I've taken all the mystery out of it. Thousands of people have already used the system for incorporation described in the book. I'll describe how you may obtain it without risk and with a valuable free bonus.

A corporation can be formed by anyone at surprisingly low cost. And the government encourages people to incorporate, which is a little known fact. The government has recognized the important role of small business in our country. Through favorable legislation incorporating a small business, hobby, or sideline is perfectly legal and ethical. There are numerous tax laws favorable to corporate owners. Some of them are remarkable in this age of ever-increasing taxation. Everyone of us needs all the tax shelter we can get!

Here are just a few of the advantages of having my book on incorporating. You can limit your personal liability. All that is at stake is the money you have invested. This amount can be zero to a few hundred or even a few thousand dollars. Your home, furniture, car, savings, or other possessions are not at risk. You can raise capital and still keep control of your business. You can put aside up to 25% of your income tax free. If you desire, you may wish to set up a non-profit corporation or operate a corporation anonymously. You will save from $300 to $1,000 simply by using the handy tear-out forms included in the book. All the things you need: certificate of incorporation, minutes, by-laws, etc., including complete instructions.

There are still other advantages. Your own corporation enables you to more easily maintain continuity and facilitate transfer of ownership. Tax free fringe benefits can be arranged. You can set up your health and life insurance and other programs for you and your family wherein they are tax deductible. Another very important option available to you through incorporation is a medical reimbursement plan (MRP). Under an MRP, all medical, dental, pharmaceutical expenses for you and your family can become tax deductible to the corporation. An unincorporated person must exclude the first 3% of family's medical expenses from a personal tax return. For an individual earning $20,000 the first $600 are not deductible.

Retirement plans, and pension and profit-sharing arrangements can be set up for you with far greater benefits than those available to self-employed individuals.

A word of caution. Incorporating may not be for you right now. However, my book will help you decide whether or not a corporation is for you now or in the future. I review all the advantages and disadvantages in depth. This choice is yours after learning all the options. If you do decide to incorporate, it can be done by mail quickly and within 48 hours. You never have to leave the privacy of your home.

I'll also reveal to you some startling facts. Why lawyers often charge substantial fees for incorporating when often they prefer not to, and why two-thirds of the New York and American Stock Exchange companies incorporate in Delaware.

You may wonder how others have successfully used the book. Not only a small unincorporated business, but enjoyable hobbies, part time businesses, and even existing jobs have been set up as full fledged corporations. You don't have to have a big business going to benefit. In fact, not many people realize some very important facts. There are 30,000 new businesses formed in the U.S. each and every month. 98% of them are small businesses; often just one individual working from home.

To gain all the advantages of incorporating, it doesn't matter where you live, your age, race, or sex. All that counts is your ideas. If you are looking for some new ideas, I believe my book will stimulate you in that area. I do know many small businessmen, housewives, hobbyists, engineers, and lawyers who have acted on the suggestions in my book. A woman who was my former secretary is incorporated. She is now grossing over $30,000 working from her home by providing a secretarial service to me and other local businesses. She works her own hours and has all the corporate advantages.

I briefly mentioned that you can start with no capital whatsoever. I know it can be done, since I have formed 18 companies of my own, and I began each one of them with nothing. Beginning at age 22, I incorporated my first company which was a candy manufacturing concern. Without credit or experience, I raised $96,000. From that starting point grew a chain of 30 stores. I'm proud of the fact that at age 29 I was selected by a group of businessmen as one of the outstanding businessmen in the nation. As a result of this award, I received an invitation to personally meet with the President of the United States.

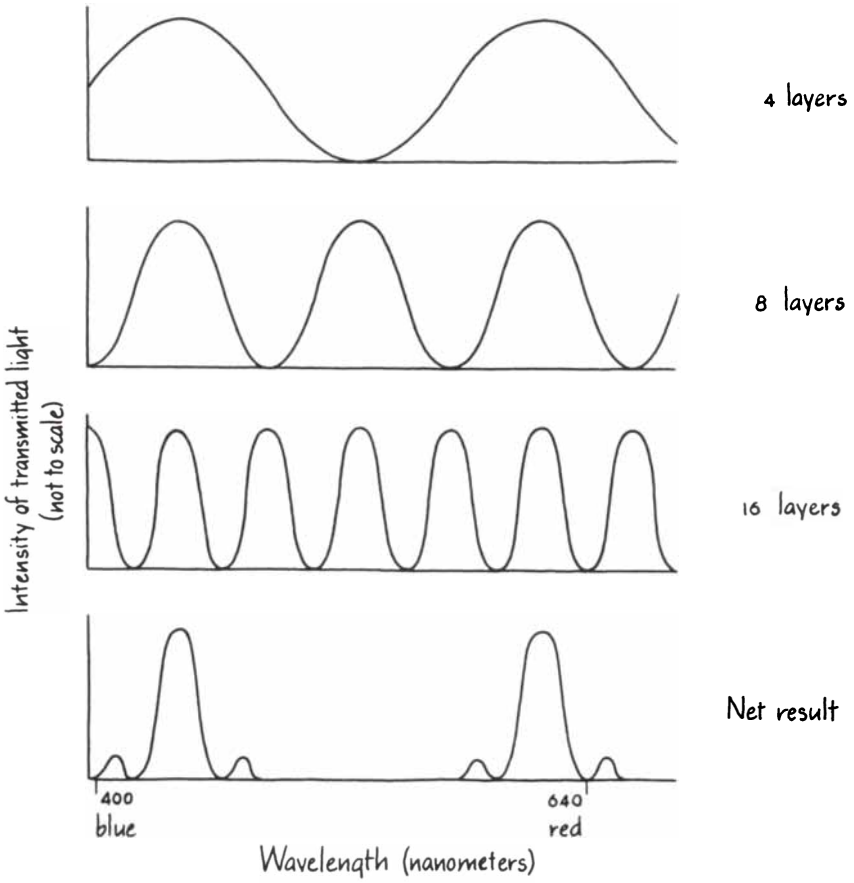
I wrote my book, **How To Form Your Own Corporation Without A Lawyer For Under $50**, because I felt that many more people than otherwise would could become the President of their own corporations. As it has turned out, a very high proportion of all the corporations formed in America each month, at the present time are using my book to incorporate.

Just picture yourself in the position of President of your own corporation. My book gives you all the information you need to make your decision. Let me help you make your business dreams come true.

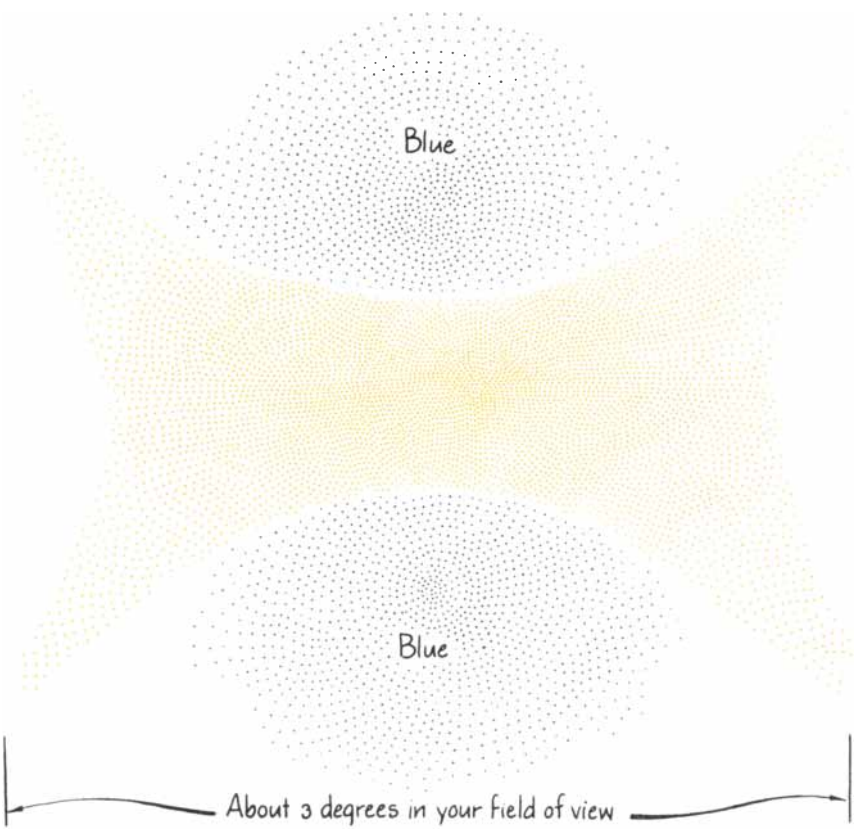
As a bonus for ordering my book now, I'll send you absolutely free a portfolio of valuable information. It's called "The Incom Plan" and normally sells for $9.95. It describes a unique plan that shows you how to convert most any job into your own corporation. You'll increase your take-home pay by up to 25% without an increase in salary or even changing jobs in many cases. If you are an employer, learn how to operate your business with independent contractors rather than employees. This means that you'll have no payroll records or withholding taxes to worry about. And you'll be complying with all I.R.S. guidelines. "The Incom Plan" includes forms, examples and sample letter agreements to make it possible.

I personally guarantee your purchase in the fairest way I know. If you feel my book is not all that I've described, return it undamaged within a full year and I'll promptly refund your money with no questions asked. If you should decide to return it, you may keep the bonus Incom Portfolio for your trouble.

To get your copy, write the words "Corporation Book and Bonus" on a plain sheet of paper, along with your name and address. Enclose a check for $14.95. There is no sales tax on your order and your purchase price is tax deductible. Mail your order to me, Ted Nicholas. c/o Enterprise Publishing Co., Dept. SA-7ZC , 1000 Oakfield Lane, Wilmington, Delaware 19810. C1016



Transmission ranges for cellophane-tape filters



Haidinger's brushes

number of layers, you will have a quarter-wave plate.

You can also make a half-wave plate out of a single layer of clear plastic tape (not "magic transparent tape," which is only semitransparent). Stick the layer on a piece of glass (for support) and then place it between two polarizing filters with their senses of polarization crossed. A particular range of colors will be transmitted; they are the colors for which the thickness of the tape is approximately a half-wave plate. My cellophane tape was approximately a half-wave plate for much of the visible spectrum except at the blue end. With crossed filters the transmitted light appeared to be white; with parallel filters the transmitted light was blue.

With a half-wave plate made out of either plastic wrapping or cellophane tape, examine the emerging light with a diffraction grating. (Such a grating, which is available from Edmund Scientific for about 50 cents, spreads the light out into a spectrum.) If the second polarizing filter is oriented to transmit the wavelength for which the device is a half-wave plate, you will find that color in the spectrum from the grating. Rotating the second filter to the perpendicular orientation blocks that color and eliminates it from the spectrum seen in the grating.

With several layers of transparent tape you can make a filter that passes light only in certain narrow ranges of wavelength. Make a stack of 16 layers of clear plastic tape on a piece of glass. To improve the transmission through the layers place a small drop of machine oil between each successive layer and between the first layer and the glass. Then orient the stack between two polarizing filters as you did with the single layer of tape.

Again examine the emerging light with the diffraction grating. The transmitted spectrum has several dark lines between colored images of the lamp's filament, indicating that the second polarizing filter is blocking several different wavelengths in the light emerging from the layers. As you rotate the second filter the dark lines are replaced by bright lines and vice versa. When the second filter is turned perpendicular to its first orientation, it passes the wavelengths it formerly blocked and blocks the ones it formerly passed.

For some wavelengths in the incident white light the stacked layers effectively constitute a half-wave plate, and their emerging polarizations are linear and flipped around either the slow axis or the fast one. For some of the other wavelengths the stacked layers are effectively a full-wave plate; their polarization is maintained. (A phase shift of any integral number of wavelengths merely causes the emerging components of the light—the ones along the slow and fast

axes—to come back in phase, so that they have the same polarization as the incident light does.) Hence one orientation of the second polarizing filter selects one of these results to transmit. Another orientation of the polarizing filter selects the other result. In either case the spectrum is incomplete; only certain wavelengths are allowed to pass.

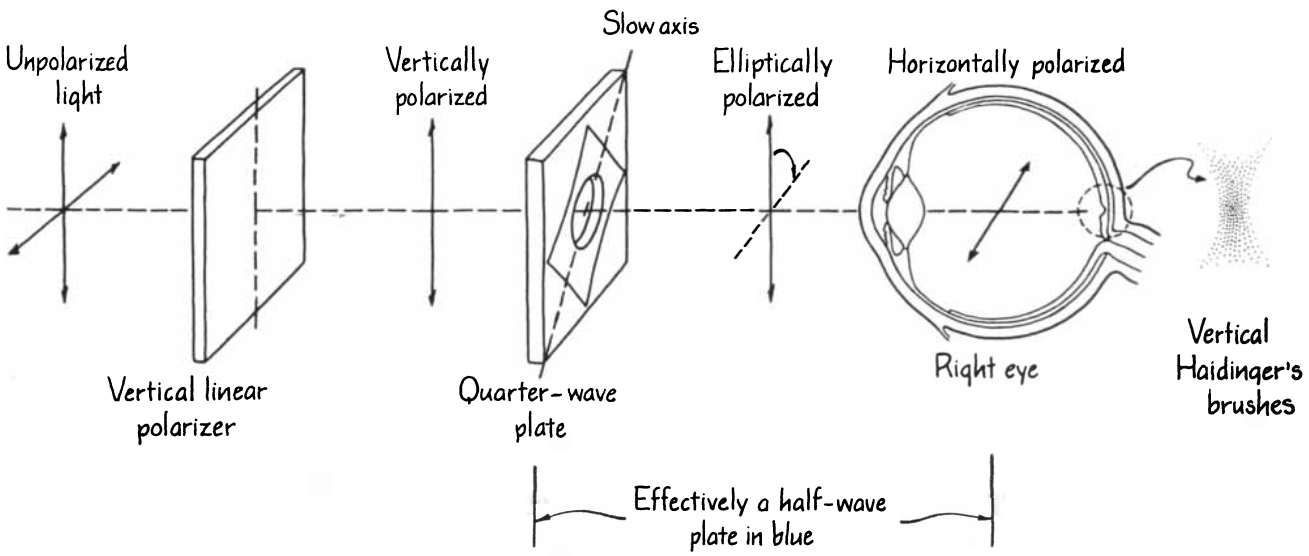
A single layer of tape gives a phase shift of about half a wavelength in the visible spectrum. To obtain a shift of a full wavelength for some colors you need a thicker stack of tape. Four layers of tape will work, but the dark lines are more apparent with 16 layers. To see this effect make two more stacks of cellophane tape, one with four layers and one with eight. For a particular orientation of the second polarizing filter the 16-layer stack passes light in fairly narrow ranges of wavelengths. The smaller stacks also pass light only in certain ranges of wavelengths, but the ranges are wider the smaller the stack is.

If white light passes through all three stacks, only light in certain narrower ranges of wavelengths survives. If more stacks were placed in the light beam, with each new stack twice the thickness of the preceding one, the surviving wavelength ranges would become increasingly narrow and the ranges would disappear until only two or so would be left in the visible spectrum. Then a final color filter could select one of them.

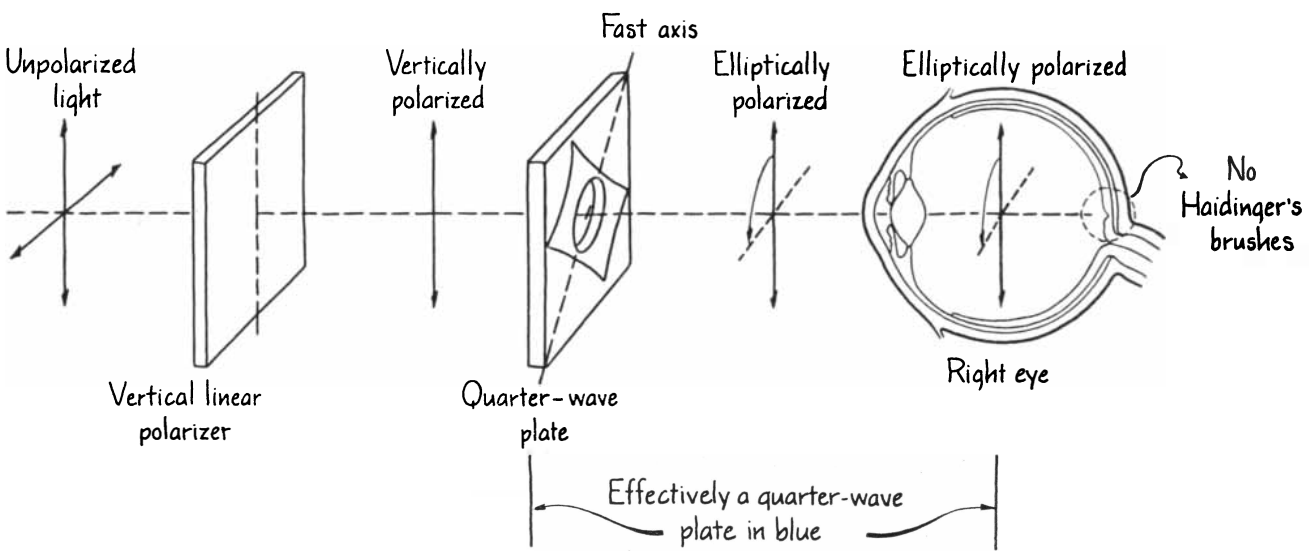
This type of band-pass filter was developed by the French astronomer B. F. Lyot in 1932 for photographing the sun. The bandwidth (the width of the wavelength range transmitted by the filter) was as small as .1 nanometer in the Lyot filters, which were made out of quartz. With the Lyot filter tuned to pass, say, one of the hydrogen-emission wavelengths and to block all other visible wavelengths the sun could be photographed in its hydrogen emission, thereby displaying the behavior of hydrogen on the sun.

Most people are able to detect polarized light with the unaided eye. Look through a linearly polarizing filter at a blue portion of the sky or at some other blue and generally featureless background. In the center of your field of vision will lie a small, faint, yellow hourglass figure subtending an angle of about three degrees. Small blue areas may also be seen at the sides of the hourglass. (They are not visible to everyone.)

The hourglass is called Haidinger's brushes, after the Austrian mineralogist Wilhelm Karl von Haidinger, who discovered the effect in 1844. The figure quickly fades unless the sense of polar-



How a quarter-wave plate and fibrils in the eye make a half-wave plate



The result with a quarter-wave plate and the eye

ization of the light incident on your eyes changes, so that you should rotate the polarizing filter slowly to keep the figure visible. The hourglass rotates in the same way.

If you cannot see the hourglass, you may be able to see it later in life. Several years ago I could see the figure in the polarized light from the sky without using any polarizing filter. Now I need the filter because the polarization of the sky is apparently not sufficient to produce the hourglass in my eye.

Although Haidinger's brushes have been discussed for a long time, their full explanation is still not available. They are probably caused by the selective absorption of light by the pigment in the macula lutea, the small region of the retina that is responsible for the greatest acuity in seeing. The pigment absorbs in the blue end of the visible spectrum, at wavelengths from 430 to 490 nanometers, but the absorption depends on the polarization sense of the incident light.

To schematize the absorption characteristics, one draws pigment elements laid out in radial lines from a center. The maximum absorption occurs along a diameter of such a pattern when the diameter is perpendicular to the sense of polarization of the light incident on the macula lutea. For example, suppose you hold the polarizing filter so that vertically polarized light enters the eye. Then the maximum absorption of the blue in the incident light takes place along a horizontal diameter. If you rotate the filter, maximum absorption occurs along other diameters, always perpendicular to whatever sense of polarization you happen to give your eye. As we have seen, material that absorbs different amounts of light along different axes is defined as being dichroic. The pigment in the macula lutea of your eye is said to be radially dichroic because of this dependence on the incident polarization.

For the sake of simplicity consider the incident light to be vertically polarized. The elimination of the blue along the horizontal by the pigment means that the retina reports to the brain the complementary color yellow as being along the horizontal. This yellow is the long axis of Haidinger's brushes. The blue clouds on each side of the hourglass have not been completely explained but are probably psychological in that the brain itself is responsible for the impression that blue is present. If you cannot find the hourglass while looking at the sky with a filter, you might try viewing a blue wall illuminated by sunlight. In any case the hourglass requires linearly polarized blue light incident on the eye in order to be seen.

Haidinger's brushes do more than indicate that part of your eye is sensitive to the polarization sense of light; they can also be employed to show that an-

other part of your eye is birefringent. To demonstrate this effect you need a linear polarizing filter and a quarter-wave plate tuned to the green or yellow. Orient the linear filter to yield vertically polarized light. Insert between it and your eye a quarter-wave plate with its slow axis slanted upward and outward. (If you are using your right eye, the slow axis should run from the bottom left to the top right.) By inserting the quarter-wave plate in this way you make the Haidinger's brushes from the linear filter suddenly change orientation by 90 degrees, from the horizontal to the vertical. If the plate's slow axis is instead slanted upward and inward (for the right eye running from the bottom right to the top left), the hourglass disappears. Using a half-wave plate (in the green or yellow) instead of a quarter-wave one gives opposite results.

Because the brushes may be difficult to see with your homemade quarter-wave plate, you might want to replace it and the linear polarizer with a circular polarizer that the Polaroid Corporation will provide to readers of this department without charge (Polaroid Corporation, Polarizer Technical Products, 20 Ames Street, Cambridge, Mass. 02139). Ask for the card "Polaroid Circular Polarizers for Contrast Improvement," to which the polarizer is taped. Several of the uses to which a circular polarizer can be put will be discussed in this department next month.

The question of why the hourglass either disappears or is flipped by 90 degrees is not settled. Apparently the effect of the inserted plate is to add to the birefringence already present in the cornea to give a net birefringence of either a quarter- or half-wave plate for blue light. If the inserted plate and the cornea are effectively a quarter-wave plate for blue light, the brush disappears. If they are effectively a half-wave plate for blue light, the brush is flipped perpendicular to the orientation the linear filter alone would give.

The cornea contains fibrils of collagen that slant mainly upward and outward. They are birefringent, with the slow axis along the length of the fiber and the fast axis upward and inward. The phase shift between the fast and slow axes is normally not sufficient to be noticed, being only about a twelfth of a wavelength in green light. If the quarter- or half-wave plate is used properly, however, the total shift of the plate and the collagen can be sufficient to change the sense of polarization of the incident blue light that is necessary to create Haidinger's brushes.

Consider the quarter-wave plate as being oriented with its slow axis upward and outward, as the slow axis of the collagen fibrils is. In the green-yellow the quarter-wave plate produces a phase shift of about a quarter of a wavelength and the collagen produces an additional shift of about a twelfth, giving a total shift of approximately a third of a wavelength in that color region. Other colors are shifted proportionately to their wavelength, the blue being shifted less than the red.

Since Haidinger's brushes depend on blue light, the shift in the blue is the interesting one. The total shift in the blue is about two-fifths of a blue wavelength. Thus the quarter-wave plate and your cornea act in combination as approximately a half-wave plate in the blue. Vertically polarized light transmitted through a half-wave plate oriented with its slow axis perpendicular to the polarization changes the polarization by 90 degrees (from vertical to horizontal in this case). Correspondingly, the brushes flip from the horizontal when no plate is in place to the vertical when the plate is inserted, provided that the plate is positioned properly.

If the quarter-wave plate has its slow axis upward and inward, the brushes are eliminated. In the yellow-green region the plate phase shifts the light by about a quarter of a wavelength in one direction and the collagen shifts it by about a twelfth of a wavelength in the opposite direction, giving a net phase shift of about a sixth of a wavelength in the green-yellow. In the blue this is a net phase shift of about a fifth of a wavelength. In other words, the combination of the collagen and the quarter-wave plate oriented in this way in effect creates a quarter-wave plate in the blue. The combination produces elliptically polarized light to illuminate the macula lutea, but light of that type does not produce Haidinger's brushes. With this sense of orientation for the inserted quarter-wave plate the brushes therefore disappear.

The effect of a half-wave plate inserted between a vertically oriented polarizing filter and your eye is similar. With the slow axis upward and outward the combined phase shift produced by the plate and your cornea is about three-quarters of a wavelength in the blue. The plate and cornea hence act as a quarter-wave plate in the blue, producing elliptically polarized light for the retina. No brushes appear. With the slow axis upward and inward the net phase shift is about half a wavelength in the blue. This time the plate and the cornea together act as a half-wave plate in the blue; they flip the sense of the incident linearly polarized light by 90 degrees and thereby rotate the Haidinger's brushes from the horizontal to the vertical.

If you can successfully see the effects with the preceding arrangements, you might want to try other relative orientations of the linearly polarized filter, the quarter- or half-wave plate and the collagen fibrils. The principles are the same, but predicting the outcome in each case would be a good test of your understanding of how quarter- and half-wave plates work.

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION (required by 39 U.S.C. 3685). *1.* Title of publication: Scientific American. *1A.* Publication number: 484780. *2.* Date of filing: September 28, 1977. *3.* Frequency of issue: monthly. *3A.* Number of issues published annually: 12. *3B.* Annual subscription price: U.S., possessions and Canada, 1 year $18; all other countries, 1 year $22. *4.* Location of known office of publication: 415 Madison Avenue, New York, N.Y. 10017. *5.* Location of the headquarters or general business offices of the publishers: 415 Madison Avenue, New York, N.Y. 10017. *6.* Names and complete addresses of publisher, editor and managing editor: Publisher, Gerard Piel, 415 Madison Avenue, New York, N.Y. 10017. Editor, Dennis Flanagan, 415 Madison Avenue, New York, N.Y. 10017. Managing Editor, none. *7.* Owner: Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Janet R. Becker, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Buckley & Co. (for Elinor Gottlieb Mannucci), c/o Fiduciary Trust Co. of New York, P.O. Box 3199, Church Street Station, New York, N.Y. 10008; Cudd & Co. for (a) Helen R. Snellenburg, Morris Duane and Raymond J. Armstrong, Trustees u/a 2/17/66 F.B.O. Helen R. Snellenburg and (b) Fred Grossman, Successor Trustee u/a 10/15/51 F.B.O. Robert L. Rosenwald, The Chase Manhattan Bank, 1 Chase Manhattan Plaza, New York, N.Y. 10015; Martin M. Davidson, 11 River Lane, Westport, Conn. 06880; Richard E. Deutsch, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Dorothy D. Eweson, c/o Brunt & Co., Box 6626, Journal Square, Jersey City, N.J. 07306; Bayard Ewing, c/o Industrial National Bank, Trust Department, 100 Westminster Street, Providence, R.I. 02903; Dennis Flanagan, c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Sidney R. Garfield, M.D., Kaiser Foundation Hospitals, 1924 Broadway, Oakland, Calif. 94612; William T. Golden, Room 4201, 40 Wall Street, New York, N.Y. 10005; E. F. Hutton & Company, Inc., 1 Battery Park Plaza, New York, N.Y. 10004; Nathan W. Levin, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Frasier W. McCann, Room 726, 10 Rockefeller Plaza, New York, N.Y. 10020; Donald H. Miller, Jr., c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Eleanor Jackson Piel, custodian Eleanor Jackson Piel, Jr., Unif. Gift Min. Act NY, 320 Central Park West, New York, N.Y. 10025; Gerard Piel, c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Jonathan Bird Piel, 320 Central Park West, New York, N.Y. 10025; Barbara K. Rosenwald, P.O. Box 3235, Norfolk, Va. 23514; Julius Rosenwald II, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Lessing J. Rosenwald, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Rowe & Co. (for Rhode Island Charities Trust), c/o Industrial National Bank of Rhode Island, Trust Department, 100 Westminster Street, Providence, R.I. 02903; Stanley Schaefer, c/o W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104; I. M. Scott, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Joan R. Scott, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; John Hay Whitney, Room 4600, 110 West 51 Street, New York, N.Y. 10020. *8.* Known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages or other securities: none. *9.* For completion by nonprofit organizations authorized to mail at special rates (Section 132.122, PSM). The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes: Not applicable to this organization. *10.* Extent and nature of circulation: *A.* Total number of copies printed (net press run): average number of copies each issue during preceding 12 months, 762,053; actual number of copies of single issue published nearest to filing date, 764,199. *B.* Paid circulation: *1.* Sales through dealers and carriers, street vendors and counter sales: average number of copies each issue during preceding 12 months, 135,349; actual number of copies of single issue published nearest to filing date, 131,100. *2.* Mail subscriptions: average number of copies each issue during preceding 12 months, 532,377; actual number of copies of single issue published nearest to filing date, 534,213. *C.* Total paid circulation (sum of *10B1* and *10B2*): average number of copies each issue during preceding 12 months, 667,726; actual number of copies of single issue published nearest to filing date, 665,313. *D.* Free distribution by mail, carrier or other means, samples, complimentary and other free copies: average number of copies each issue during preceding 12 months, 20,660; actual number of copies of single issue published nearest to filing date, 22,925. *E.* Total distribution (sum of *C* and *D*): average number of copies each issue during preceding 12 months, 688,386; actual number of copies of single issue published nearest to filing date, 688,238. *F.* Copies not distributed: *1.* Office use, leftover, unaccounted, spoiled after printing: average number of copies each issue during preceding 12 months, 14,233; actual number of copies of single issue published nearest to filing date, 10,139. *2.* Returns from news agents: average number of copies each issue during preceding 12 months, 59,434; actual number of copies of single issue published nearest to filing date, 65,822. *G.* Total (sum of *E, F1* and *2*—should equal net press run shown in *A*): average number of copies each issue during preceding 12 months, 762,053; actual number of copies of single issue published nearest to filing date, 764,199. *11.* I certify that the statements made by me above are correct and complete. (Signed) Donald H. Miller, Jr., Vice-President and General Manager. *12.* For completion by publishers mailing at the regular rates (Section 132.121, Postal Service Manual). 39 U.S.C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates." In accordance with the provisions of this statute, I hereby request permission to mail the publication named in item *1* at the phased postage rates presently authorized by 39 U.S.C. 3626. (Signed) Donald H. Miller, Jr., Vice-President and General Manager.

INTRODUCING A FIENDISHLY SEDUCTIVE FOUR-DOOR. THE 1978 DODGE DIPLOMAT.

PRICED AS SHOWN: $5195.*

1978 DODGE DIPLOMAT FOUR-DOOR	
Base price	$ 5147
Six-cylinder engine	Standard
Power front disc brakes	Standard
4-speed manual transmission†	Standard
Deluxe wheel covers	Standard
Power steering	Standard
Transverse torsion-bar suspension	Standard
Vinyl roof	Standard
White sidewall radial tires	$ 48
Total	$ 5195*

The 1978 Dodge Diplomat four-door offers big-car value and comfort... in a manageable new size. And for one low price, you get a lot of beautiful car.

As you can see, most of the equipment that helps make a car a real value... is standard on the Diplomat. But some things have to be experienced to fully appreciate this car's value. Like Diplomat's outstanding driveability. And the overall performance of Diplomat's lively Super Six engine.† This power plant with a four-speed manual transmission† gives Diplomat EPA mileage estimates of 25 mpg highway and 17 mpg city. (Your mileage may vary according to your car's condition, its equipment, and your driving habits. California mileage lower.)

See your Dodge Dealer today and find out why more and more people are buying or leasing Dodge Diplomats.

*Sticker price as shown, not including taxes and destination charges. California price higher.

†In California, 225-1 six-cylinder engine standard and automatic transmission required at extra cost.

Dodge

A PRODUCT OF CHRYSLER CORPORATION

The 2 ounce vacation.

(With the taste of the Virgin Islands.)

Enjoy a taste of the Virgin Islands tonight. Mr. Boston Virgin Islands Rum. Perfectly clear, dry and every bit as sunny as the Islands that distill it. Makes you wish you were there.

Let Mr. Boston make your party.

Virgin Islands Rum, 80 Proof, Imported and bottled by Mr. Boston Distiller, Boston, Mass., Owensboro, Ky., Albany, Ga. © 1977.

ANNUAL INDEX

The following index lists all the authors and articles that appeared in SCIENTIFIC AMERICAN *during 1977. Also indexed are "Mathematical Games" and "The Amateur Scientist."*

AUTHORS

ARTICLES

MATHEMATICAL GAMES

THE AMATEUR SCIENTIST

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

A CONSECUTIVE-DIGIT PRIME. Joseph S. Madachy in *Journal of Recreational Mathematics,* Vol. 4, No. 2, page 100; April, 1971.

CONSECUTIVE-DIGIT PRIMES—AGAIN. Joseph S. Madachy in *Journal of Recreational Mathematics,* Vol. 5, No. 4, pages 253–254; October, 1972.

CONSECUTIVE-DIGIT PRIMES (ROUND 3). Raphael Finkelstein and Judy Leybourne in *Journal of Recreational Mathematics,* Vol. 6, No. 3, pages 204–206; Summer, 1973.

THE JUPITER EFFECT. John R. Gribbin and Stephen H. Plagemann. Macmillan, 1974.

PRODUCT TECHNOLOGY AND THE CONSUMER

PATTERNS OF ENERGY CONSUMPTION IN THE UNITED STATES. Stanford Research Institute, prepared for the Office of Science and Technology, Executive Office of the President. Government Printing Office, January, 1972.

PRODUCT PERFORMANCE IN AN AFFLUENT SOCIETY. Lewis M. Branscomb in *Product Quality, Performance, and Cost: A Report and Recommendations Based on a Symposium and Workshops Arranged by the National Academy of Engineering at the Seventh Annual Meeting, April 29 and 30, 1971.* National Academy of Engineering, 1972.

SERVICE-LIFE EXPECTANCY OF HOUSEHOLD APPLIANCES: NEW ESTIMATES FROM THE USDA. Marilyn Doss Ruffin and Katherine S. Tippett in *Home Economics Research Journal,* Vol. 3, No. 3, pages 159–170; March, 1975.

THE NUCLEOTIDE SEQUENCE OF A VIRAL DNA

A RAPID METHOD FOR DETERMINING SEQUENCES IN DNA BY PRIMED SYNTHESIS WITH DNA POLYMERASE. Frederick Sanger and Alan R. Coulson in *Journal of Molecular Biology,* Vol. 94, No. 3, pages 441–448; May 25, 1975.

OVERLAPPING GENES IN BACTERIOPHAGE ϕX174. B. G. Barrell, G. M. Air and C. A. Hutchison III in *Nature,* Vol. 264, No. 5581, pages 34–41; November 4, 1976.

NUCLEOTIDE SEQUENCE OF BACTERIOPHAGE ΦX174 DNA. Frederick Sanger, G. M. Air, B. G. Barrell, N. L. Brown, Alan R. Coulson, John C. Fiddes, C. A. Hutchison III, P. M. Slocombe and M. Smith in *Nature,* Vol. 265, No. 5595, pages 687–695; February 24, 1977.

THE MOTION OF THE GROUND IN EARTHQUAKES

ELEMENTARY SEISMOLOGY. Charles F. Richter. W. H. Freeman and Company, 1958.

MECHANISM OF THE CHILEAN EARTHQUAKES OF MAY 21 AND 22, 1960. George Plafker and J. C. Savage in *Geological Society of America Bulletin,* Vol. 81, No. 4, pages 1001–1030; April, 1970.

EARTHQUAKE SHAKING AND DAMAGE TO BUILDINGS. Robert A. Page, John A. Blume and William B. Joyner in *Science,* Vol. 189, No. 4203, pages 601–608; August 22, 1975.

STRONG-MOTION RECORDINGS OF THE CALIFORNIA EARTHQUAKE OF APRIL 18, 1906. David M. Boore in *Bulletin of the Seismological Society of America,* Vol. 67, No. 3, pages 561–577; June, 1977.

THE EPIDEMIOLOGY OF INFLUENZA

THE GREAT EPIDEMIC. Adolph A. Hoehling. Little, Brown & Co., 1961.

THE INFLUENZA VIRUSES AND INFLUENZA. Edited by Edwin D. Kilbourne. Academic Press, Inc., 1975.

EPIDEMIC AND PEACE, 1918. Alfred W. Crosby, Jr. Greenwood Press, 1976.

THE RETINEX THEORY OF COLOR VISION

COLOR VISION AND THE NATURAL IMAGE: PART I. Edwin H. Land in *Proceedings of the National Academy of Sciences,* Vol. 45, No. 1, pages 115–129; January, 1959.

COLOR VISION AND THE NATURAL IMAGE: PART II. Edwin H. Land in *Proceedings of the National Academy of Sciences,* Vol. 45, No. 4, pages 636–644; April, 1959.

INTERACTION OF THE LONG-WAVE CONES AND THE RODS TO PRODUCE COLOR SENSATIONS. John J. McCann and Jeanne L. Benton in *Journal of the Optical Society of America,* Vol. 59, No. 1, pages 103–107; January, 1969.

LIGHTNESS AND RETINEX THEORY. Edwin H. Land and John J. McCann in *Journal of the Optical Society of America,* Vol. 61, No. 1, pages 1–11; January, 1971.

DISCLINATIONS

CRYSTAL DISLOCATIONS: ELEMENTARY CONCEPTS AND DEFINITIONS. F. C. Frank in *The Philosophical Magazine: A Journal of Theoretical, Experimental and Applied Physics,* Vol. 42, No. 331, pages 809–819; August, 1951.

THEORY OF CRYSTAL DISLOCATIONS. Frank R. N. Nabarro. Clarendon Press, 1967.

THE GEOMETRY OF DISCLINATIONS IN CRYSTALS. William F. Harris in *Surface and Defect Properties of Solids, Vol. 3: A Review of the Recent Literature Published up to April, 1973.* The Chemical Society, London, 1974.

NEGATIVE WEDGE DISCLINATIONS OF ROTATION 2π RADIANS AND TOPOLOGICAL CHANGES OF MEMBRANOUS SYSTEMS. William F. Harris in *The Philosophical Magazine: A Journal of Theoretical, Experimental and Applied Physics,* Vol. 32, No. 1, pages 37–41; July, 1975.

WEAVER ANTS

THE INSECT SOCIETIES. Edward O. Wilson. The Belknap Press of Harvard University Press, 1971.

WEAVER ANTS: SOCIAL ESTABLISHMENT AND MAINTENANCE OF TERRITORY. Berthold K. Hölldobler and Edward O. Wilson in *Science,* Vol. 195, No. 4281, pages 900–902; March 4, 1977.

COLONY-SPECIFIC TERRITORIAL PHEROMONE IN THE AFRICAN WEAVER ANT *OECOPHYLLA LONGINODA* (LATREILLE). Berthold K. Hölldobler and Edward O. Wilson in *Proceedings of the National Academy of Sciences,* Vol. 74, No. 5, pages 2072–2075; May, 1977.

A CELTIC FARMSTEAD IN SOUTHERN BRITAIN

EXCAVATIONS AT LITTLE WOODBURY, WILTSHIRE, PART I: THE SETTLEMENT REVEALED BY EXCAVATION. Gerhardt Bersu in *Proceedings of the Prehistoric Society,* Vol. 6, pages 30–111; 1940.

THE IRON AGE SETTLEMENT OF GUSSAGE ALL SAINTS. Geoffrey Wainwright and Mansel G. Spratling in *Antiquity,* Vol. 47, pages 109–130; 1973.

IRON AGE COMMUNITIES IN BRITAIN: AN ACCOUNT OF ENGLAND, SCOTLAND AND WALES FROM THE SEVENTH CENTURY B.C. UNTIL THE ROMAN CONQUEST. Barry Cunliffe. Routledge & Kegan Paul, 1974.

THE IRON AGE IN LOWLAND BRITAIN. Derek W. Harding. Routledge & Kegan Paul, 1974.

THE AMATEUR SCIENTIST

POLARIZED LIGHT. William A. Shurcliffe and Stanley S. Ballard. D. Van Nostrand Company, 1964.

WAVES. Frank S. Crawford. McGraw-Hill Book Company, 1968.

HAIDINGER'S BRUSHES AND PREDOMINANT ORIENTATION OF COLLAGEN IN CORNEAL STROMA. C. C. D. Shute in *Nature,* Vol. 250, No. 5462, pages 163–164; July 12, 1974.

The attractive VideoBrain console is only 15" X 12" X 14". Its durable black surface is mar-proof and water resistant. A mere 10 lbs., it consumes a miserly 10 watts of power. Unique 360 degree joysticks (4 max.) with trigger button allow complex arcade-type games.

LECTRO-MEDIA LTD.
introduces the

FIRST FAMILY COMPUTER

It's VideoBrain, the family computer that's educational, entertaining and helpful at home.

Introducing VideoBrain , the first family computer. A truly multi-purpose instrument designed to be used in the home or office for education, household management and entertainment. "Speaking" to the Brain is as simple as using a typewriter (and it will even teach you to type). VideoBrain responds to you by answering on your own TV screen in 16 colors, musical tones, and flashing messages. The Brain plays arcade-type games so sophisticated that your tennis player follows through on his point-winning stroke, a ballboy runs on the court to give you the ball, while the crowd in the stands follows the ball back and forth.

POWERFUL TEACHING AIDE

VideoBrain makes learning exciting. Subjects such as math, music composition, and computer programming are made easy with automatic adjustment of difficulty levels. Challenging word games improve your spelling and vocabulary. VideoBrain will teach your child the invaluable skill of typing. The Brain knows that it is necessary for you to converse with it or any other computer.

VIDEOBRAIN - THE HOME MANAGER

Imagine an appliance that balances your checkbook, turns the coffeepot on and the air conditioner off at preset times. Or stores messages, answers your phone, and tells you date, day of week, and exact time at a glance to your TV. All this will be possible with the Brain and expansions soon to be available.

VIDEOBRAIN - GROWING WITH YOU

Because the Brain is a true computer, it is designed to expand and never ever becomes obsolete. Future cartridges will add to your library collection of education, entertainment, and utility programs. With an expansion module, the Brain will store your business or professional records on dual audio cassettes, edit and then type reports or bills on your IBM Selectric. It will answer your telephone and send written material over it, or pay bills.

COMPUTER CHESS AND CHECKERS

Dr. Arthur Samuels, Stanford University's professor emeritus has taught your VideoBrain to play varying levels of chess and checkers against you, or keep score while you play against another Brain owner over the phone.

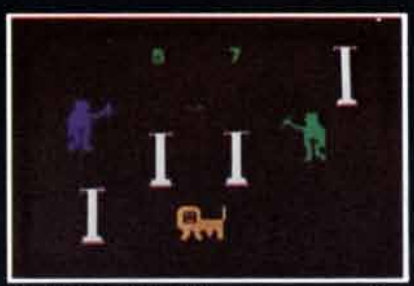
DESIGNED BY COMPUTER EXPERTS

The brains behind the Brain are two experts in the computer field: Dr. David Chung, the inventor of Fairchild F-8 microprocessor (the Brain's cortex), and Dr. Albert Yu, Intel's integrated circuit genius. Together they have reduced the Brain's incredibly complex circuitry to just 8 IC's on a single circuit board. This guarantees years of maintenance free performance. Two of these chips are proprietary to the VideoBrain Computer Co., and produce the 16-color TV graphics. So fine, in fact, (160 X 400 matrix) that it would take 320K bits of core to refresh the screen if conventional techniques were used.

COMES READY TO USE

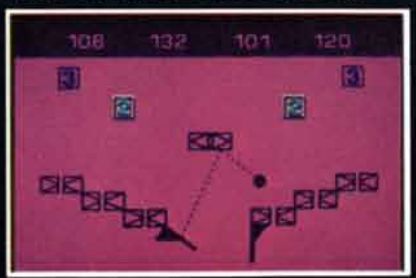
Your VideoBrain comes complete with alphanumeric keyboard. Unlike other products announced with overly optimistic plans for expansion, your VideoBrain comes ready to teach complex word skills and allows keyboard interplay. VideoBrain will allow data storage on audio cassettes; telephone and peripheral communications through an optional communications module. A complete home control module is also under design, both to be announced early in 1978 with realistic delivery dates.

II FREE WITH VIDEOBRAIN



Gladiators–384 different games on three battlefields–a coloseum in ancient Rome, a modern day football field, or intergalactic space. Play against opponent or computer.

Wordwise 1–Word Building Exercise and Game. VideoBrain selects an assortment of ten letters and challenges you to make words out of them. Build vocabulary.



VideoBrain Cartridge Library

OPTIONAL LIBRARIES--$19.95 ea.

Music Teacher 1–Teaches sheet music reading and writing for any keyboard instrument. Turn your computer into a musical instrument with a four octave range.

Wordwise 2–Touch Typing and Cypher, an exciting wordgame. The VideoBrain makes typing fun. VideoBrain checks accuracy and calculates speed. Cypher scrambles any title or phrase entered. Ideal party game.

Blackjack–A real Nevada-style game. 1 or 2 players against the house or the clock.

Pinball–Colors flash, buzzers sound, bumpers and gutters. 1 to 4 players.

FUTURE CARTRIDGES

Math Tutor 1–Helps teach basic math. Parent or teacher inputs.

Checkers and Chess–Developed by Dr. Arthur Samuels, Stanford's father of computer board games.

Cash Management–Complete package for home finances.

Finance Package–Make business decisions with full color graphics.

Also Planned: *Bridge Trainer, Stock & Real Estate Analysis.*

Call Toll-Free

800-523-2906

PA. RES. Call: (215)-925-9090

LECTRO-MEDIA LTD.
P.O. BOX 1770 sa
PHILA., PA. 19105
Philadelphia's Computer Store
Showrooms: **NewMarket**
America's Oldest Market
Front & Pine

YES, Send me the amazing VideoBrain home computer for $495. Don't forget my 2 free carts. Addt'l cartridges $19.95. I am paying by check or money order, send me a free optional cartridge of my choice as my cash discount.

Shipping: Unit $10., Each cart.$1. PA Res: add 6% sales tax. AIR: CA,WA,OR,HI,AD,Houston, Dallas: add $2. Canada: add $3. No CODs. St. address Only. 15 day trial. If not fully satisfied with your VideoBrain Home Computer, return it in 15 days for full refund. Unit pending FCC approval. Limited quantities available.

Who do you call to find the home of your dreams in less than 60 seconds?

Finding the home of your dreams used to require determination, stamina and lots of luck.

Used to.

Now, Realtors across the country have found a breakthrough way to make homebuying a whole lot easier.

They call *PRC Realtronics*—a powerful computerized real estate service developed for Realtors by PRC.

How to visit 9,000 homes for sale. By computer.

With a PRC Realtronics terminal in his office, a Realtor today can help a homeseeker scan through as many as 9,000 local property listings in seconds.

To help each buyer find the best possible buy.

And help everyone—buyer, seller, Realtor—save a staggering amount of effort and time.

An extraordinary solution? Perhaps.

But with the professional skills of 5,000 technical specialists worldwide, designing practical answers to large-scale problems is what PRC is all about.

The largest diversified professional services company in the world.

In fact, for every problem you can think of, chances are there's a PRC

In cities like Houston, homebuyers today can ask local Realtors* for the "perfect" home. Size, price, location, built-ins, special extras—up to 250 specific features. Scanning thousands of listings for a home to match the request, a powerful computer system developed for Realtors by PRC can pinpoint it in seconds.

*Realtor is a registered trademark of the Natl. Assn. of Realtors

team of specialists already working on answers.

Planning the optimum assembly sequence for a major new auto plant.

Monitoring Space Shuttle facilities at Cape Kennedy.

Helping Chicago turn solid waste into electricity.

Installing the world's most comprehensive computerized cost measurement system for private industry.

The list of PRC projects goes on. And on.

Planning, engineering, architecture. Management consulting. Information services. We're here to help.

The point is, there's no problem too technical, too complex or too big for Planning Research Corporation to handle.

The question is, what could PRC do for you? Planning Research Corporation. 1850 K Street, N.W., Washington, D.C. 20006. Call (202) 293-4700.

Call

prc

Planning Research Corporation

Dewar's "White Label" to all, and to all a good Scotch.

Dewar's never varies.

BLENDED SCOTCH WHISKY • 86 8 PROOF • ©SCHENLEY IMPORTS CO., N. Y., N. Y.

GIFT WRAPPED AT NO EXTRA COST.