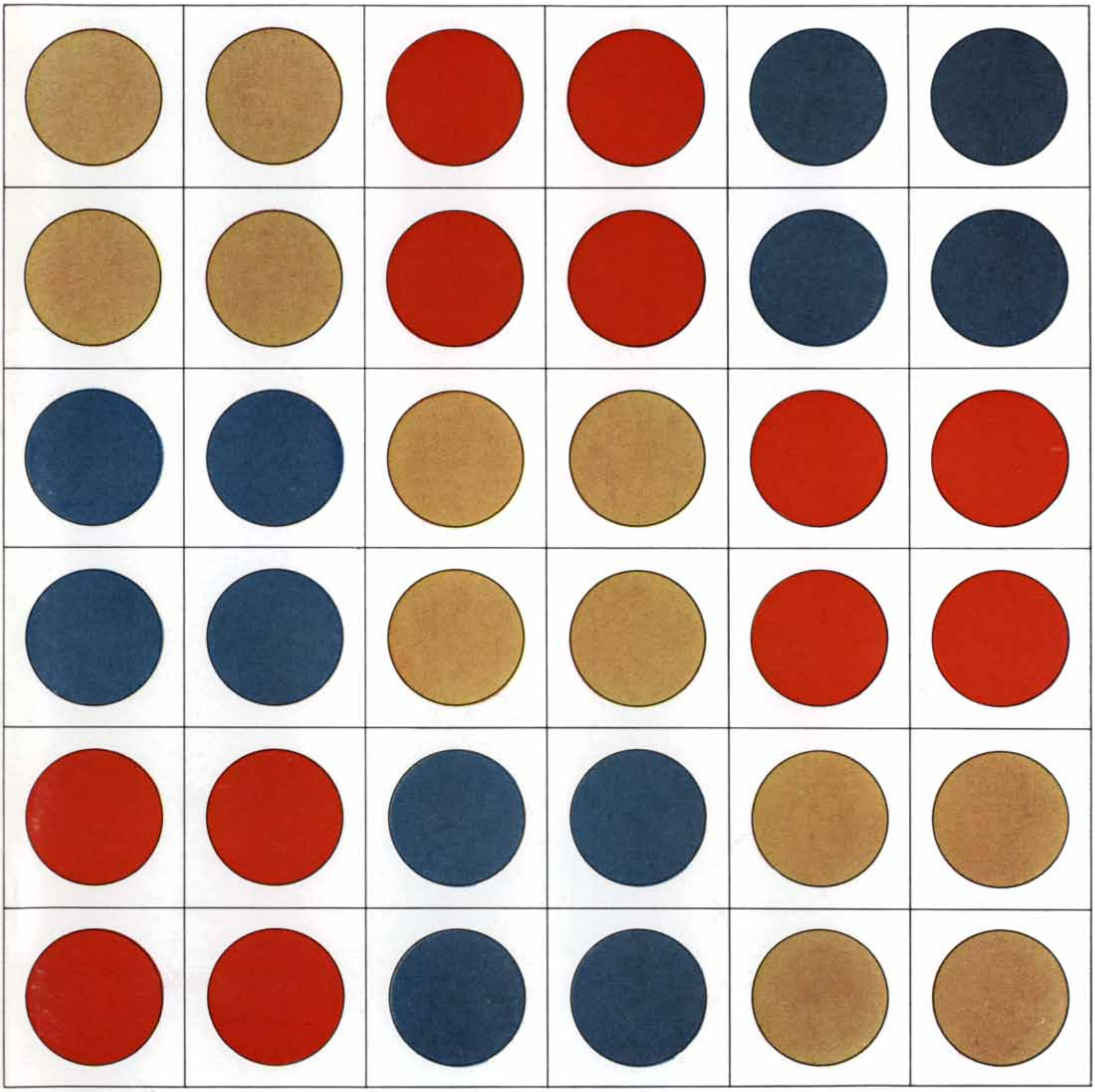
SCIENTIFIC AMERICAN



INDUCTIVE REASONING

ONE DOLLAR

November 1969

It's all the port new breed of

you'll need for this cargo ships.

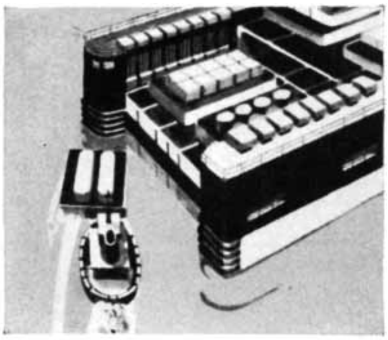
General Dynamics is building a fleet of new barge-carrying freighters for the Lykes Bros. Steamship Co., Inc. They may turn out to be the most important development in cargo ships since the last sail was furled. They can turn just about any waterfront area into a port.

Each ship has a capacity of 38 barges. The barges, each carrying over 480 tons of cargo, are towed to the ship and lifted by a stern elevator to one of three decks. They are then stowed automatically, and the ship speeds to its next destination at over 20 knots.

When it arrives, no docking is necessary. Barges are off-loaded the same way. Thus, 18,000 tons of cargo can be discharged in 13 hours, leaving the mother ship free to sail. A conventional freighter would be tied up for 195 hours to unload the same cargo. Cargo handling is cut from six operations to two.

Approximately three city blocks long, they're the biggest common cargo carriers ever built for the United States merchant fleet. We are building three of these Lykes-designed ships for delivery into international service in 1971.

Building advanced ships is not new to General Dynamics. Many of the most modern ships in the U.S. Navy have slid down our ways,



including the newest class of ammunition ships and fleet oilers.

It's just one example of what technology can accomplish when it's handed a problem.

At General Dynamics, we put technology to work solving problems from the bottom of the sea to outer space...and a good bit in between.

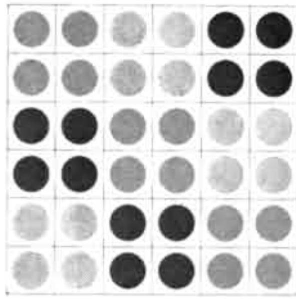
GENERAL DYNAMICS

Established 1845 SCIENTIFIC AMERICAN November 1969 Volume 221 Number 5

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1969 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $10 PER YEAR.



THE COVER

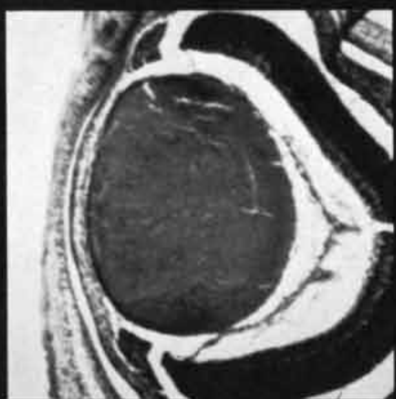
The design on the cover is one that might typically be used in a new inductive-reasoning game called Patterns (see "Mathematical Games," page 140). A player known as the Designer creates a pattern on a six-by-six grid. The other players try to discover what it is by asking about a few squares, framing a hypothesis and testing it. The Designer tries to give his pattern enough symmetry or other form of order so that at least some of the players are able to discover it. The pattern on the cover has bilateral, or mirror, symmetry about the diagonal axis from top right to bottom left.

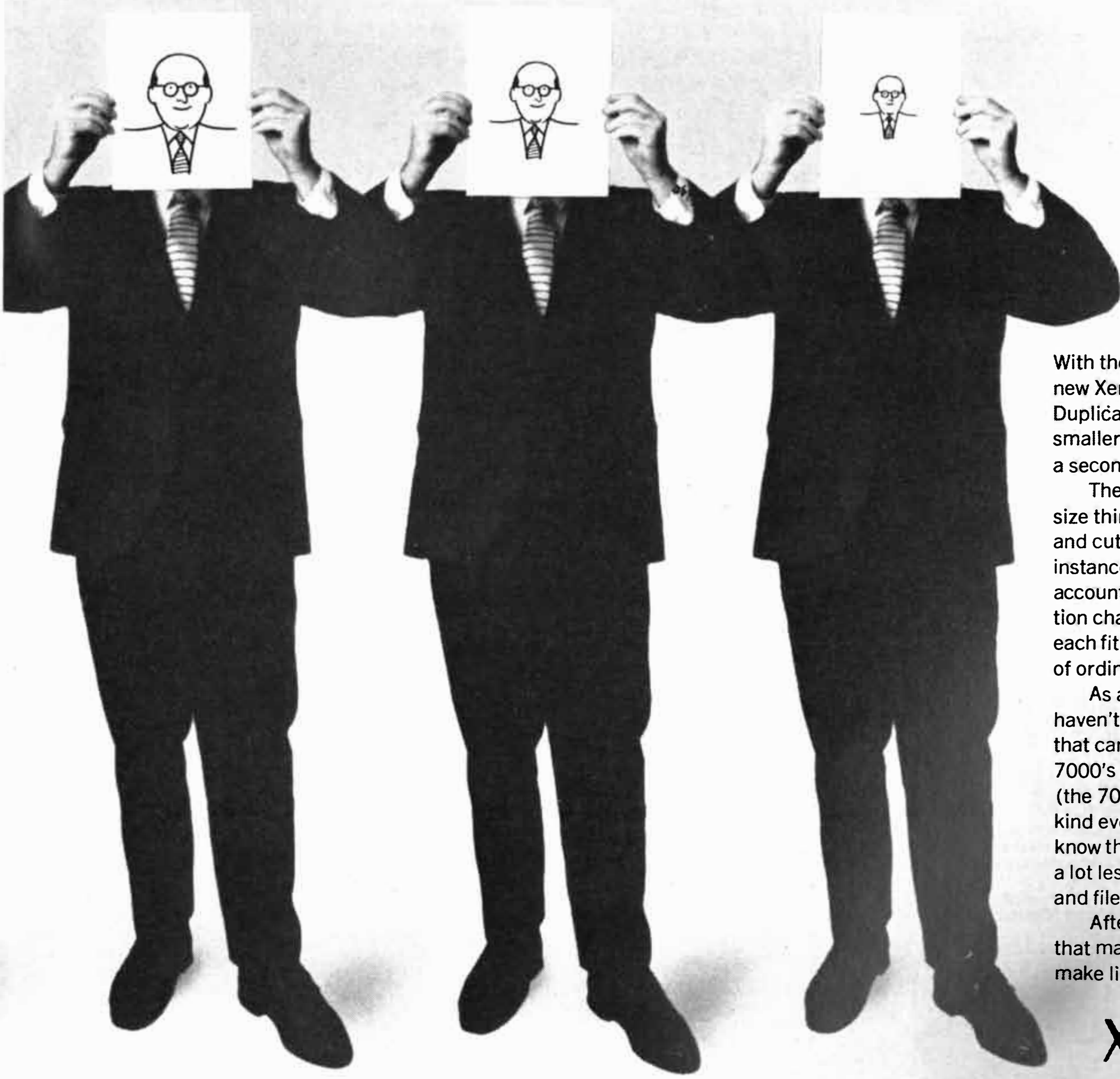
THE ILLUSTRATIONS

Cover illustration by Alan D. Iselin

Page	Source
24–27	Bunji Tagawa
28	Edward Lowenstein, Massachusetts General Hospital
29	Bunji Tagawa
31–36	Allen Beechel
37	Energy Conversion Devices, Inc.
38–40	Allen Beechel
41	Allen Beechel (*top*), Energy Conversion Devices, Inc. (*bottom left*)
43	Lee Boltin and Museum of the American Indian
44–51	Eric Mose
52	José M. Cruxent, Venezuelan Institute for Scientific Investigations
70	Sol Mednick
72–74	Dan Todd
75	Victor E. Ragosine
76–77	Dan Todd
78	Dan Todd (*top*), Victor E. Ragosine (*bottom*)
79–80	Victor E. Ragosine
87–94	Jerome Kuhl
102	Thomas Prentiss
104–105	Thomas Prentiss and Miklos Pinther
106–118	Thomas Prentiss
120	Cecil E. Hall, Massachusetts Institute of Technology
122–124	Enid Kotschnig
126–127	Bunji Tagawa
128–129	Bunji Tagawa (*left*), Crawford H. Greenewalt (*right*)
130	Crawford H. Greenewalt (*top*), Allen Beechel (*middle*), Bunji Tagawa (*bottom*)
131	Crawford H. Greenewalt (*top*), Allen Beechel (*middle*), Bunji Tagawa (*bottom*)
132	Bunji Tagawa
133	Bunji Tagawa (*left*), Allen Beechel (*right*)
134	Crawford H. Greenewalt (*top*), Allen Beechel (*middle*), Bunji Tagawa (*bottom*)
136	Crawford H. Greenewalt (*top, third from top*), Allen Beechel (*second from top, bottom left*), Bunji Tagawa (*bottom right*)
138	Bunji Tagawa (*top*), Allen Beechel (*middle, bottom*)
140–146	Alan D. Iselin
151–155	Roger Hayward


CANON BOOSTER
Canon





LETTERS

Sirs:

In your issue of January, 1969, you published a very interesting article by C. R. Taylor on the physiology of elands and oryxes and the ways by which they can survive extremes of heat and dryness in Africa. We learned many interesting things from the experiments made by Dr. Taylor.

However, we disagree with some of his final words. He arrives at the conclusion that there are "seemingly insurmountable problems" in "tapping the potential of antelope meat." In his opinion antelopes could not compete with cattle under "domestication, fencing and concentrated feeding" and it is not practical to breed them in deserts.

Apparently Dr. Taylor forgot other possibilities of ranching the eland successfully. One is to keep them in savannas, where cattle do not fare well, at least during the dry season. This is being tried in several places. Another possibility would be to keep them mixed with cattle, chiefly in weedy pastures where they may browse (although they also graze), perhaps eating plants not used by cattle. Pearsall, Darling, Matthews, Worthington, Talbot, Stewart, Dasmann and other researchers found that several different species of animals can exploit the African savannas better than only one species such as cattle (R. F. Dasmann in *African Game Ranching,* pages 34–39; 1964).

Experiments to domesticate the eland are under way in several parts of Africa, in the U.S.S.R., in the U.S. and now also in Brazil. We received our first pair of these animals last January, thanks to the help of Mr. John Perry and other friends. Here a quarter of our territory is covered by savannas. The eland, which Dr. Taylor proved to be so well adapted to warm and tough conditions, may perhaps fare better than cattle in such an environment.

PAULO NOGUEIRA NETO, SC.D.

Assistant Professor
Department of Zoology
University of São Paulo
São Paulo, Brazil

Sirs:

Roger Revelle's lucid article introducing your issue on the ocean [SCIENTIFIC AMERICAN, September] states of the Hawaiian Islands that Captain Cook "was as astonished as we are today that the Spanish Manila galleons had never found them."

This statement, while it is effectively true, is not strictly so. The account of Lord Anson's voyage, published in London in 1748, contains a copy of a captured chart showing a group of islands in the latitude of Hawaii, although 10 degrees too far east. The original of this chart, which is said to be in Spain, noted to the effect that the islands were discovered and described by Juan Gaetano in 1555. There can be no reasonable doubt that this was Hawaii, since no islands in that latitude exist to the east, and the longitude error in those prechronometer days is not surprising. The Hawaiian Islands were evidently not seen by Spaniards (except by accident) because of the ocean wind systems. Ships bound from Manila for Acapulco could make their easting only by reaching far to the north, avoiding the foul northeast trades and running before the fair prevailing westerlies.

F. E. ROMBERG

Austin, Tex.

Scientific American, November, 1969; Vol. 221, No. 5. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and treasurer.

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

Advertising correspondence should be addressed to Allan Wittman, Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 20 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York. N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 357B.

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

NAME

NEW ADDRESS

OLD ADDRESS

HISTORY

After 60 years Allen-Bradley appoints Schweber Electronics — their 1st distributor for electronic components.

Seated, Seymour Schweber, President of Schweber Electronics (left) and Arloe Paul, President of Allen-Bradley.

A kit containing brochures, specifications and price list of Allen-Bradley Electronic Components is now available. Write to Schweber Electronics, Westbury, New York, 11590.

Call any of these locations for immediate delivery of A-B components:

WESTBURY, NEW YORK: 516/334-7474 ☐ WALTHAM, MASS: 617/891-8484 ☐ ROCKVILLE, MD: 301/427-4977 ☐ CLEVELAND, OHIO: 216/333-7020
HOLLYWOOD, FLA: 305/927-0511 ☐ HUNTSVILLE, ALA: 205/539-2756 ☐ TORONTO, CANADA: 416/925-2471

SCHWEBER ELECTRONICS

Rolls-Royce unlimited.

△ **El Fantastico**
When it left Manaus, Brazil, it was just an SRN6 hovercraft. Two thousand miles later it emerged as El Fantastico, the only powered craft of any type to travel the whole length of the Orinoco to the sea. En route it conquered some of the most dangerous rapids in South America.

The journey was sponsored by Britain's Geographical Magazine. Its purpose was exploration of the tributaries of the Orinoco, deep in the South American jungles.

The SRN6 incidentally proved the enormous value of hovercraft to future exploration. A hovercraft can cross virtually any surface, including dry land, swamps and rapids. It rides above the surface on a cushion of air created by enormous horizontal fans. It moves very briskly, too. El Fantastico averaged a respectable 37 knots on the Rio Negro. Its Rolls-Royce Gnome engine performed faultlessly throughout.

The hovercraft's relative indifference to hostile terrain or water suggests all sorts of uses. Canada is operating the SRN6 for passenger service, coast guard duties and arctic exploration. Its Rolls-Royce engine gives a top speed of 72 miles per hour.

The world's largest hovercraft, the SRN4, is in regular service as a ferry across the English Channel. It has four Rolls-Royce engines, carries 254 people and 30 cars across at 72 miles per hour.

Fire when ready ▷
The latest thing in armor is plastic. It's rustproof, clankless, and it keeps you alive even if someone three paces away shoots you with both barrels of a 12-gauge shotgun, a Colt .45 or a 9mm Luger. Both the material and the armor are made by Rolls-Royce (Composite Materials) Limited.

◁ **407.5 miles per hour in 1931** This beautiful aircraft is the Supermarine S.6. It set the world speed record in 1931 at an incredible 407 miles per hour. It won the coveted Schneider Trophy for Britain in 1929 and 1931. Its engine, a Rolls-Royce Type R, was a minor miracle of forced-draft engineering.

In 1929, Rolls-Royce had only seven months in which to develop an engine. Working incessantly, they transformed their 825-hp V-12 Buzzard into a supercharged giant of 1,900 hp, which powered the S.6 to victory at 328.63 mph.

In 1931 they crammed five years' work into nine months, evoked a heroic 2,350 hp from the R. With this version the S.6 won the race at 340.8 mph and retired the trophy, and later the same year set the world speed record. The S.6 now lives in retirement at the Science Museum in London.

△ **Badger with mole speeds pipelaying** Rolls-Royce diesel engines power a remarkable new machine, called the Badger, that digs pipe trenches ten times faster than conventional equipment, at half the cost, and with almost none of the usual mess.

The Badger has a narrow vertical blade with a mechanical mole at its bottom edge. The blade passes through the ground like a plow, making only a narrow slit on the surface. But underground the mole burrows a tunnel up to twenty-four inches in diameter, and pulls the pipe in after it.

Designed by Hudswell Yates in England, the Badger is mounted on an International Harvester crawler chassis. On their first U.S. job, Badger machines installed over 50,000 feet of plastic water pipe in a single day.

△ **TriStar engine breathes through a tunnel** This great sinuous nostril is the air intake for one of the Lockheed TriStar's three Rolls-Royce RB.211 engines. One of many advantages of this configuration is easier ground servicing. TriStar customers include Eastern, TWA, Delta, Northeast, Air Canada and Air Jamaica; 181 TriStars are on order.

ROLLS ROYCE

Rolls-Royce Limited
Derby, England
Rolls-Royce
Aero Engines, Inc.
551 Fifth Avenue
New York, N.Y. 10017

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

NOVEMBER, 1919: "The foremost item of astronomical interest this month is the report which comes from England of the successful outcome of the observations made during the eclipse of last May to determine if rays of light passing close to the sun are deflected from their course. The observed facts are easy enough to understand. At the time of the eclipse the sun stood in a region of the heavens which is rich in bright stars. When the sun's disk was obscured by the moon, these stars became visible and could be photographed—the duration of totality being long enough to permit several successive exposures. On the plates so obtained the apparent positions of the stars could be accurately measured and compared with those on other plates of the same regions taken when the stars were visible at night, remote from the sun. The comparison showed beyond question that those stars whose images fell near the occulted sun were apparently shifted in position, in comparison with the remoter stars in the field, which served as standards of reference—the shift being in all cases away from the sun's center. The astronomers who were responsible for the investigation—Prof. Eddington of Cambridge and Dr. Crommelin of Greenwich Observatory—are thoroughly familiar with work of this sort and have undoubtedly taken every precaution to ensure accuracy. What this discovery means is obviously that rays of light which pass near the sun are deflected toward it—so that looking back along the deflected ray toward the star the latter appears to be shifted away from the sun. The most remarkable feature of the situation is that this deflection, both in direction and amount, had been predicted theoretically by Einstein, a German physicist, who has long been recognized by students of mathematical physics to be worthy of the high place which will doubtless now be his by acclamation."

"When recently interviewed at Amsterdam, Anthony Fokker, the Dutchman who figured so prominently in German wartime aviation, stated that in 1916 the German army authorities asked him to make a cheap airplane capable of flying about four hours, to be steered by wireless and to carry a huge bomb. It was intended to send these machines aloft in groups, to be controlled by one flier. Fokker said he prepared the plans, but the German war office decided to make the machines in government factories, with the result that they bungled along for months. Then, in the summer of 1918, they gave a huge order for wireless-controlled airplanes to M. Fokker, and he was just ready to manufacture them in wholesale quantities when the armistice was signed."

SCIENTIFIC AMERICAN

NOVEMBER, 1869: "As we approach the 17th of November, the day appointed for the final opening of the Canal of Suez, the interest felt in Europe and America in this vast enterprise increases with every new report of its advance toward completion. A few days more, and the two seas—the Sea of Corals, or the Mediterranean, and the Sea of Pearls, or the Red Sea—will be joined by a water route of 26 feet in depth and 328 feet in width, except at El Guisr, Serapeum and Chalouf, where the canal measures only 196 feet. The greater part of the expense of the works, conducted with as much patience as courage, has been borne by Egypt, while France will carry off the triumph and England may in time derive the greatest profit."

"The report of the Secretary of the Regents of the Smithsonian Institution, Professor Henry, states that at the last session of the board it was resolved to memorialize Congress, asking that the usual appropriation of $4,000 for the maintenance of the National Museum might be increased to $10,000, and also that $25,000 might be appropriated toward fitting up the large room in the second story of the building for the better exhibition of the Government collections. The request was refused and only the usual appropriation was made. In view of the fact that $4,000 was the sum appropriated when the museum was under the charge of the Patent Office, that since its removal to the Institution its size has been trebled, that the currency is greatly depreciated, and that the amount expended since the fire of 1865 is over $140,000, the greater part of which was for the accommodation of the National

Museum, it is hard to see why the moderate request of the Regents should not have been granted."

"There is a great deal said nowadays under the captivating title of 'Social Science,' but much of what is said and written warrants a doubt of even the existence of such a science. Still more does it warrant the doubt that those who attempt the discussion of social topics have, even admitting the existence of such a science, ever mastered the first rudiments of it. The wordy and weak discussions which have filled up the time of the 'Social Science Conventions' have not availed to fix public attention upon social evils more strongly than before they were uttered. The few suggestions made for reform and the correction of acknowledged existing evils have been of the most impracticable kind and showed most glaringly superficiality of thought in those who offered them. If there be not now, it is high time there ought to be such a thing as social science. Until some prophet arises capable of grappling with this subject from a physical and biological, as well as a political and legal, point of view, and beginning down upon hardpan shows how society may be constructed in harmony with all the conditions of pure living, regardless of creeds, conventionalities or traditions, let us not flatter ourselves that such a thing as social science exists."

"It seems that the subtile force of electricity, which has annihilated space in intercommunication, is now to be called in to ameliorate the condition of that large and meritorious class of community, women who support themselves by work with sewing machines, and to make the operation of the sewing machine in the family no longer a task but a luxury. All who have witnessed the operation of Gaumé's Electro-Magnetic Engine, ourselves among the number, are convinced that it must eventually be largely employed as a motor for this purpose. All philanthropists must join us in wishing success to an invention so well calculated to do good. As we shall shortly illustrate and describe this machine at length, we shall not at this time enter into its details. Suffice it to say that the numerous obstacles which have barred the way to success in this field seem all removed, and that the cheap, compact motor so long sought is at last obtained."

"The last trip of the *Pereire* from Havre to New York was performed in eight days and 16 hours. This is the fastest trip on record."

THE NEW

THE ROVER 3500S COMBINES THE BEST OF TWO ENGINEERING WORLDS: AMERICAN POWER & COMFORT—EUROPEAN HANDLING

ROVER engineers devoted themselves to the task of designing a car that combined the best features of European cars with the best features of American cars:

European road-holding, safety, craftsmanship, elegance and economy;

American power and ease of operation—V-8 engine, automatic transmission, power steering, power brakes, power windows, integrated heating and air conditioning.

What's more, Rover engineers insisted the car prove itself by rigorous testing *here* in America. All too often imported cars, even luxury cars, are not particularly suited to the conditions of American motoring. Some imports lack power, others steer heavily, some are too hot in summer or hard to start in winter. For this reason, three years have been spent testing the 3500S on American roads.

One engineering team, for example, responsible for developing the air conditioning system to American conditions, drove 16,000 miles in temperatures ranging from 90°F. to 113°F., from Death Valley to rush hour traffic in Phoenix. Another team spent winter days and nights in Canada to make certain the 3500S would start on the coldest of mornings. It did. Down to minus 44°F. (A typical log of one of the hot weather tests is shown below.)

Every feature of the Rover 3500S was evaluated from this particular point of view. What works best in American driving conditions? The result may be fairly called the first fully American European car.

Here are some of the features:

1. A lightweight aluminum V-8 engine, famous for its smoothness, economy and reliability. (78% of Americans buy cars with V-8 engines. But any American wishing a V-8 with a European chassis available in this country has to pay over $10,000 for it, except in the case of the new Rover 3500S. *Its price:* $5,398*.) What's more, the Rover V-8 is exceptionally economical (our test cars have averaged over 17 m.p.g.), nice for a car capable of 117 m.p.h.

2. Standard equipment on the Rover 3500S is a three speed Borg Warner automatic transmission—easy to service in this country—and electric windows by Delco, also an American company. (The refined air conditioning system is optional, but only $478.)

3. The variable power steering, also standard equipment, automatically adjusts itself to the driver's need. That is, on the straight-away, it lets you "feel" the road. In the turns or when parking, the power increases with the degree of the turn. The tougher the spot, the more help you get.

ROVER V-8

4. The car rides like a limousine but does not pay for its extraordinary smoothness on the straight-away with wallowing in the turns.

5. The 3500S embodies the famous Rover approach to automobile safety. Passengers sit within a rigid steel cage, affording them tremendous protection in case of accident.

6. The 3500S power disc brakes front and rear are recognized for their resistance to fade. As an added safety feature, there is a dash light to warn you when brake pads need replacing, or if the brake fluid gets low.

7. Also standard equipment is the "Icelert," an electronic sensing probe, which causes a warning light to flash, signaling the driver that the temperature is such that ice could be forming on the road.

8. The tires on the 3500S are gold band radial ply, which provide better road holding and longer life.

There are many other fascinating facts about the new 3500S.

If you would like to know more about it, please mail in the coupon below for a 16-page brochure.

The Rover 3500S represents a new statement in automobiles, bringing together for the first time the very best qualities of European and American cars. We believe that it is what a great many people have been waiting for.

*Suggested retail price at East and West Coast ports of entry.

LOG: 21st August through 10th September, 1968

This test car covered 7,677 miles through 24 states, this trip. The gas consumption was 17.7 miles per gallon. Temperature and altitude ranges were 113°F. to 32°F., −277 ft. to 14,110 ft., respectively.

August		Miles	Average M.P.H.	M.P.G. (U.S.)
22nd	Left Teaneck for York via N.J. and Penna. turnpikes	239	60.8	17.9
23rd	Visited York and travelled to Harrisonburg, Va. on Hwy. 81	201	57.4	18.3
24th	Travelled to Birmingham on Hwys. 81, 75, 59 and Rt. 11	638	57.9	17.25
25th	To New Orleans on Hwys. 59, 10 and Rt. 11	364	62.6	16.10
26th	To Fort Worth, on Hwy. 20 and Rts. 61, 190 and 71	539	46.7	17.90
27th	To El Paso on Hwys. 20, 10 and Rt. 80	611	(not noted)	17.65
28th	To Tucson on Hwy. 10	352	(not noted)	19.90
29th	To Los Angeles via Gila Bend and Yuma, Hwys. 10, 8, 5 and Rt. 80	551	56.0	18.60
30th	Stayed in L.A.	14	(not noted)	(not noted)
31st	To Las Vegas via Death Valley and Beatty, Hwy. 15, Rts. 127, 190, 58, 95	481	48.7	16.6

September		Miles	Average M.P.H.	M.P.G. (U.S.)
1st	To Williams, Ariz. via Hoover Dam, Rts. 93 and 66	395	62.8	16.85
2nd	To Raton, N.M. via Albuquerque, Rts. 66, 84, and 85	748	60.9	19.6
3rd	To Colorado Springs, Col., Hwy. 25, Rts. 160 and 85	150	65.2	20.3
4th	Climbed Pikes Peak twice and explored. 18.5 mile climb of Pikes Peak took 53 mins. from Toll gate to summit 14,110 ft., following traffic	129	(not noted)	20.3
5th	Travelled to Des Moines, Hwys. 25 and 80	793	75.0	17.1
6th	To Toledo via Rockford, Ill., Hwys. 80, 50, Rt. 51 and Toll road	705	66.3	17.1
7th	To New York on the Turnpike	610	65.7	19.3
8th	Travelled to Pound Ridge	129	(not noted)	(not noted)
9th	Tests in New York City traffic, Garment District and out to Teaneck. Travelled 10 mi. in 2 hrs. 53 mins. during City Traffic Testing	28	(not noted)	(not noted)

Rover Division
British Leyland Motors Inc.
600 Willow Tree Road, Leonia, New Jersey 07605

Gentlemen: Please send me a four-color brochure on the ROVER 3500S, and tell me the location of a dealer near me.

My name is: ____________

My address: ____________

City__________ State__________ Zip__________

THE AUTHORS

PETER M. WINTER and EDWARD LOWENSTEIN ("Acute Respiratory Failure") are anesthesiologists: Winter is associate professor of anesthesiology at the University of Washington School of Medicine and Lowenstein is in the department of anesthesia of the Massachusetts General Hospital and the Harvard Medical School. Winter was born in Russia and came to the U.S. in 1938 after a stay in Austria. He was graduated from Cornell University and obtained his M.D. at the University of Rochester. He did his postdoctoral work at the Massachusetts General Hospital. He writes: "My special areas of clinical and research interest are respiratory intensive care, the toxic effects of oxygen, and high-pressure physiology. By inclination I would prefer mountain climbing and sailing, but one does have to make a living." Lowenstein did his premedical work at the University of Cincinnati and received his M.D. at the University of Michigan.

H. K. HENISCH ("Amorphous-Semiconductor Switching") is professor of applied physics at Pennsylvania State University and associate director of the university's materials research laboratory. He also has served as consultant to several firms, including the Polaroid Corporation and English Electric Valve Co. Ltd., and currently is consultant to Energy Conversion Devices, Inc. Henisch came to the U.S. in 1963 from the University of Reading in England. His professional interests include semiconductors, transistor physics, contact properties, phosphors and electroluminescence. He has written three books on these topics; a fourth, *Crystal Growth in Gels,* will be published next year. Henisch is also an associate of the Royal Photographic Society. In collaboration with his wife he has recently completed a pictorial and literary essay on chipmunks that will appear next year as a book under the title *Chipmunk Portrait.*

JOSÉ M. CRUXENT and IRVING ROUSE ("Early Man in the West Indies") are respectively head of the department of anthropology at the Venezuelan Institute for Scientific Investigations and Charles J. MacCurdy Professor of Anthropology at Yale University. For many years before taking his present position Cruxent was director of the Museum of Natural Sciences in Caracas and professor of anthropology at the Central University of Caracas. Born in Barcelona, he studied both art and archaeology there. Leaving Spain at the outbreak of the civil war, he taught art in Belgium and Venezuela before turning to anthropology. He is the excavator of Nueva Cadiz, the first Spanish town in South America, and of many other sites in Venezuela. Rouse received his bachelor's degree from Yale in 1934, took his Ph.D. there in 1938 and has been on the faculty ever since. He has made a number of excavations in the Caribbean area, collaborating with Cruxent in Venezuela.

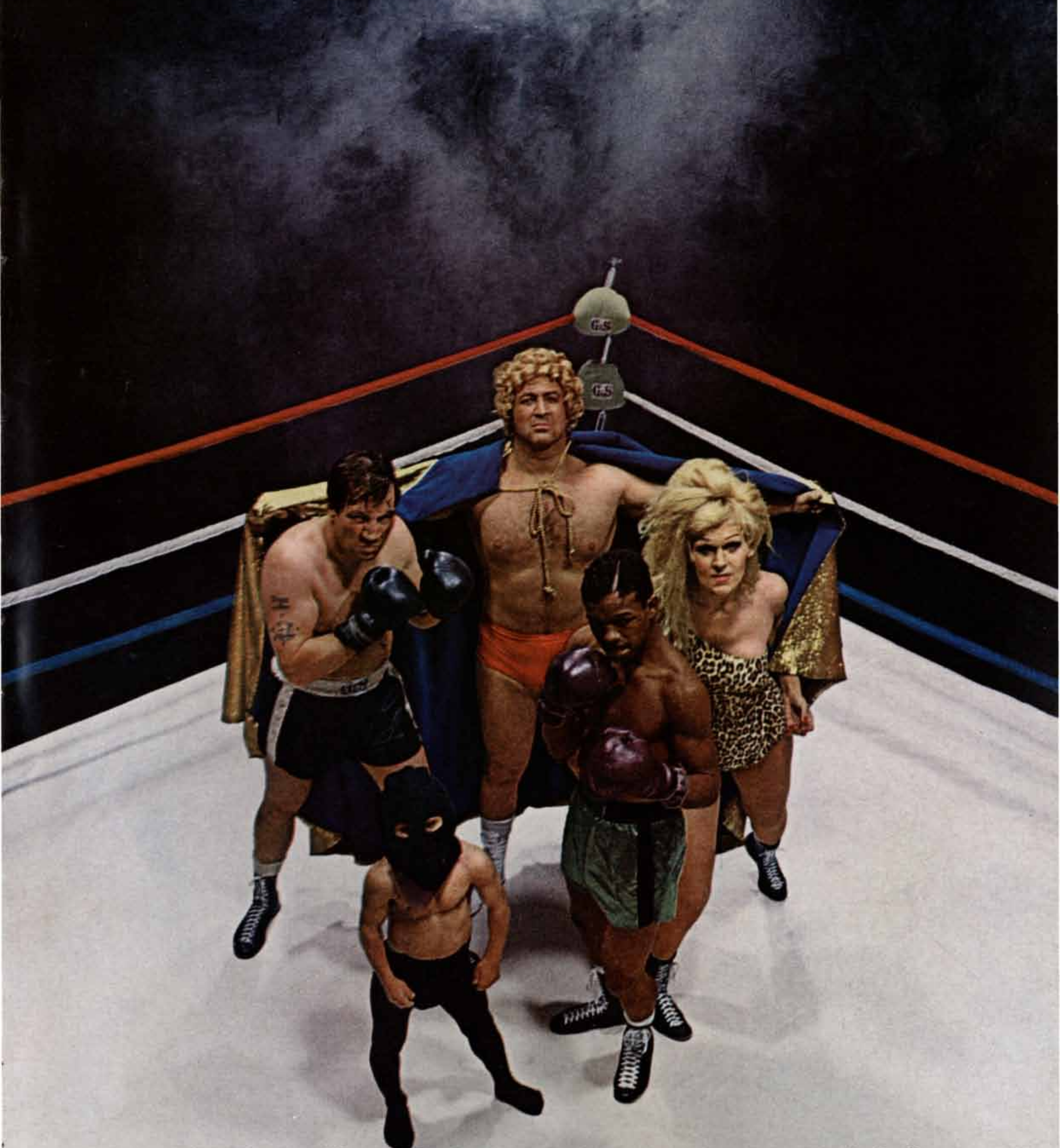
VICTOR E. RAGOSINE ("Magnetic Recording") is manager of the research department of the Ampex Corporation and assistant director of research for the company. He was graduated from Columbia College in 1939 and stayed on at Columbia University for graduate work in physics. During World War II he worked for the Navy as a civilian. Between the end of the war and 1962, when he joined Ampex, he held a number of research positions in industry. "In addition to my current interests in magnetic, electron-beam and laser recording," he writes, "I have done research in radioisotope power sources, dielectrics and thin films."

IMRE TÓTH ("Non-Euclidean Geometry before Euclid") is a member of the faculty of the University of Bucharest who is at present visiting professor at the University of Frankfurt am Main. "My life was deeply marked by war and the Nazi occupation," he writes. "In 1939 I joined the resistance movement: I wrote and printed manifestoes at noon and distributed them at night; I participated in organizing strikes and acts of sabotage. My activities, unlike my zeal in sport and music, seem to have been successful, at least in the opinion of the authorities, who arrested me many times." Tóth, who was graduated from the University of Cluj in Romania in 1949 and received his Ph.D. from the University of Bucharest in 1968, works in mathematics and philosophy. He writes that he also likes "preclassical music and jazz, lyrical poetry, plays of Bernard Shaw and Bertolt Brecht and the novels of Swift, Sterne and Tolkien."

EGON OROWAN ("The Origin of the Oceanic Ridges") is professor emeritus of mechanical engineering at the Massachusetts Institute of Technology. He was graduated from the Technical University of Berlin in 1929 and received his doctorate in physics there in 1932. For several years he was a research worker in England, serving at the University of Birmingham from 1937 to 1939 and at the Cavendish Laboratory of the University of Cambridge from 1939 to 1945. From 1945 until 1947 he was a Nuffield research fellow at Cambridge, and from 1947 to 1950, when he joined the faculty at M.I.T., he was reader in the physics of metals at Cambridge. His work has involved the study of plasticity, fracture and strength, the fatigue of metals and the static fatigue of glass.

RICHARD LOSICK and PHILLIPS W. ROBBINS ("The Receptor Site for a Bacterial Virus") began their collaboration at the Massachusetts Institute of Technology, where Losick received his Ph.D. last year and Robbins is professor of biochemistry. Losick is now a junior fellow of the Society of Fellows at Harvard University. He is studying the mechanism of gene regulation during the differentiation of bacteria into spores. Robbins was graduated from DePauw University in 1952 and received his Ph.D. from the University of Illinois in 1955. Before going to M.I.T. in 1959 he was an assistant professor at the Rockefeller Institute. He is currently spending a sabbatical year at the Imperial Cancer Research Institute in London, learning how to transform cultures of animal cells with tumor-causing viruses.

CRAWFORD H. GREENEWALT ("How Birds Sing") is chairman of the finance committee of E. I. du Pont de Nemours & Company; he was president of the company from 1948 to 1962 and chairman of the board from 1962 to 1967. He is also well known for his ardent pursuit of several avocations. At one time he specialized in growing orchids, and that led to an interest in photography as a means of recording the growth of plants. Later he became interested in photographing birds in flight. In particular he has photographed hummingbirds, a pursuit that has taken him to many countries and resulted in his developing new photographic equipment. His book *Hummingbirds* was published in 1960; in 1961 he received a prize from the American Philosophical Society for a paper on hummingbirds that was chosen as the best paper published in the proceedings of the society during 1960. His book *Bird Song: Acoustics and Physiology* was published last year. Greenewalt began his career with Du Pont in 1922, immediately after he had received his bachelor's degree in chemical engineering from the Massachusetts Institute of Technology.



How'd you like to be in the ring with them every night?

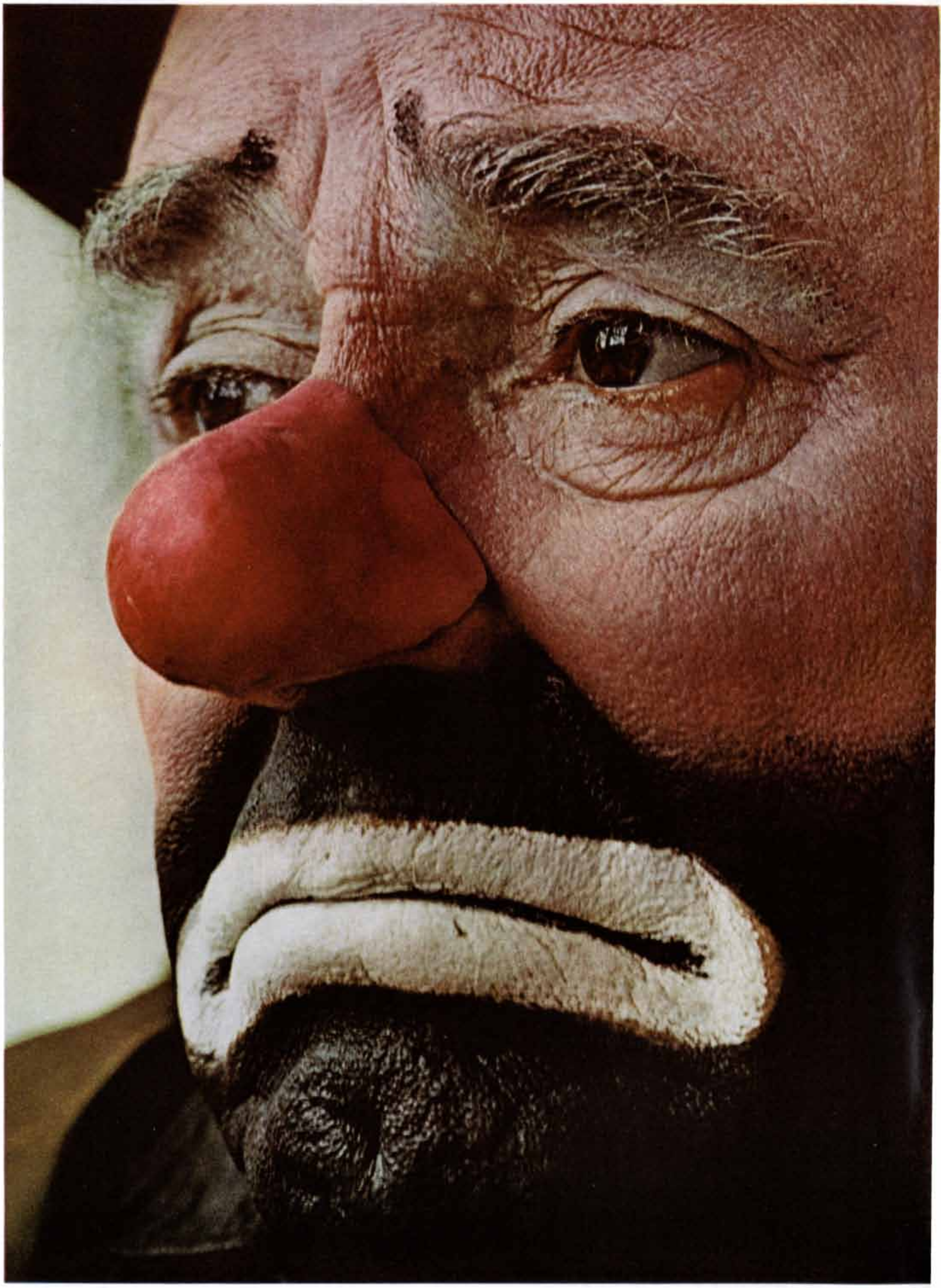
Night after night the canvas is in there getting its living daylights beaten out. As the competition gets tougher, the canvas has to get better. That's why today more canvas is being made of a blend of Fiber 410® rayon and cotton. Rayon gives it more strength and weighs in lighter than cotton. It makes for stronger and lighter weight tents, tarps, boat covers and awnings, too.

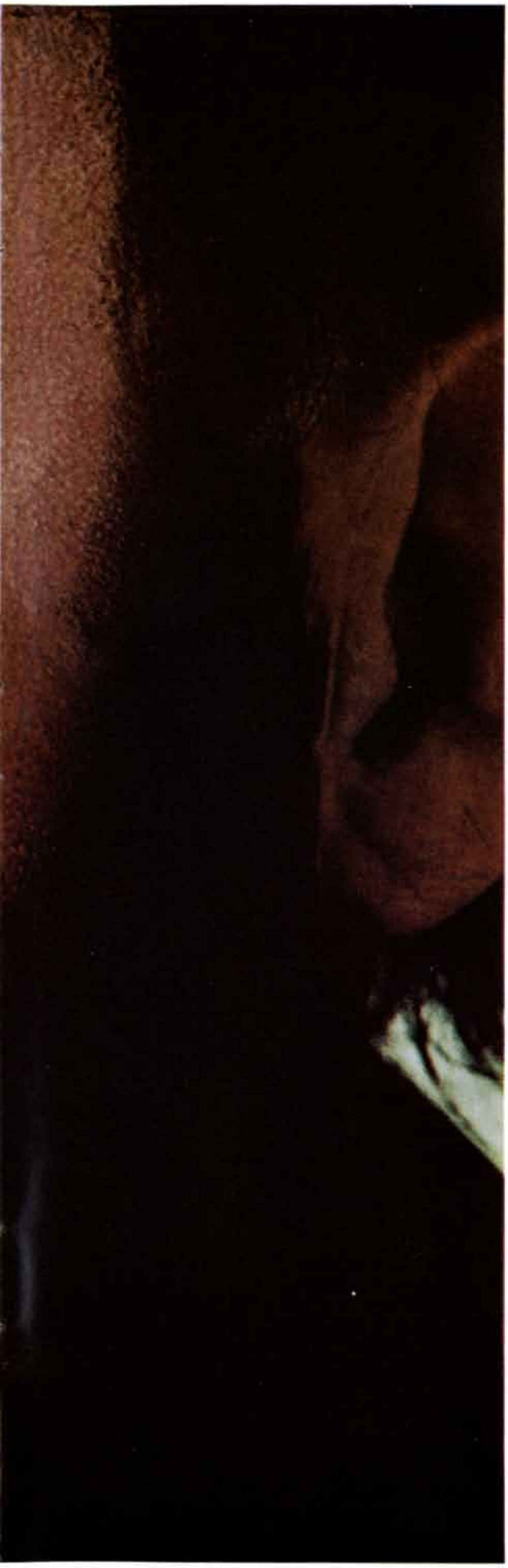
It took a lot of training to get our Avisco® rayon where it is today. We can design it any way you want it. Being man-made, the odds are in our favor that we can engineer it exactly to your specifications. We control it every inch of the way. So every inch of fiber will perform according to plan.

If you've got a plan of your own maybe Avisco rayon can show you the ropes. Just call and let us show you what we're made of. Empire State Building, New York, New York 10001.

fmc® CORPORATION **AVISCO FIBERS**

Is there anything we can't make them do?





Emmett Kelly is happy about our silicones.

To create "Weary Willie" takes a lot of cosmetics or makeup. Once they were hard to apply. They caked, crinkled, gave performers an unnatural look.

Today, cosmetics with silicones give a smooth natural look. Easy to apply, easy to remove.

Union Carbide has developed silicones for an amazing range of products and purposes.

Examples: add silicones to paints and they become easier to apply; cover surfaces evenly, uniformly. Put them in spray-on starch and irons won't stick to cloth. Coat ovens with them and grease wipes off easily. Used on leather, shoes become water-repellent. Added to polishes, they give cars and furniture a quick, long-lasting shine.

Silicone rubber? It stands the extreme temperatures of outer space. And silicones are great for making things stick together that ordinarily wouldn't stick together at all. Like making glass and metals adhere to many materials.

What will we do next with the amazing silicones? There's no telling. We never stop experimenting. And at Union Carbide we experiment with practically everything.

UNION CARBIDE THE DISCOVERY COMPANY

Carl Friedrich Gauss (1777-1855) — *Mosaic Portrait by Emily Syminton DeGroat From the Planning Research Corporation Collection*

"He made contributions of fundamental significance not only in almost every leading field of pure mathematics, but also in astronomy, geodesy, electricity, and magnetism."[1]

[1]David Eugene Smith, *A Source Book in Mathematics,* p. 292, McGraw-Hill, New York, 1929.

INTERACTIONS OF DIVERSE DISCIPLINES

Of the 2500 professional men and women on the staff of Planning Research Corporation, 194 are mathematicians. Their colleagues represent the classical sciences, the behavioral, management and computer sciences, economics and nearly all branches of engineering. These disciplines, interacting on multidisciplined teams, form the most powerful of all problem solving tools and they are currently being employed by industry and governments in 25 nations.

The Corporation's Annual Report, which gives an overview of capabilities in new facilities planning and management planning, will be sent to interested executives. Address Mr. Stanley L. Gendler, Senior Vice President.

PLANNING RESEARCH CORPORATION

Home office: 1100 Glendon Avenue, Los Angeles, California 90024

An Equal Opportunity Employer. Candidates are invited to write to the Administrator for Professional Staffing.

Established 1845 SCIENTIFIC AMERICAN November 1969 Volume 221 Number 5

Acute Respiratory Failure

This crisis of life has come to be regarded as a distinct clinical entity. In respiratory intensive-care units it is now being treated by teams of specially trained physicians and supporting personnel

by Peter M. Winter and Edward Lowenstein

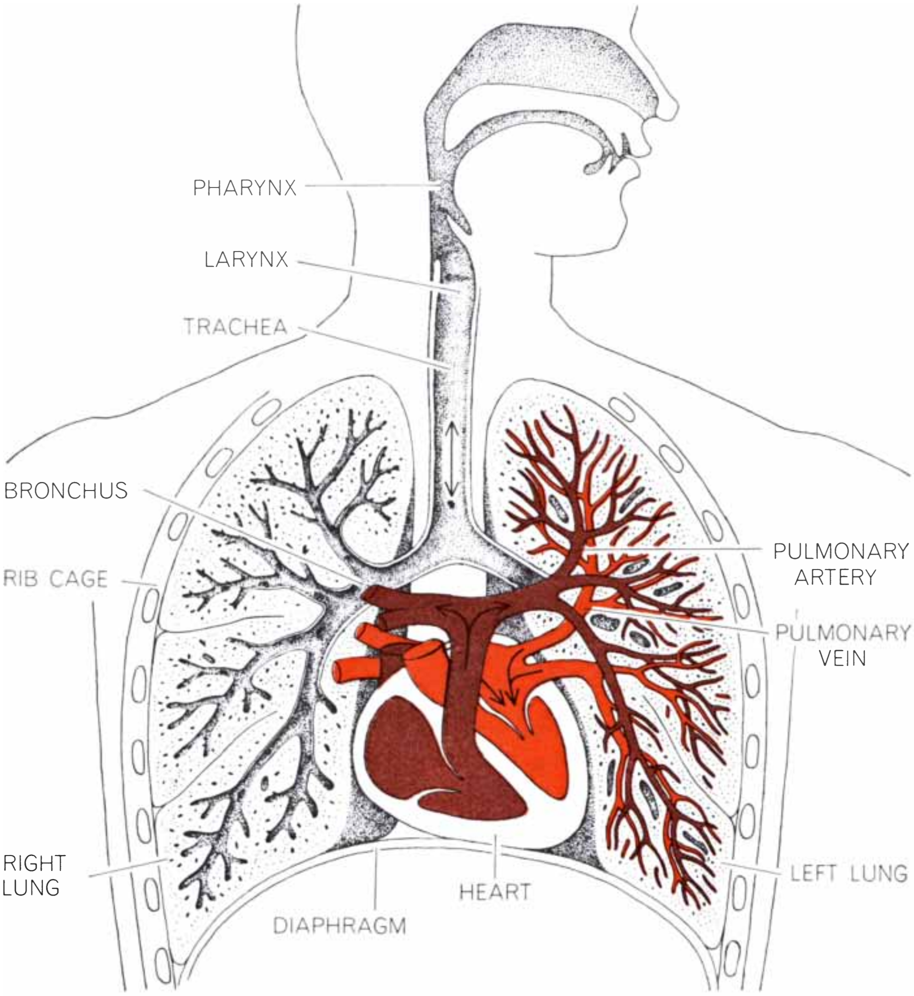
The failure or serious impairment of a bodily function that is essential for life is logically regarded as being fatal. One such function is that of the lungs, which deliver oxygen to the blood and remove carbon dioxide and must keep the blood level of both gases within strict bounds. Yet respiratory failure is now reversible in a large percentage of cases if proper treatment is provided. Such treatment is available in respiratory intensive-care units: properly equipped hospital facilities directed by a new kind of medical specialist, the intensivist, and manned by teams of trained physicians and supporting personnel.

The increasing capability of respiratory intensive care in reversing lung failure is the result, first of all, of an increasing discourse between respiratory physiologists and physicians who treat patients. It is not sudden breakthroughs in physiology that have made the difference but rather the fact that data that have long been available are now being brought to bear through active intervention to preserve the life of critically ill patients.

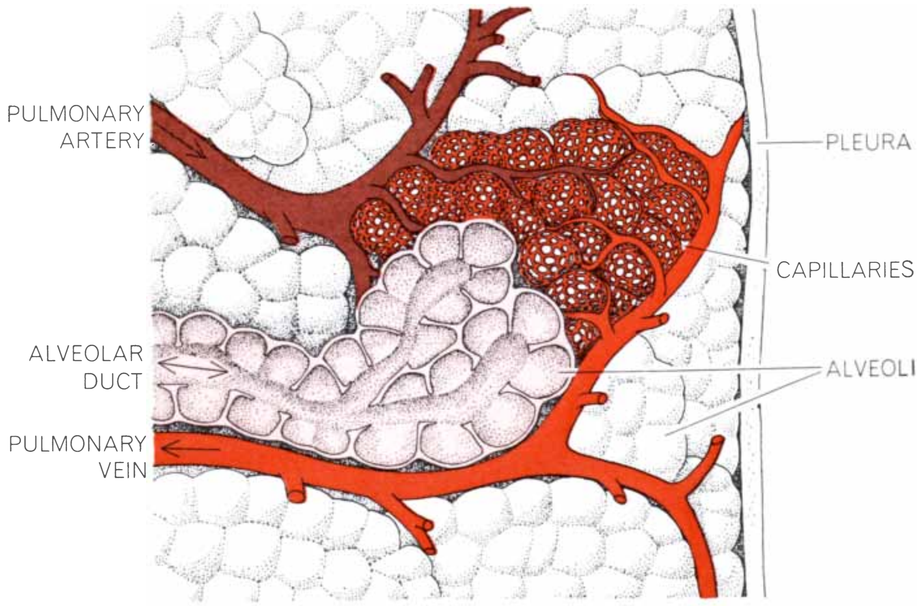
The advance in measurement technology and medical engineering is another reason respiratory failure can now be treated. Because the lung (like most organs) has a vast functional reserve, deterioration may occur long before any trouble is apparent; what later appears to be sudden respiratory failure may be only the final event of a long process. The ability to detect and treat this deterioration early is the key to respiratory intensive care; it derives from the fact that quantifying procedures formerly confined to the research laboratory can now be performed quickly, frequently and accurately in a clinical setting during a medical emergency. The treatment of acute respiratory failure is probably as close to being a quantitative science as any field of clinical medicine can be today; in this situation precise measurement approaches or exceeds in importance the "clinical judgment" that for so many years has been the prime quality of the good physician.

In order to understand the way the respiratory system may fail it is first necessary to know how it works normally [*see top illustration on next page*]. The primary purpose of the lung is to bring air and blood into intimate proximity so that the necessary influx of oxygen and removal of carbon dioxide can take place. This is achieved by the combined efforts of two pumping systems, one moving a gas and the other a liquid; the two systems bring the two fluids into apposition so that they are separated by a thickness of only 1.5 microns. It is at this interface that the passive diffusion by which all gas exchange in the lungs is accomplished takes place. The right ventricle of the heart pumps all the blood through the 150-milliliter volume of the pulmonary capillary circulation. The total surface area over which this blood is spread is enormous: approximately 750 square feet. The blood is thus presented as a critically thin film to the air, which has been pumped into the lungs by the diaphragm and chest muscles. The air and blood are brought into apposition at some 300 million alveoli, tiny blind pouches surrounded by the pulmonary capillaries [*see bottom illustration on next page*]. Of the air that is breathed in, 60 to 80 percent takes part in gas exchange in the alveoli. The rest remains in "dead space" in the trachea, the bronchi and the smaller conducting tubules of the lung and does not come into contact with pulmonary capillaries.

In the normal lung venous blood, which has given up oxygen to the tissues and has taken up carbon dioxide from them, perfuses the pulmonary capillaries. The air in the alveoli is rich in oxygen and low in carbon dioxide. For less than one second they are in apposition, and during that time all pulmonary gas exchange takes place: the lung equilibrates, or makes equal, the partial pressures (tensions) of the two gases in the air and in the blood [*see illustration at top left on page 25*]. The optimal function of the lung is dependent on the correct matching of ventilation (the delivery of inspired air) and perfusion (blood flow through the lung). While the actual distribution of pulmonary blood and alveolar gas is exceedingly complex, for simplicity the matching may be thought of as being nearly perfect in the healthy state, and in this situation the



IN THE LUNGS air and blood are brought into close proximity by two pumping systems so that oxygen can enter the blood and carbon dioxide can leave it. Here the right lung is drawn to show the respiratory tree through which air moves and the left lung is drawn to show the pulmonary vessels through which deoxygenated blood (*dark color*) is moved to the lung and oxygenated blood (*bright color*) is carried away. Inspiration occurs when the diaphragm descends and the chest expands, creating a partial vacuum into which air flows; expiration is caused by the reverse motions. If these motions are stopped abruptly, asphyxia occurs; if they are partially impeded, more subtle changes take place in the lungs.



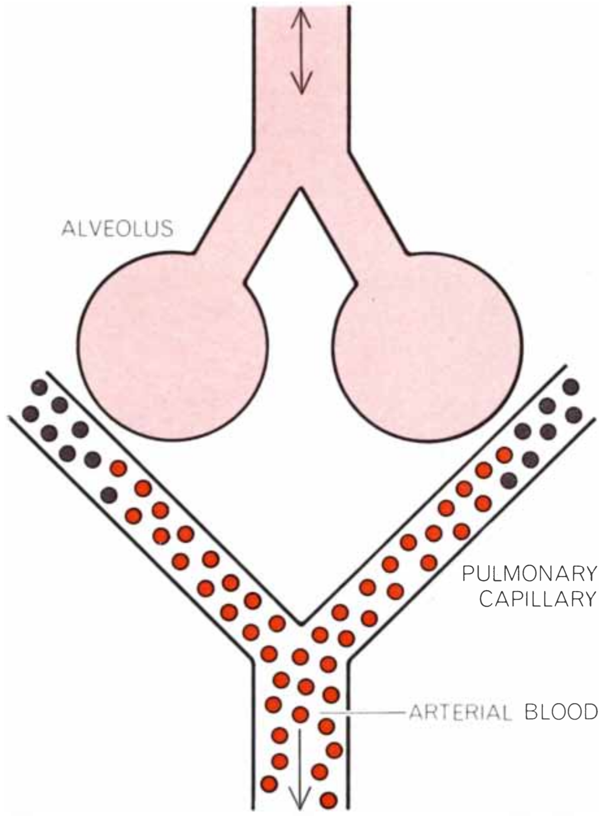
GAS EXCHANGE is effected across the thin membrane of some 300 million alveoli, tiny sacs surrounded by pulmonary capillaries. In this drawing some of the grapelike alveoli are shown with capillaries and some without them, and some are shown in cross section.

uptake of oxygen and the elimination of carbon dioxide are maximally efficient.

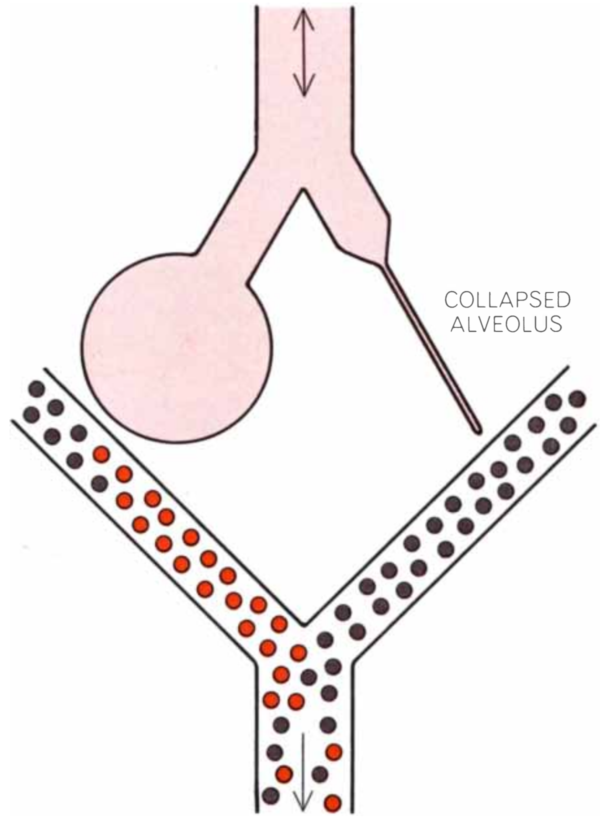
From this simple model one can predict what changes will take place at the alveolar-capillary level in various diseased states and what the results of each such change will be. Pathology manifests itself as changes in the tensions of oxygen and of carbon dioxide in the blood, and therefore accurate and frequent measurement of these changes is absolutely essential in the quantitative treatment of respiratory disease. A lack of oxygen results in tissue damage and ultimately in tissue death. High carbon dioxide levels depress the circulation. The carbon dioxide level is also a major determinant of the acid-base status of the body, and disturbance of this balance may adversely affect many organ systems. Respiratory failure can therefore be defined as the inability to maintain viable blood homeostasis, or stability, with respect to oxygen, carbon dioxide and acidity.

The collapse of alveoli may be the commonest cause of arterial hypoxemia, or a low oxygen level in blood, in surgical patients [*see illustration at top right on opposite page*]. The pain of an incision may prevent the periodic deep breaths that normally keep alveoli from collapsing. When, as a result, a number of alveoli collapse and no longer present air to the blood, deoxygenated venous blood is admitted directly into the arterial circulation, functionally bypassing the lung. This blood, containing little oxygen, dilutes the blood that has come from normal alveolar-capillary units and has the proper high level of oxygen; the degree of this venous admixture is measured by the difference between the alveolar and the arterial oxygen levels. The progressive collapse of more and more alveoli leads to an increase in venous admixture and thus to arterial hypoxemia. Less and less oxygen is taken up from the lungs, causing decreased delivery to such critical tissues as the brain, heart and kidneys. This process may sometimes be dramatic; often, however, it is insidious, with few warning signals given until shortly before death. It can only be diagnosed in time if one measures the alveolar and arterial oxygen tensions or has some approximation of these measurements.

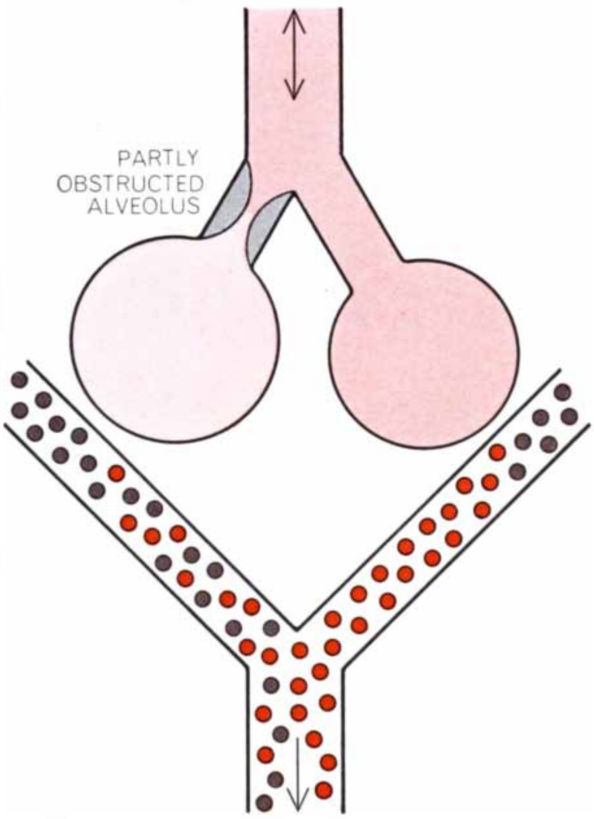
The partial obstruction of alveoli is a more common cause of arterial hypoxemia [*see illustration at bottom left on opposite page*]. The result is the same as for collapsed alveoli, although the fundamental pathology is different. In this case the alveolus is not collapsed and may actually be excessively large, but it



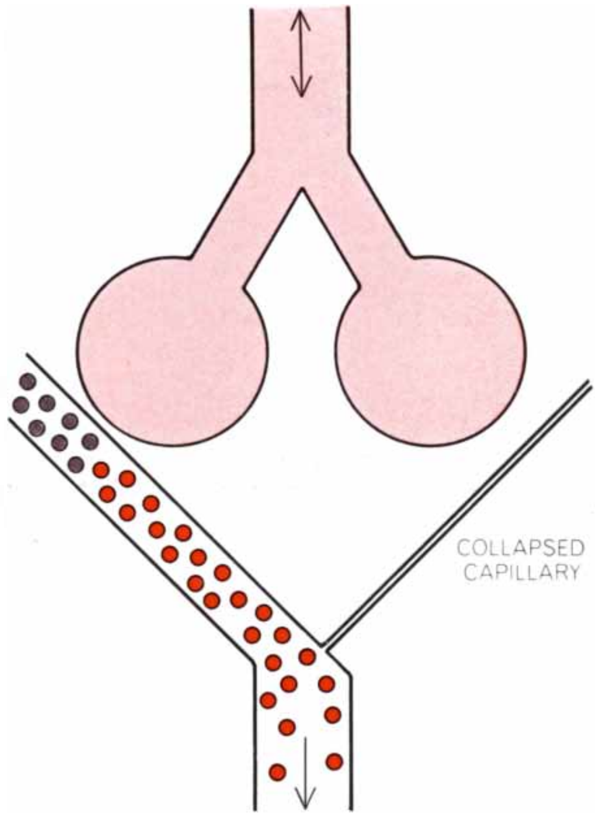
NORMAL INTERACTION between alveolus and capillary is shown schematically. Venous blood (*gray disks*) with low oxygen tension and high carbon dioxide tension is brought to the alveoli. Oxygen (*light color*) diffuses into the capillaries, combines with hemoglobin and returns oxygen tension to arterial levels (*bright colored disks*). At the same time carbon dioxide is removed.



COLLAPSED ALVEOLI are common after major surgery. The blood that passes a collapsed alveolus (*right*) does not take part in gas exchange, in effect bypassing the lung. Low in oxygen tension, it mixes with the oxygenated blood from normal alveoli. A lower than normal oxygen tension is therefore found in arterial blood. If it is severe enough, this "venous admixture" can cause tissue damage.



PARTIALLY OBSTRUCTED ALVEOLI are common in chronic pulmonary disease. Oxygen is removed from the obstructed alveolus by the blood more rapidly than it can be replaced by ventilation. The mean oxygen tension in the alveolus is reduced and the blood is only partially oxygenated, again leading to venous admixture.



COLLAPSED CAPILLARIES do not affect oxygenation, but alveoli that are not perfused cannot take up carbon dioxide from the blood and become alveolar dead space. Total ventilation must be increased in order to compensate for the dead space. If it is not, carbon dioxide is retained and the acid-base balance is disturbed.

is underventilated in relation to its perfusion: oxygen is extracted by pulmonary capillary blood more rapidly than it is replaced by air entering the alveolus. This maldistribution of ventilation and perfusion is the prime cause of hypoxemia in such diseases as pulmonary emphysema, extreme obesity and certain diseases of the heart valves.

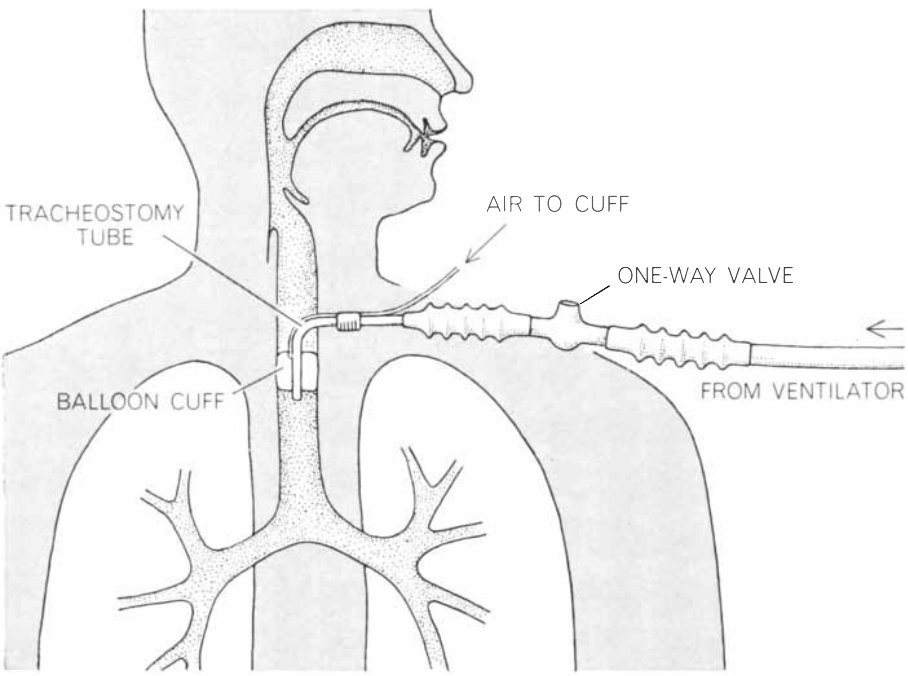
The elimination of carbon dioxide is the other major function of the lung. Metabolic activity determines how much carbon dioxide is produced and therefore how much needs to be eliminated. For normal arterial carbon dioxide levels to be maintained the correct volume of fresh gas must pass into and out of the lung; the adequacy of ventilation, as opposed to oxygenation, is therefore determined by measuring carbon dioxide tension in the blood. Carbon dioxide accumulation is common in the ventilatory depression that accompanies drug poisoning or in injuries that produce instability of the chest wall.

Ventilation at a level of about five liters per minute is normal, but conditions may exist in which even two or three times this volume cannot provide adequate ventilation. The reason is that only the portion of a breath that actually reaches perfused alveoli takes part in gas exchange. The air that remains in the anatomical dead space in the bronchial tree does not take part in carbon dioxide elimination and hence is wasted ventilation from a functional point of view. Anatomical dead space remains relatively constant and comprises essentially all the dead space in normal man, but it should be clear that any ventilation of nonperfused alveoli is also wasted. This is called alveolar dead space, and it may increase dramatically in disease if a pathologic state leads to the collapse of pulmonary capillaries without affecting the alveoli [*see illustration at bottom right on preceding page*]. The reason is clear. If an alveolus is ventilated but not perfused, the entire volume of that alveolus becomes dead space. Massive increases in dead space occur with severe decreases in blood pressure, hemorrhage or obstruction in the pulmonary circulation. Under these circumstances as much as 70 to 80 percent of each breath may be wasted. More and more gas must be moved by the respiratory muscles to overcome this inefficiency in carbon dioxide elimination. When dead space exceeds about 60 percent, ventilatory requirements can no longer be met and at that point intervention is mandatory to prevent death from exhaustion or from the retention of carbon dioxide.

The ratio of dead space to the total volume of each breath is thus an index of inefficiency in one aspect of lung function, the excretion of carbon dioxide. The difference between the tension of oxygen in the alveolus and in arterial blood is an index of inefficiency in the other prime aspect of lung function, oxygen uptake. The derangement of oxygen uptake and of carbon dioxide excretion usually occur together; indeed, until the development of blood-oxygen measurements an elevated carbon dioxide level was considered a prerequisite of respiratory failure. Nevertheless, as the preceding discussion of the separate mechanisms makes clear, it is quite possible for hypoxemia to occur alone. Isolated hypoxemia is a more insidious process than hypoxemia combined with carbon dioxide retention, and can only be recognized through measurements. As we have indicated, detection and treatment before symptoms appear are the most important factors in the decrease of mortality from lung failure.

When the body's mechanisms for maintaining blood-gas homeostasis are overwhelmed by a disease process, respiratory failure is a fact and assistance in ventilation becomes mandatory to preserve life. Simple methods such as mouth-to-mouth rescue breathing may suffice for a short time, but in order to ventilate a patient for more than a few hours it is essential that a safe and reliable connection be made between the patient's trachea and a mechanical ventilator. A tube may be passed through either the mouth or the nose and down into the trachea. Such an endotracheal tube is a satisfactory solution for several days, but if ventilatory assistance is prolonged, there is a risk that the tube may damage the larynx. To preclude this a tracheostomy is performed: an incision is made in the neck and a tube is passed directly into the trachea below the level of the larynx [*see illustration at left*]. Both methods give the physician positive control of the airway. Modern endotracheal tubes have inflatable balloon cuffs, which provide the gas-tight fit that is necessary for the precise control of ventilatory volumes and for protecting the lower airway and the lungs from vomitus and other harmful foreign materials.

Almost all ventilators currently in use are of the "intermittent positive pressure" type. Either a piston or a high-pressure gas source forces the proper mixture of gases into the patient's lungs through the endotracheal tube. Expiration is brought about by passive recoil. Ventilation is gauged by analysis of the



VENTILATOR delivers the proper gas mixture through intermittent positive pressure: gas is driven into the respiratory tract until a specific quantity has been delivered or until a specific pressure has been attained, and expiration is achieved by the passive recoil of the lungs and chest wall. Here a tracheostomy has been performed and the tube passes through an incision in the patient's neck. The inflatable balloon cuff ensures a gastight fit in the trachea, so that the desired humidity, gas mixture and respiratory volume can be maintained.

PRECEDING EVENT	PATIENTS (NUMBER)	SURVIVAL (PERCENT)	LENGTH OF STAY (DAYS)
SURGERY			
HEART SURGERY	82	51	32
REMOVAL OF THYMUS GLAND	43	100	13
OTHER CHEST SURGERY	78	69	13
ABDOMINAL SURGERY	150	44	14
NEUROSURGERY	60	55	8
TRAUMA			
CRUSHED CHEST	47	87	32
MISCELLANEOUS			
DRUG OVERDOSE	60	85	8
NEUROMUSCULAR DISEASE	57	75	57
EMPHYSEMA	43	53	9
OBESITY, CARDIAC ARREST PNEUMONIA, BURNS AND OTHERS	98	72	15
TOTAL	718	65	20

RESPIRATORY UNIT of the Massachusetts General Hospital, directed by Henning Pontoppidan, had an overall survival rate of 65 percent in treating 718 patients in acute respiratory failure from 1961 through 1968, as shown here. The rate rose steadily, reaching 81 percent in 1968. At the same time the patients' average stay in the unit increased from 12 days in 1961 to 20 days in 1968, the year for which length-of-stay data are given here. The vast majority of these patients would have died without the respiratory intensive care available in the unit. Only the emphysema patients and some chest-surgery patients had primary disease of the lungs; in the others the failure of lung function was secondary. It is clear that the survival rate is highest when there is no intrinsic lung damage.

carbon dioxide level in the blood. If ventilation is inadequate, the ventilatory rate or volume is increased to bring the carbon dioxide tension back to normal. Since either gross overventilation or underventilation is harmful, the carbon dioxide level is monitored frequently.

In the correction of hypoxemia the function of the ventilator is to increase the oxygen tension enough to compensate for the collapse or obstruction of alveoli, that is, for intrapulmonary shunting or for maldistribution as between ventilation and perfusion. To do this the oxygen concentration in the mixture supplied to the patient can be enriched beyond the normal 21 percent of ambient air. The eventual return of blood-oxygen tension to normal must then be confirmed by direct measurement. On the one hand, an inadequate oxygen tension in the inspired gas will clearly be ineffective in alleviating the hypoxemia; on the other, an excessively high tension may cause severe and potentially lethal damage to the lung. This makes it mandatory that the patient receive only as much oxygen as is needed to bring the blood oxygen up to the normal level. In the most severe situations, however, ventilation with even pure oxygen will not return blood oxygen to adequate levels. As one approaches this stage, the toxic effect of oxygen on the lung must be ignored; what is crucial is that such vital organs as the brain, kidney and heart do not undergo irreversible damage. Fortunately, as we discovered in our laboratories, arterial hypoxemia provides partial protection against pulmonary oxygen toxicity. Only when oxygen delivery by the blood to the tissues has been assured is it reasonable to consider the poisonous effects of oxygen.

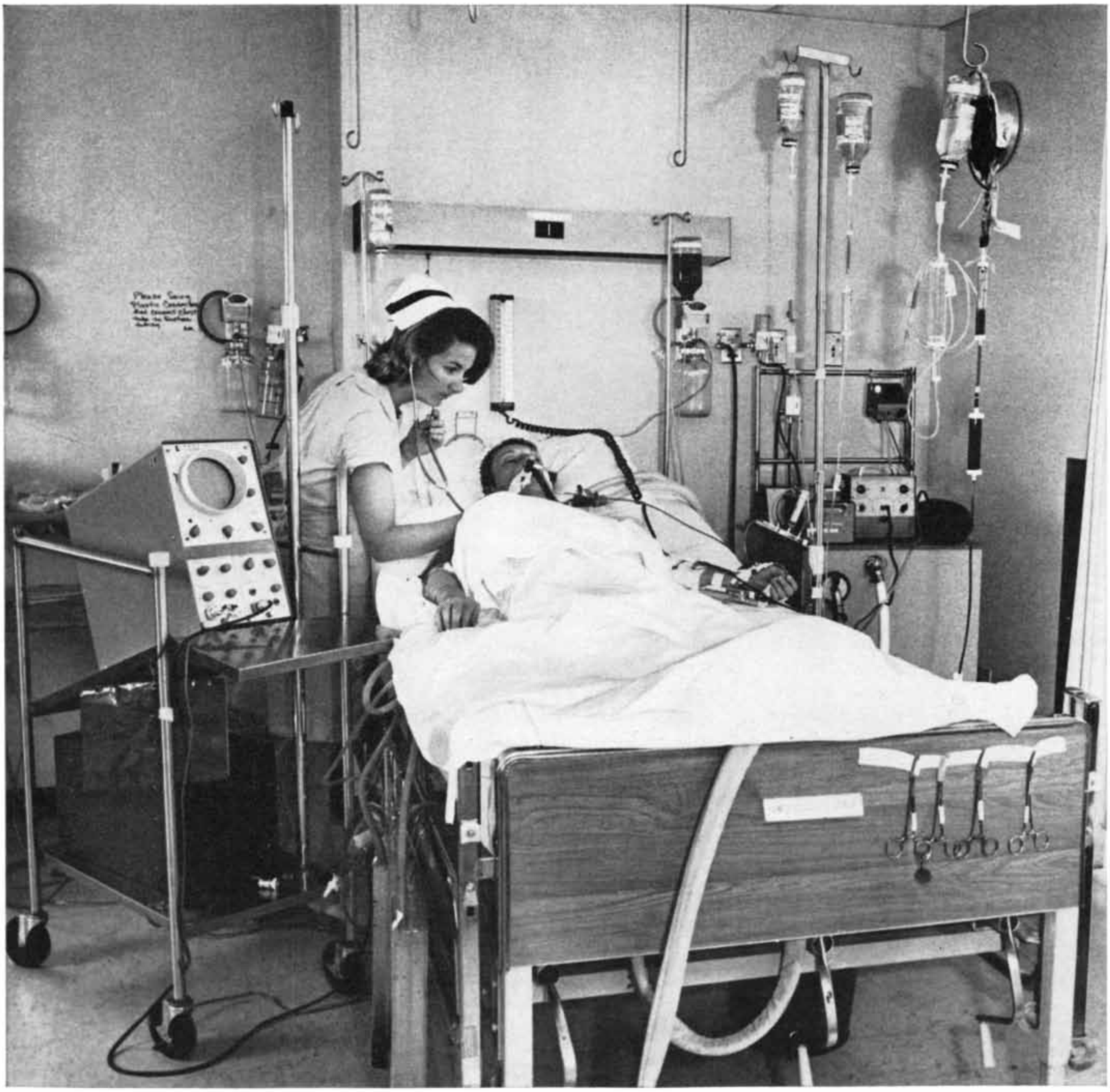
As the patient's condition either improves or deteriorates, the requirements for both ventilation and oxygenation change. These changing needs can be met only by manipulating the ventilating system in response to repeated measurements of blood-gas tensions. In this way one attempts to preserve life while the underlying disease is treated or spontaneous recovery takes place. Artificial ventilation must be continued until a patient is again physically able to oxygenate his blood and eliminate carbon dioxide to meet his bodily requirements.

Respiratory intensive care does not consist exclusively of ventilation. Many other problems must be dealt with continually. Patients so ill that they cannot breathe almost invariably have serious problems in other organ systems that require attention. In addition, the treatment of respiratory failure may in itself cause serious and potentially lethal difficulties. For example, the normal humidification system of the lungs is interfered with by artificial ventilation. One must therefore take steps to humidify the inspired gas in order to avoid drying, encrustation and ultimate obstruction of the airways. The endotracheal tube may invade adjacent tissues or blood vessels or be the cause of later scarring and narrowing of the air passages. The probability of pulmonary infection is increased during artificial ventilation, and so the airway must be treated as a sterile surgical field. As we have indicated, the very oxygen that is saving life may cause damage to the lungs, and this must be guarded against. The pressure needed to inflate the lungs artificially may interfere with the circulation, and so precise control of the ventilatory pattern is frequently critical. The hormonal balance of the body is greatly affected by mechanical ventilation, sometimes leading to the retention of water in the lungs and thus to further interference with oxygenation. In the course of treatment each of these potential complications of therapy must be identified, prevented if possible or else dealt with when it appears.

The application of physiological knowledge to a clinical problem and the interrelated manipulations needed to maintain homeostasis in the face of severe respiratory failure clearly make for a complex undertaking. The complexity has resulted in the creation of what is really a new discipline. The physician without special training in respiratory intensive care is no more qualified to do this work than one without proper training would be to do cardiac surgery. Training programs and special facilities have therefore been developed, with emphasis on a team approach.

The need for such an approach to the care of the critically ill patient exemplifies a subtle but most significant change in medical care. For many years a patient was the sole responsibility of the physician who specialized in the illness that brought the patient to seek medical attention. This remained true even when problems that developed in the further course of this disease carried the original doctor far beyond his area of particular competence. He might call in other specialists for consultation, but he retained control of the case.

Now, however, a patient may be categorized by degree of illness as well as by type of illness. Ideally a special team now assumes primary responsibility for the approximately 5 percent of the patients in a general hospital who are in imminent danger of dying. This team should be headed by an "intensivist," one who restricts his practice to the care of such patients, and it should include physicians from many other specialties. A large group of trained ancillary medical personnel should also be included. In the area of respiratory failure this group

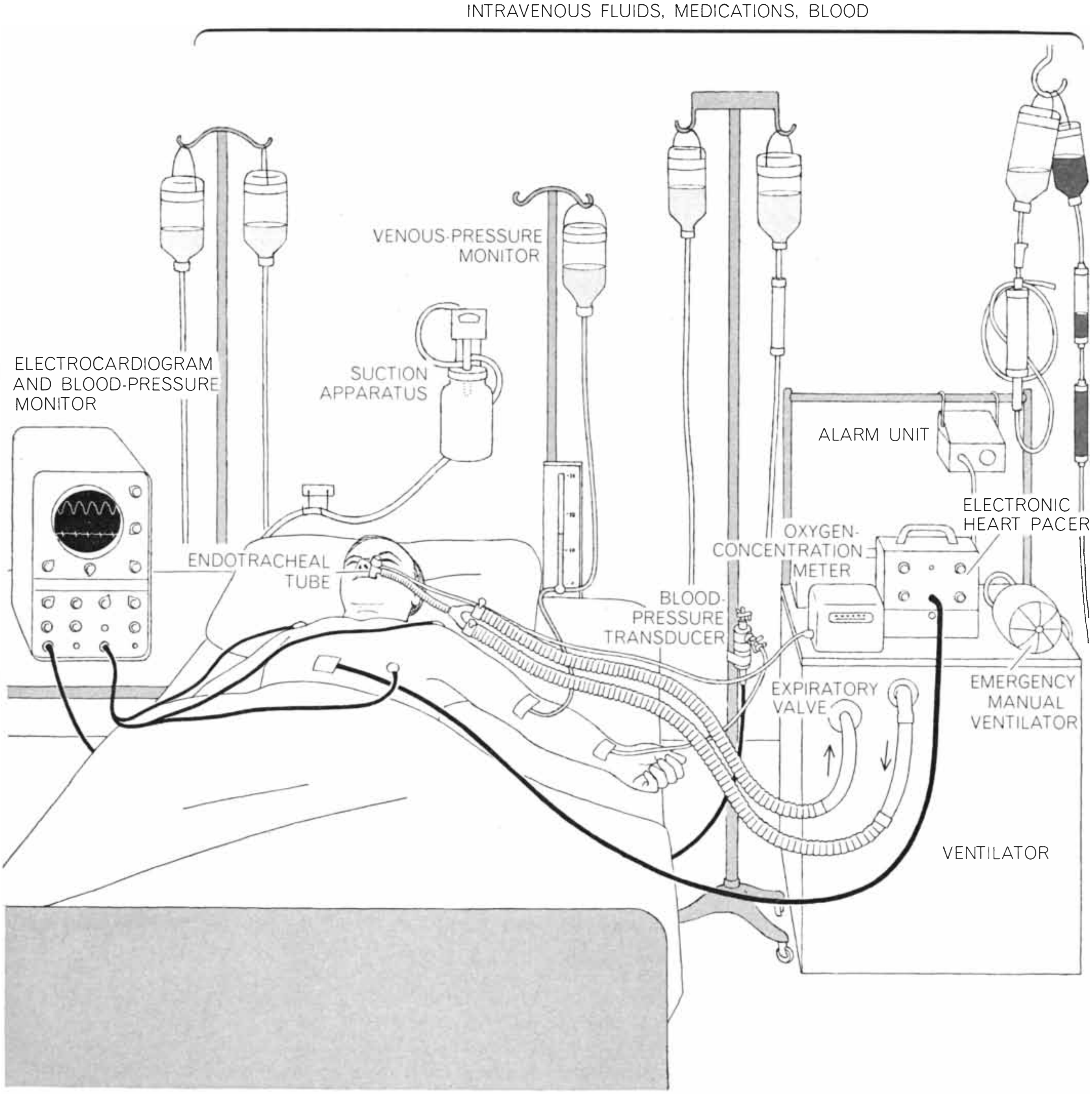


PATIENT in a respiratory intensive-care unit is attended by a specially trained nurse and surrounded by a maze of equipment. The patient, who had undergone a heart-valve operation, is being ventilated by a volume-limited device via an endotracheal tube.

should include nurses specializing in respiratory care, clinical laboratory technicians, pulmonary physiotherapists and inhalation therapists. Finally, the treatment facility itself must be specially designed to satisfy requirements that are quite different from those of the ordinary hospital ward. For all these reasons, the prolonged treatment of severe respiratory failure should not be undertaken lightly. The commitment is enormous, and in our judgment long-term respiratory care can be carried out successfully only in large hospital centers. Plans are now being formulated in several areas for the routine transfer of patients who need this intensive treatment to respiratory care units.

Failure of lung function is often not the result of primary lung disease, as statistics from the Respiratory Unit of the Massachusetts General Hospital make clear [*see illustration on page 27*]. Instead it is often secondary to some other condition such as poisoning, neurologic disease, peritonitis, injury to the chest, heart failure or the process of surgery itself. The intensivist specifically trained in the management of ventilation must therefore work in close cooperation with a specialist concerned with the patient's primary disease.

Respiratory care is now being practiced by physicians whose original training was in anesthesiology, surgery, internal medicine or pediatrics. The attention being paid to the respiratory system by these specialists in varied disciplines testifies to the ubiquitous nature of respiratory failure: At some stage stress associated with almost any disease can impair lung function and the restoration of that function may be the crucial step in restoring the patient to health. The combined efforts and diverse viewpoints of the members of intensive-care teams should continue to increase the rate of recovery from respiratory failure.



EQUIPMENT for respiratory intensive care is diagrammed. The alarm is for ventilator failure. Blood pressure is transmitted from a wrist artery to the transducer and thence to the monitor. The suction apparatus is for removing secretions from the trachea.

AMORPHOUS-SEMICONDUCTOR SWITCHING

Solid-state devices that are glassy rather than crystalline can be employed to control the flow of electric current. They are not yet in commercial use but are under intense study all over the world

by H. K. Henisch

The behavior of electrons in solids has probably absorbed more man-hours of technological investigation than any other physical phenomenon. Out of these investigations has risen the electric power industry, the telecommunications industry and most recently the computer industry. Since the invention of the transistor in 1948 an enormous amount of electronic research has been devoted to studying the behavior of electrons in extremely pure crystals of semiconducting substances, notably germanium and silicon. Last year alone more than 10,000 technical papers were published on semiconductors and related electronic devices, and the worldwide annual production of transistors approached a billion units.

Within the past few years the electronics industry has been asked to consider the possibility that semiconducting devices of a novel kind can be made from solid substances that are neither crystalline nor highly pure. These substances are in fact amorphous; thus in the traditional view they are the very antithesis of the crystalline materials that had seemed to be uniquely suited for the making of solid-state electronic components. The new amorphous devices have not yet found their way into commercial use, but they are under intense study in many of the world's major electronics laboratories. Perhaps the most important examples—and certainly the ones that have received the most public attention—are the threshold and memory devices named Ovonic switches by their inventor, Stanford R. Ovshinsky. The threshold switch is a two-terminal component that can have two states in an electrical circuit: an almost nonconducting ("off") state and a conducting ("on") state. Normally the device is in the "off" state, but when the voltage across it reaches a certain threshold value, it flips over to the "on" state and thus acts as a switch. On removal of the applied voltage the "off" state is immediately restored. The memory switch behaves in much the same way, except that it will remain in the "on" state once it has been switched on, even though the applied voltage subsequently drops below the threshold value. This "on" state is normally stable, but the device can be returned to the "off" state by passing a sustained pulse of high current. The response of the electronics industry to the possibilities presented by the new amorphous devices has not been warm, which is perhaps understandable considering the large sums the industry has spent in developing the technology of crystalline semiconductors.

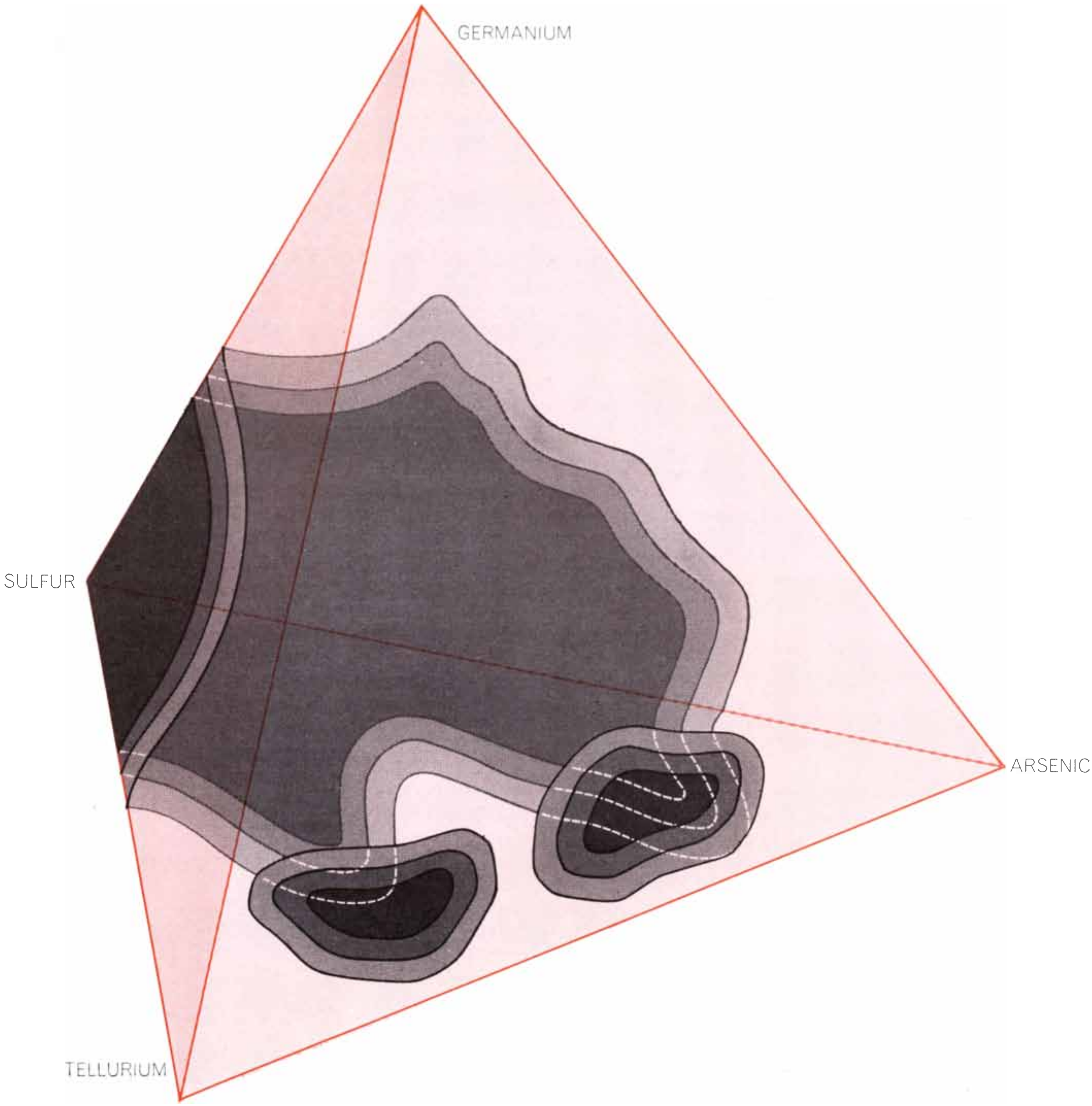
In this article I shall describe the new components and explain why they work as they do. The explanations are still tentative because the world of amorphous matter is still imperfectly understood; indeed, most physicists and electrical engineers have long regarded it as alien territory. Nevertheless, the operative mechanisms and the relations between amorphous and crystalline materials are slowly being clarified.

Today these two categories of solids are no longer regarded as being sharply distinct. At the beginning of the century physicists were busy pointing out that crystals are orderly arrays of atoms, ions or molecules. More recently they have been just as busy insisting that the order is limited, and that imperfections are always present because of impurities or native irregularities, or both. At the same time physicists have come to recognize that even the liquid state, in which atoms and molecules wander about freely, must have a certain degree of short-range order. How else could one explain the fact that many crystalline solids—metals and semiconductors as well as insulators—show no change in electrical properties when they pass through the melting point and lose their crystal structure?

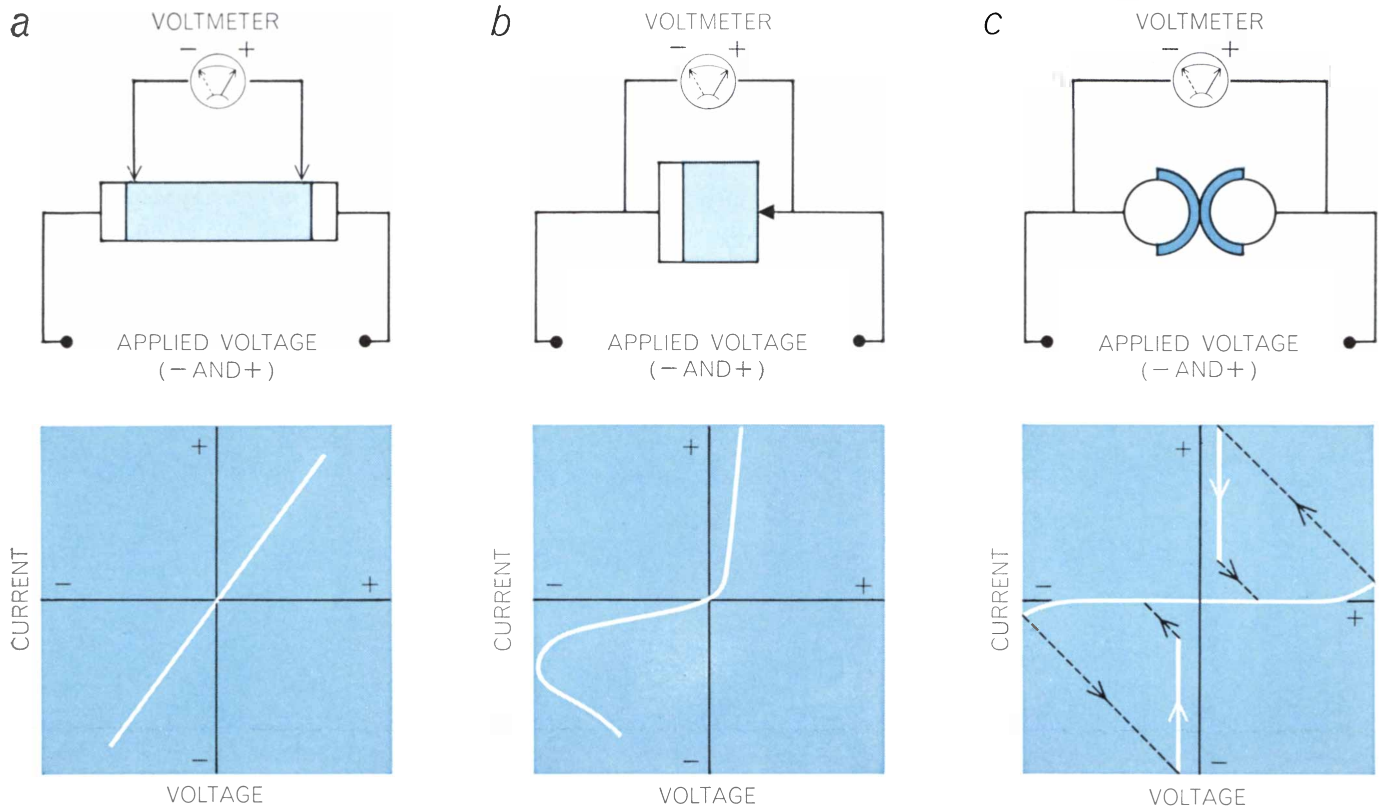
Moreover, it is now known that certain organic compounds (for example para-azoxyanisole, para-nitrotoluene and cholesterol) can exist as assemblages in which molecules are strongly bonded in geometric patterns within a plane but in which different planes are related to one another in an essentially random way because the forces between them are weak. Such substances, often called liquid crystals, exhibit some characteristics of crystals and some of liquids, which goes to show that crisp definitions and rigorous distinctions are hard to come by in this field.

Striking similarities in behavior between crystals and noncrystalline solids are now plentiful; they cover electrical, optical and luminescent properties. Russian investigators, for instance, have demonstrated in amorphous solids two phenomena long known to occur in crystals: radiative recombination (the emission of light when electrons combine with electron-deficient sites) and thermally stimulated currents, or "current glow" (a transient electrical conduction process arising at certain temperatures from the release of "trapped" charge carriers). These observations signify much more than provocative parallels between two types of substance. They mean that at least some of the experimental procedures that have yielded sophisticated insights into the nature of semiconducting crystals can also be applied to matter in the amorphous state. Of course, some processes observed in crystalline semiconductors cannot be reproduced in amorphous materials because the latter contain electrons traveling in mean free paths that are much shorter than those found in, say, germanium or silicon. Many other crystals suffer from the same limitation, however, and as a result the differences between the two kinds of solids have increasingly come to be regarded as differences in degree.

To some of those engaged in unraveling the electronic behavior of amorphous matter the present period bears a remarkable resemblance to an earlier one in the history of electronics: when the problem of detecting the feeble electrical signals of early radio transmissions was solved with the help of lead sulfide crystals. One took a promising-looking specimen of this natural semiconductor, galena, and gripped it tightly in the jaws of an alligator clip. A thin crimped wire, affectionately known as a cat's whisker, was then fixed to some insulated part of the radio receiver and allowed to touch the crystal lightly. The result was a diode detector, that is, a circuit element that in effect is rapidly switched on and off by an oscillating signal such as a radio wave. It was a device of monumental unreliability that nevertheless sustained radio for many years. In its day it was a mighty invention, a practical switch that could outperform the fastest mechanical commutator by many orders of magnitude. To be sure, its operating



GLASSES USEFUL AS SEMICONDUCTORS typically contain from two to four elements. The most promising materials are those known as chalcogenide glasses, indicating they contain at least one element from group VIa of the periodic table: oxygen, sulfur, selenium and tellurium. The tetrahedron represents the chalcogenide system consisting of germanium, sulfur, tellurium and arsenic. Every point not at an apex represents some mixture consisting of two or more elements. Points on an edge consist of two elements; points on a face consist of three elements, and points inside the solid consist of all four elements. On each face the region shown in gray indicates the composition of ternary mixtures that will form glasses when they solidify. Mixtures falling inside the solid have not yet been completely evaluated. Near the boundary of the gray regions a mixture may form a glass or not depending on various factors such as how fast the material is chilled. Thus thin films can often be made in glassy form, whereas a bulk sample of the same composition would solidify in crystalline form. The data used here were supplied by R. Fleming and E. A. Fagen.



OHM'S LAW AND DEPARTURES FROM IT can be illustrated by three kinds of device: a bulk semiconductor (*a*), a crystal rectifier (*b*) and an "Ovonic" threshold switch (*c*). Each produces a distinctive relation when applied voltage is plotted against the flow of current. Only the first device behaves according to Ohm's law, which states that current flow is directly proportional to applied voltage. One form of crystal rectifier was the crystal-and-cat's-whisker device used as a detector in the early days of radio. The voltage-current curve is continuous, and its precise shape beyond the voltage turnover point depends on the time variation of the applied signal and on the nature of the load in series with the device. In contrast the curve produced by the amorphous threshold switch is normally symmetrical and discontinuous. When the voltage reaches a certain threshold, the resistance drops suddenly. The switch represented here consists of two graphite electrodes coated with a chalcogenide glass (*color*). The switch can also be made with thin-film electrodes as shown in the diagrams on page 34 and in the photomicrograph of an array of Ovonic switches on page 37.

principle remained obscure for decades, so that systematic development was long out of the question, but it provided a glimpse of the potential that was ultimately fulfilled by the transistor and other solid-state devices.

There is another debt we owe to the crystal detector of early radio: its discovery shattered the unquestioning faith of the scientific community in Ohm's law. For many decades this law, which states that current flow is proportional to the applied voltage, had been taught and accepted as a simple, immutable truth. The crystal diode showed how extremely nonlinear this relation could actually be. The refutation of Ohm's law as a general principle marked the end of an era and saved the science of electricity from getting stuck in one of its blindest alleys.

The electrical behavior of the new amorphous devices departs even more radically than the crystal diode from the behavior described by Ohm's law [*see illustration above*]. This has led to a new attitude toward amorphous materials. Today amorphous threshold and memory devices threaten the accepted view that the job of switching electronic signals is best performed by well-ordered semiconducting crystals containing carefully measured traces of impurities.

Work on threshold and memory switches in noncrystalline structures (originally amorphous oxide films with electrolyte contacts) was pioneered by Ovshinsky in 1958, at a time when this field attracted little interest among scientists at large. Two years later Ovshinsky established a company, Energy Conversion Devices, Inc., which has since been engaged in the systematic development of amorphous switching devices and of control, information-storage and printing systems based on them. Ovshinsky demonstrated switching action based on ionic processes in amorphous solids in 1959; in those early devices switching was sensitive to the polarity of the applied voltage. Amorphous threshold and memory switches in which the operation is insensitive to polarity are covered in patents filed by Ovshinsky beginning in June, 1961. Late in the same year Jacob F. Dewald, W. R. Northover and A. D. Pearson of the Bell Telephone Laboratories applied for a patent relating to negative-resistance devices and memory devices also based on amorphous materials; the reset characteristics of the devices they described were sensitive to polarity.

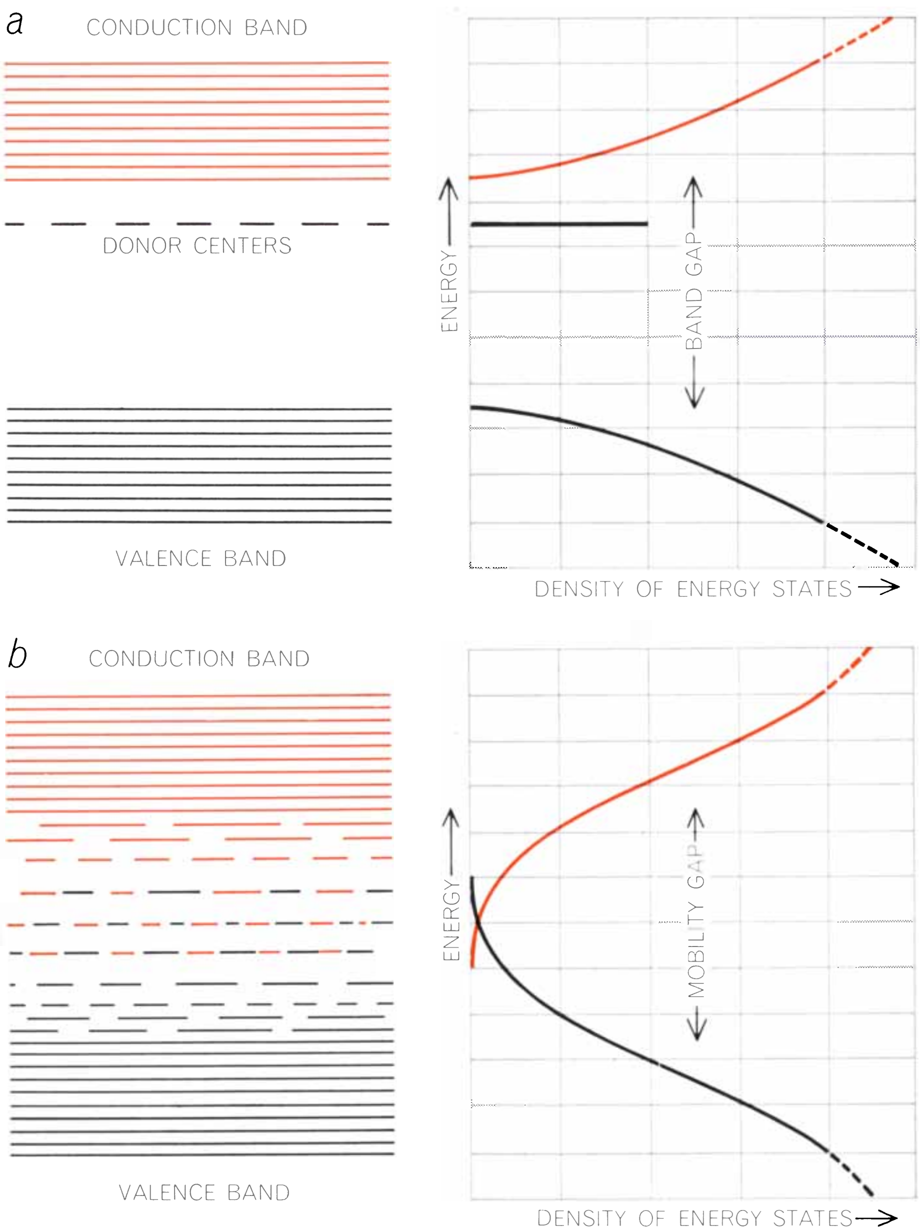
Among the many types of noncrystalline materials since investigated are amorphous oxides (including oxides of vanadium, tungsten, phosphorus, germanium and silicon) and chalcogenide glasses, which can be regarded as inorganic polymers. The term chalcogenic is applied to any of the elements in Group VIa of the periodic table: oxygen, sulfur, selenium and tellurium. The chalcogenide glasses include binary systems (for example germanium-tellurium), ternary systems (various three-component mixtures of germanium, arsenic, tellurium, silicon, selenium, zinc and cadmium) and quaternary systems composed of the same elements. One quaternary system of particular interest consists of germanium, sulfur, tellurium and arsenic [*see illustration on preceding page*]. Made by melting such components in an evacuated quartz tube, the chalcogenide glasses lend themselves to a variety of subsequent processing techniques: they can be cast, extruded, rolled, hot-pressed, blown, painted, deposited by evaporation in a vacuum or deposited by sputtering.

The essential properties of Ovonic devices do not depend critically on the fabrication method; moreover, they are remarkably insensitive to accidental trace impurities. Obviously the processes for manufacturing them must be carefully optimized to yield devices with precisely repeatable characteristics and long operating life, but the ultraclean regimes established for the fabrication of single-crystal semiconductors do not seem to be necessary.

The quaternary chalcogenide glasses should contain a large number of unsaturated chemical bonds capable of engaging and satisfying the valence requirements of any impurity. This factor and the possibility of internal rearrangements without lattice constraints are believed to preclude the impurities from acting as donors or acceptors of electrons, as they would in crystalline semiconductors. In this sense, at least, such glasses must be regarded as almost intrinsic semiconductors. This implies that negative charge carriers (electrons) and positive charge carriers ("holes") are present in roughly equal numbers, but it does not necessarily mean that the two types of carrier are equally mobile. There are other observations that support this view; chalcogenide glasses exhibit reasonably distinct wavelength limits in their ability to absorb light, and their electrical conductivity varies exponentially with temperature over a wide temperature range. In two-terminal assemblies, before threshold is reached the current also varies nonlinearly with voltage. When plotted in normalized form, all the voltage-current data fall on the same curve, regardless of temperature [*see bottom illustration on page 36*].

The various observations suggest two general conclusions. First, the electrons in such amorphous materials can be said to have energy relations (a "band structure") somewhat like those prevailing in crystals, in spite of the absence of long-range order. Second, the prevailing disorder leads to the existence of additional energy levels, probably discontinuous, that destroy the sharpness of the band edges and indeed lead to a certain amount of band overlap. Such levels also exist in crystalline semiconductors, but normally they are not so numerous or so widely spread out [*see "a" in illustration below*]. For the amorphous materials used in threshold and memory switches these additional energy levels appear to be of critical importance.

There is a widespread belief that insulators are insulators because they suffer from a shortage of free charge carriers (such as electrons), but this is only part of the explanation, and sometimes it is not the most significant part. If it were the whole truth, extra electrons entering through the contacting electrodes would be able to support conduction processes to a much greater extent than is actually observed. The fact is that insulating solids have localities where any extra charge carriers that may come from the contacts are trapped and rendered immobile. Acting together, such carriers form a space charge that tends by electrostatic repulsion to prevent additional carriers from entering; thus the space charge gives rise to a potential barrier under the contact. In this way conduction through the bulk of an insulator is effectively inhibited, no matter what kind of electrodes are applied. To



BAND MODELS are symbolic representations of how electrons behave in solids. Electrons are confined to the regions designated by the closely spaced parallel lines, which represent energy states. In a crystalline *n*-type (negative charge carrier) semiconductor (*a*) there is a sharply defined energy gap in which there is one discontinuous level corresponding to an impurity that can donate electrons. In amorphous solids the bands are smeared out, and in many (*b*) there is believed to be an overlap where the levels are discontinuous. There is no "band gap" in the normal sense; instead there is a "mobility gap" where electron movement is difficult but still possible. The mobility-gap model is based on work by Morrel H. Cohen, Hellmut Fritzsche, Stanford R. Ovshinsky and independently by Sir Nevill Mott.

be realistic one must say that insulators are insulators in large part because they contain a high concentration of trapping centers. In some cases the nature of these centers is well understood; more often their character (whether it is due to impurities, to disorder or to other factors) remains to be identified. Their functional presence, however, cannot be in doubt. The energy levels that are occupied by trapped charge carriers are in fact the additional discontinuous levels referred to above.

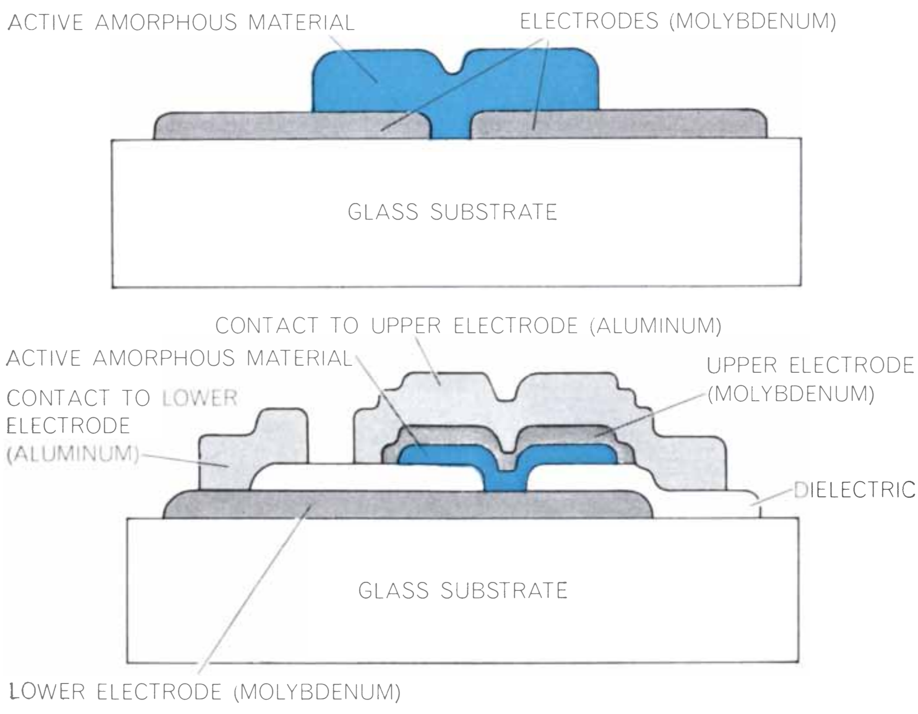
What would have to be done if one wanted to make a given insulator into a really good conductor without changing it chemically? One possibility would be to supply it somehow with more free charge carriers. This is what happens when an insulator is illuminated to make it photoconductive. The process calls for radiation that is strongly absorbed and that is accordingly unable to penetrate very far into the solid. Moreover, except in a few substances the survival rate (or lifetime, as it is called) of light-generated carriers in the free state is low. As a result the effects of radiation are somewhat limited and confined to regions close to the surface.

There is, however, another way to convert an insulator into a conductor. It depends on the fact that a substance in which *all* charge traps are filled would allow any new incoming charge carriers to take part in conduction processes, just as if no traps were present at all. This is so because an electron trap can accommodate only one electron, and it ceases to be a trap once it has captured its quarry. Of course, to make conduction possible the substance would also have to be free of the potential barriers created by space charges. In other words, it would have to be electrically neutral everywhere, or very nearly so. If traps are present, this calls for negative and positive charge carriers to be trapped in equal concentrations. The surprising conclusion is that a material in this state would behave something like a metal.

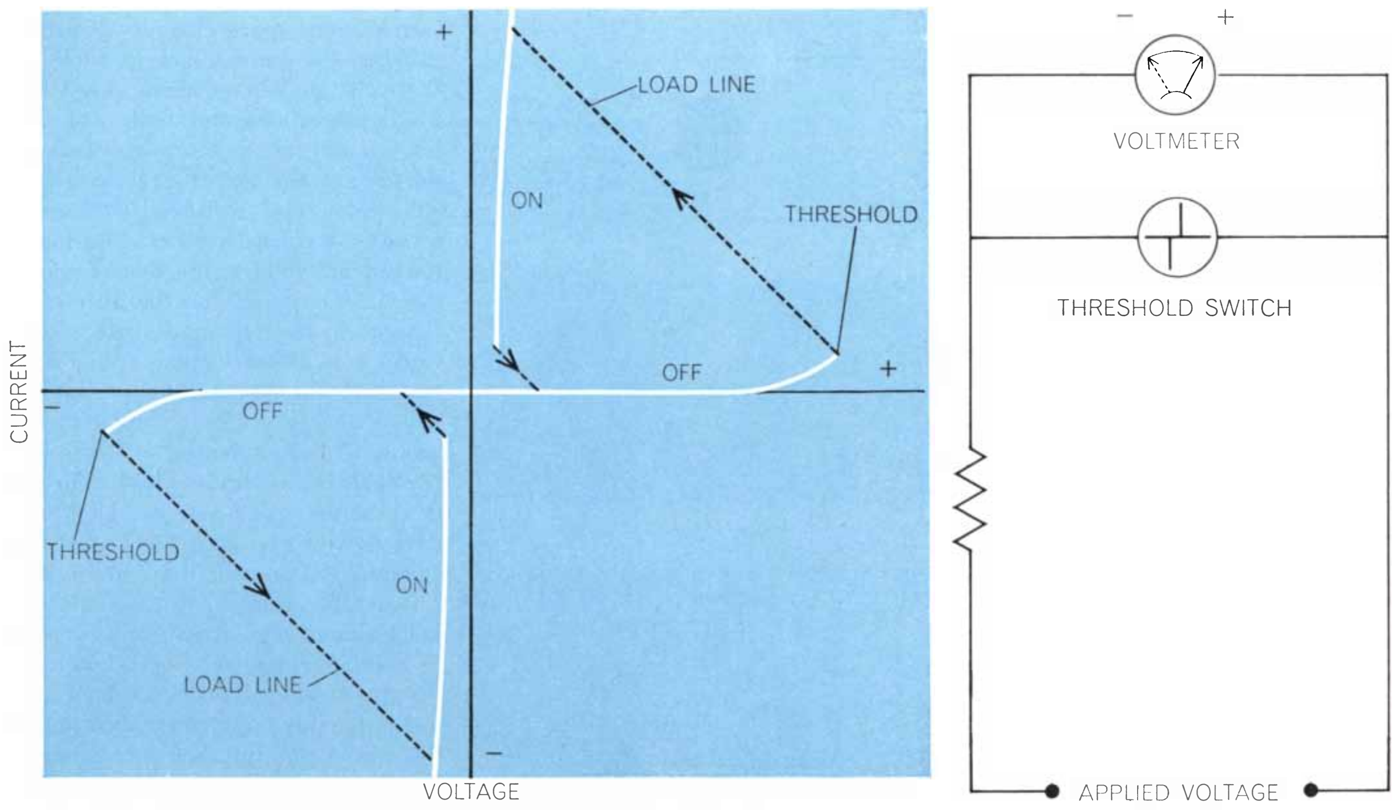
Does this mean that any insulator can readily be changed into a quasi-metal and back again? Not at all. A normal crystalline insulator could not satisfy the neutrality condition because its traps originate from impurities or defects in the crystal, or both. An exact balance between electron traps and hole traps would be entirely unlikely. Such a charge balance, however, is just the kind of situation envisioned in a recently developed model for the electronic band structure of amorphous semiconductors. It postulates the existence of two sets of discontinuous trapping states that constitute extensions of the conduction and valence bands and overlap near the center [*see "b" in illustration on preceding page*]. In this model the band gap found in crystalline semiconductors is replaced by a "mobility gap." The deeper discontinuous levels derived from the conduction band would serve as electron traps; those derived from the valence band, as hole traps. It can be shown that, in the very nature of these bands, the number of electron traps and hole traps formed in this way would be automatically equal, thus satisfying the conditions for exact charge compensation.

The next question is: How can the fully compensated state be brought about? This leads to a surprising number of possibilities. As a result switching processes that are superficially alike may arise from mechanisms that are essentially different, even within the framework of the model described. Depending on the nature of the material used (the choice is by no means confined to chalcogenide glasses), its thickness and its temperature, it is appropriate to ascribe varying degrees of importance to space charge, to the strength of the electric field, to temperature and to other effects. Each set of assumptions has its difficulties and each its optimum range of plausibility. Disappointing as this may be to the tidy mind in search of a unique theory, it adds variety and sophistication to the subject and should eventually broaden the inherent technical opportunities. Switching effects that are occasionally observed in crystalline materials, of course, demand different models and interpretations.

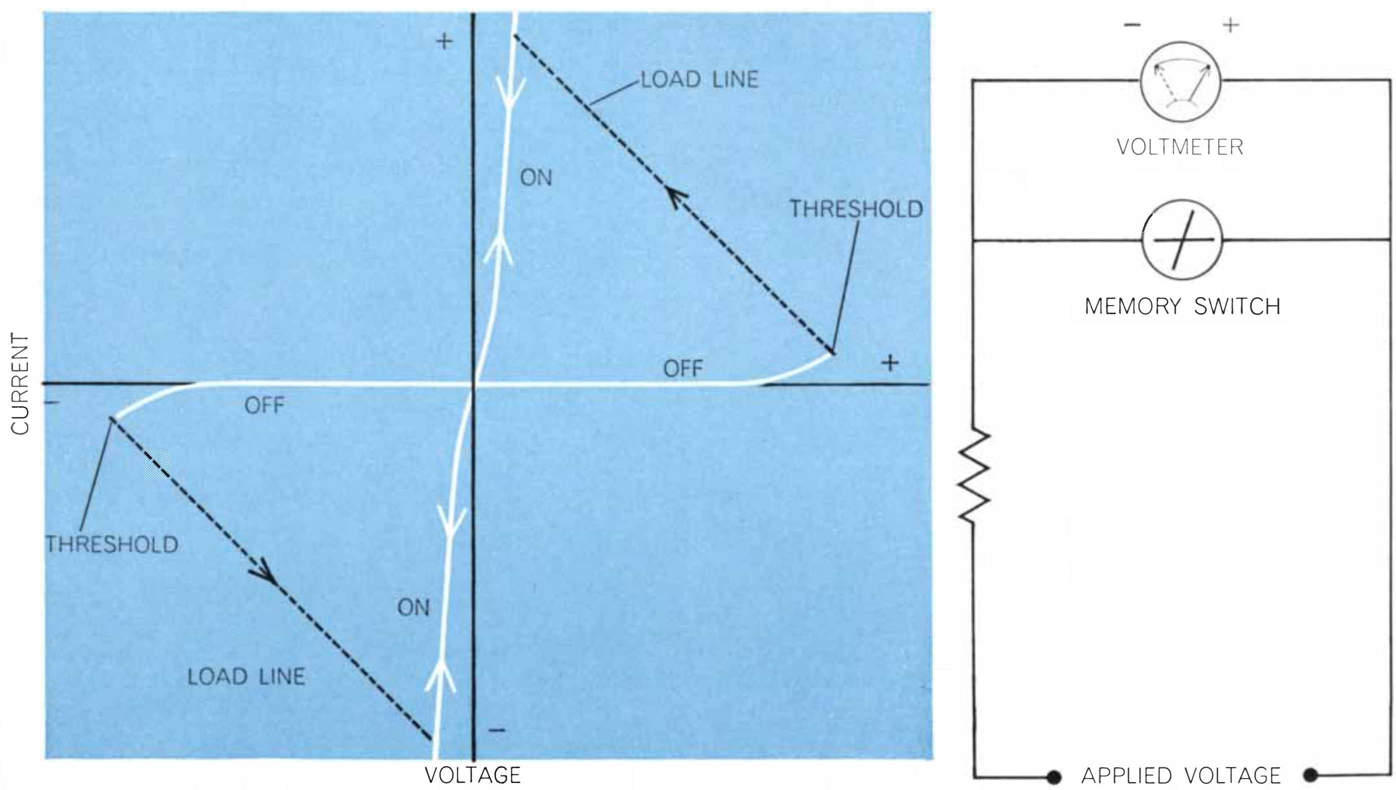
Now let us consider what happens when a current begins to flow through a two-terminal system that includes a moderately thick film of amorphous matter, which in bulk form would be a good insulator. At some distance from the contacts the current will be composed of nearly equal numbers of electrons and positive holes drifting in opposite directions, as is normal in an almost intrinsic semiconductor. Because of the high fields prevailing and the asymmetric energy relations at each contact, however, the current at the cathode will be composed preferentially of electrons and the current at the anode preferentially of holes. As a result of trapping, a negative space charge could build up near the cathode and a positive one near the anode, leading to a distortion of the internal field. It is not necessary to assume that a charge carrier, once it is trapped, will remain in captivity forever. Various mechanisms are at work that will tend either to free it or to neutralize its charge, but as long as these processes are less than completely effective under the prevailing influx of new



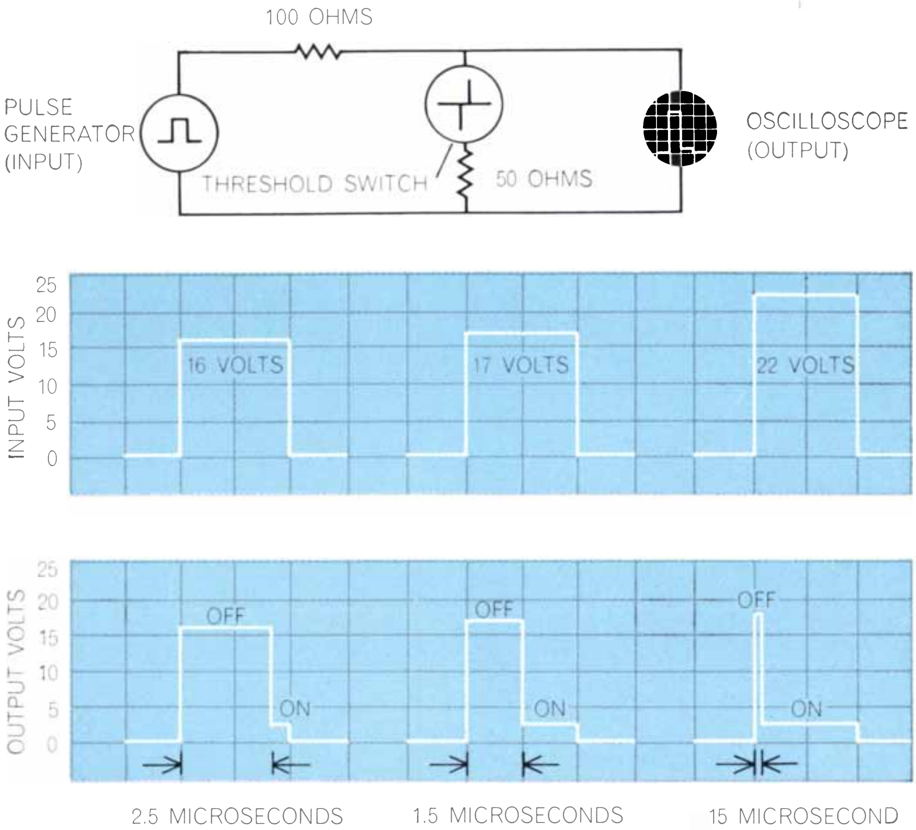
THIN-FILM OVONIC SWITCHES can take various forms. In one of the simplest (*top*) an amorphous semiconductor serves merely as a bridge between two molybdenum electrodes. In a more elaborate "pore" device (*bottom*) the semiconductor fills a pore that is sandwiched between two electrodes. These in turn are overlain by aluminum contacts. The devices, which were designed by R. E. Neale, are drawn here about 200 times actual size.



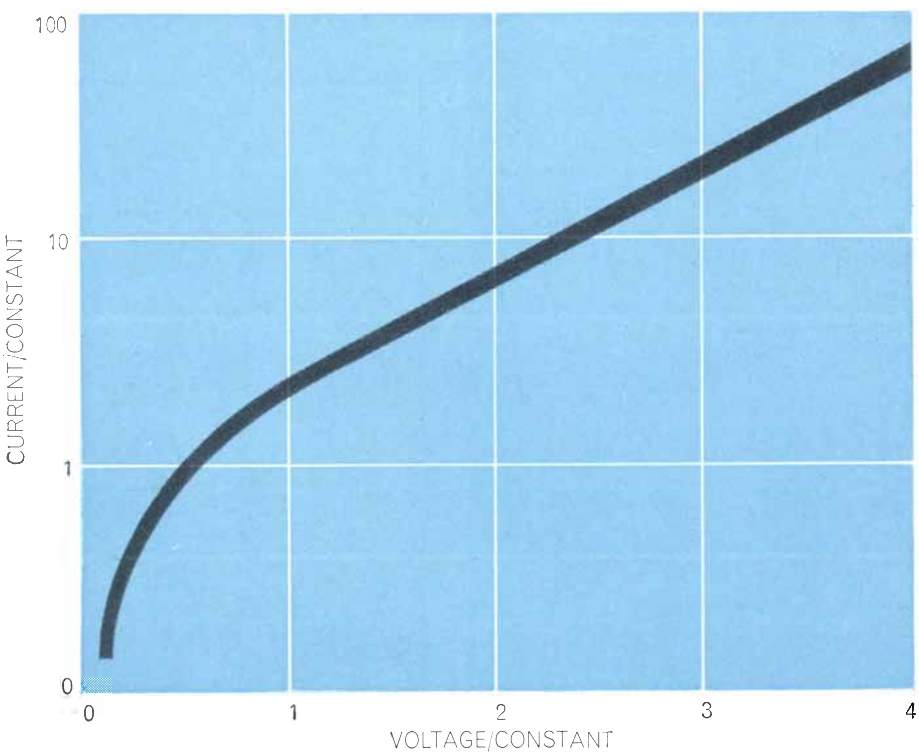
OVONIC THRESHOLD SWITCH exhibits a voltage-current relation with three well-defined regions. At first current flow is small and approximately proportional to voltage. As the threshold point is approached the current flow increases. When the threshold voltage is exceeded, the device switches to a holding region in which the voltage is virtually constant while the current can vary widely. This represents the "on" characteristic. After a brief delay the switching process itself takes place in less than a nanosecond (billionth of a second). Time intervals of this small magnitude make heating implausible as the principal operating mechanism. When the current after breakdown is reduced below a certain limit, the "off" characteristic of the threshold switch is quickly restored.



OVONIC MEMORY SWITCH is a version of the threshold switch in which the "on" state is stable. This state is produced by a high-voltage pulse that generates an appreciable amount of heat in the device and leads to a partial ordering of the amorphous material. This makes the material permanently conducting, and current can then pass through it in either direction. A different electrical treatment equivalent to melting, but not necessarily as drastic, can restore the disorder and return the switch to the "off" state. The "on" state of a memory switch involves a different conduction mechanism than that at work in the "on" state of a threshold switch.



SWITCHING DELAY in a typical amorphous threshold switch varies depending on the strength of the input pulse. If a rectangular pulse of 16 volts is applied to a typical 16-volt unit, it will switch to the "on" state in 2.5 microseconds. Raising the pulse to 17 volts shortens the delay, and raising it to 22 volts shortens it still further. After an "on" period a 16-volt switch requires about .5 microsecond to recover its normal "off" characteristics.



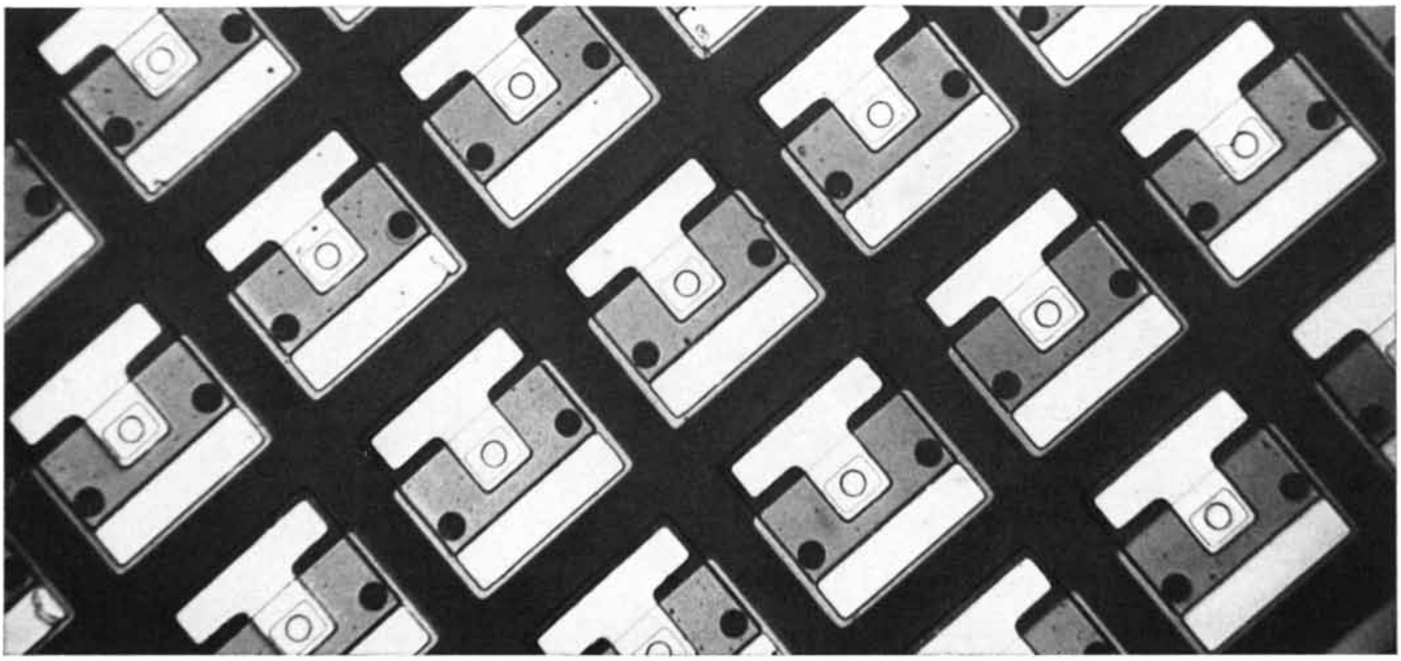
CONDUCTANCE OF AMORPHOUS THRESHOLD SWITCHES before switching occurs is related to voltage in a manner that depends on temperature. At any given temperature the flow of current increases almost exponentially over a considerable voltage range. The curve shown here has been normalized by dividing current and voltage by temperature-dependent constants. It covers measurements from −95 degrees Celsius to 176 degrees, suggesting that over this range, at least, a common mechanism underlies the current-voltage relation. The data were obtained by P. J. Walsh, R. Vogel and E. J. Evans of Picatinny Arsenal.

carriers some space charge will persist.

When the current flow is small, the space charges will remain localized close to the electrodes and both will soon reach quasi-steady-state conditions. As the current increases, the space charges will widen and will tend in time to spread toward each other. The higher the applied voltage, the shorter will be the time required for the two space charges to overlap in a small region. Once such an overlapping region is established, charge neutrality will prevail within it and the local conductivity will be high. This in turn means that the electric field will redistribute itself. The voltages across space-charge regions that have not yet overlapped will inevitably increase and so will the current flow. The situation is inherently unstable. The space-charge regions will rapidly come to overlap completely (or almost so), thereby making the entire volume neutral and highly conductive. According to this model, the threshold for switching can be viewed as the onset of overlap [*see top illustration on preceding page*].

The small conductance observed below the threshold voltage is found to be proportional to the contact area, as one would expect. After breakdown the situation is quite different. The observed conductances are then independent of the contact area, showing that the region carrying the current is small compared with the smallest practical electrode area (say half a ten-millionth of a square centimeter). This tiny current-carrying region may properly be termed a microplasma, by analogy with breakdown processes in silicon rectifiers. The shape of the voltage-current characteristic suggests that the diameter of the current-carrying channel is not actually constant but increases with increasing current. From a practical point of view these observations mean that the active area of a switching device should be small in order to achieve the maximum current change during switching. There is, however, a limit to how small this area can be made, because unwanted edge fields become difficult to avoid as the site of the electrode becomes smaller.

The low resistance observed after switching is sustained only as long as the current exceeds a certain minimum value. If it falls below this value (as it does at some stage when the applied voltage is withdrawn), the operating point returns to the prebrakedown high-resistance characteristic. This form of operation is typical for threshold switches. The point of return to the "off," or high-resistance, state is interpreted as the condition under which full trapping



ARRAY OF OVONIC THIN-FILM SWITCHES is enlarged 65 diameters in this photomicrograph. The complete array, prepared by photolithographic methods, contains 2,500 switches in an area one inch square. The active amorphous region, which does the actual switching, is the small circle at the base of the "T." Its diameter in this unit is 30 microns, but for certain uses it can be as small as five microns. The array is an experimental unit made by Energy Conversion Devices, Inc., a firm founded by Ovshinsky in 1960.

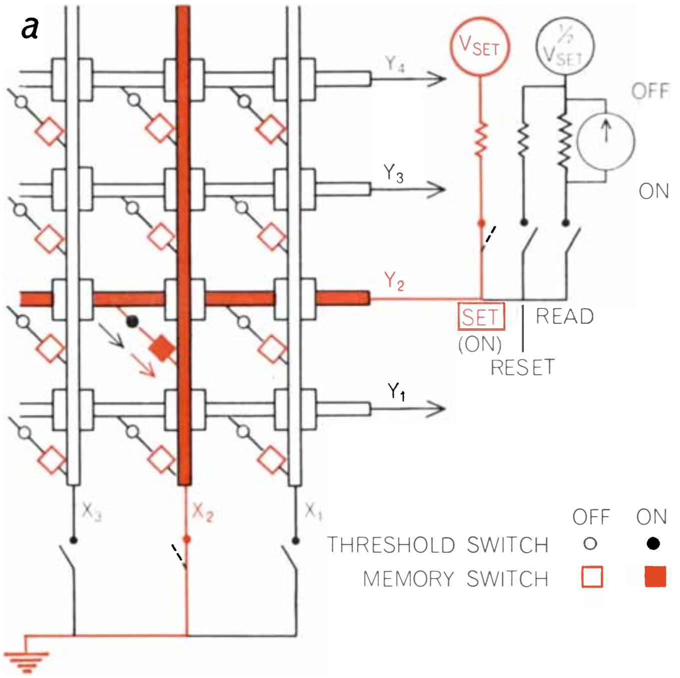
levels can no longer be maintained against the competing mechanisms that tend to empty them.

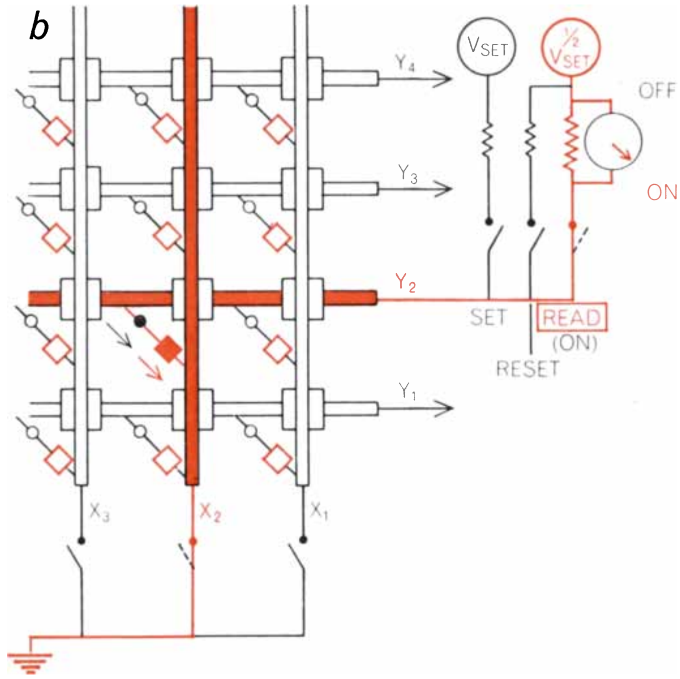
To what extent and under what practical conditions such a picture of the operation of the switching device can be sustained remains to be seen. In principle it accounts for the switching as such, and also for the time delay that is known to precede the switching process. By raising the applied voltage above the threshold value the time delay can be reduced from several microseconds to about a nanosecond (billionth of a second).

In accordance with this model of the switching operation the device would be highly polarized in the immediate prethreshold condition, that is, one side would have to be positively charged and the other negatively charged. On this point the observations available to date are not completely unambiguous. It is believed that some space-charge effects make themselves felt over a wide temperature range, but the conditions under which they exercise a controlling influence in device performance remain to be ascertained. Meanwhile alternative interpretations are certainly available that do not demand this high-degree of prethreshold polarization. One such model ascribes major importance to the time-dependent temperature profile between the electrodes. Close to the midpoint the temperature would be at a maximum and the resistivity at a minimum. The development of this profile would lead to a progressive redistribution of the electric field. This in turn would give rise to greatly enhanced high-field conduction through electrode barriers and thus eventually to switching. On this basis the breakdown itself would be electronic in character; the prebreakdown delay would be mainly thermal. Another model ascribes principal importance to high-field effects in the interior of the device and to the nonequilibrium carrier distributions to which such fields give rise. It accounts well for the apparent increase in conductance that is observed as the threshold point is approached. On this basis the prebreakdown delay would arise from the time required for charge carriers to gain energy from the field in the presence of traps.

All these models have certain stock ingredients in common, but others have been advanced that regard the switching process in entirely different terms. One such model ascribes switching to a heat-induced phase change that is propagated through the medium with the speed of sound. Another suggests that the amorphous medium undergoes a phase transformation, initiated when valence electrons are liberated by the high electric field. A third attributes switching to the formation of conducting filaments. Still another interprets switching as a purely thermal process, the simple outcome of local heat dissipation in a material with a highly temperature-dependent resistivity. Thermal switching is certainly possible but, as far as practical switching devices are concerned, it is an implausible mechanism because it would involve too much heating in too short a time. Of course, some heating is inevitably superimposed on other processes, and this makes quantitative checks on purely nonthermal theories less straightforward than one might wish. When threshold switches are operated by brief pulses of current (the most likely mode of operation), the detailed characteristics must necessarily depend on the subtle interplay of electronic and thermal mechanisms, each with its own set of time constants. In thin films (thinner than two or three microns, say) electronic processes should predominate; in thicker films, with smaller fields and poorer heat dissipation, slower thermal effects may tend to take over.

In certain amorphous systems switching results from the rearrangement of ions in the material, but these systems are fairly slow-acting as well as polar in character. One amorphous memory switch of this type consists of silicon oxide to which one electrode of gold and one of aluminum are attached. It is of considerable interest because it exhibits a negative resistance in the "on" condition. Ionic rearrangement is not the switching mechanism in chalcogenide glasses, but the migration of ions from the contacts is occasionally responsible for the destruction of devices, particularly when they are operated with direct





RANDOM-ACCESS MEMORY ARRAY combines amorphous threshold switches and memory switches. This illustration shows the operation at one crossover point, X_2Y_2, in a large array. At the outset all memory switches are in the "off" position. The "set" voltage (V_{set}) is applied to the crossover point X_2Y_2 by closing the switch at X_2 and the "set" switch leading to Y_2 (*a*). When the electrical signal reaches the crossover point, it fires the threshold switch, allowing enough voltage to reach the memory switch to turn it to the "on" state also. To "read," or sense, the state of the memory switch at X_2Y_2 one closes the switch at X_2 and the "read" switch leading to Y_2 (*b*). A voltage only half as large as the "set" voltage is now applied to the X_2Y_2 crossover, but it is enough to fire the

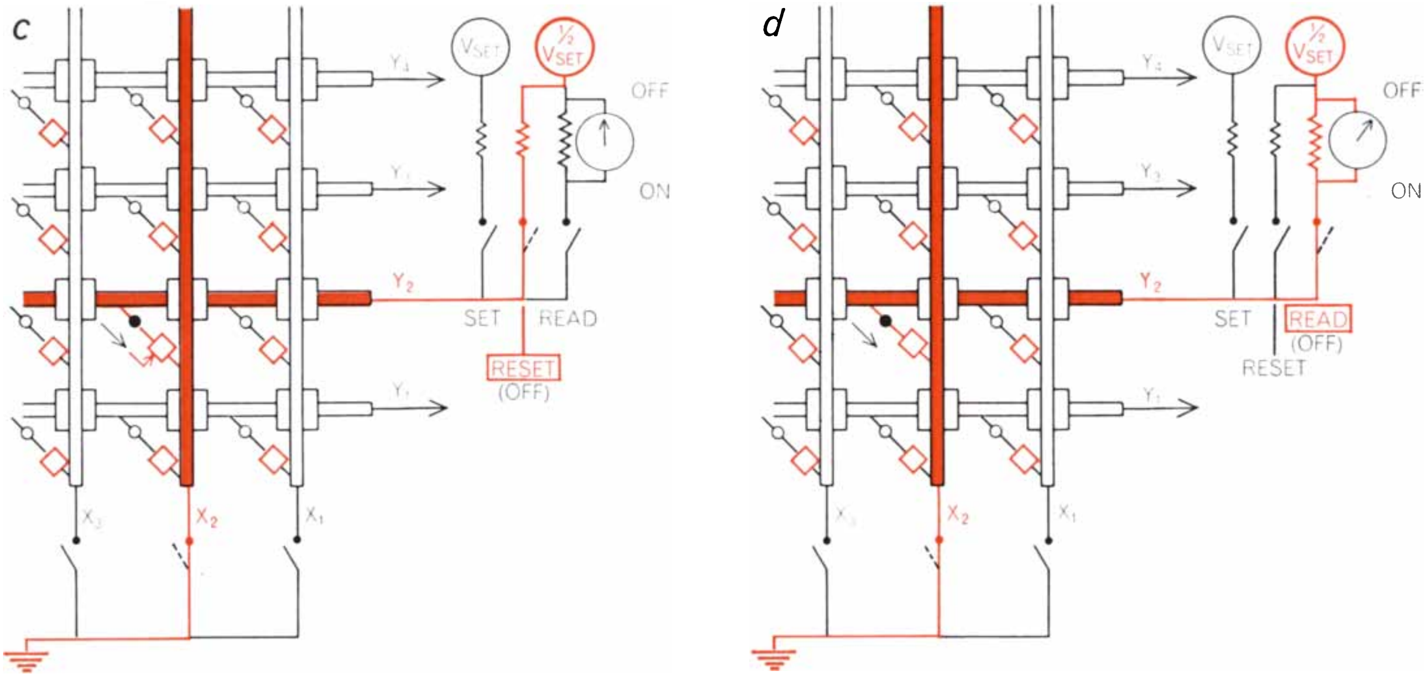
current. The success of a device therefore depends in large measure on a judicious choice of electrode materials. Chemically reactive electrodes and electrodes that diffuse easily through the glass must be avoided. When such electrodes are used, they lead to major instabilities, asymmetric operation and short life. There is also some danger of devitrification of the glass films under excessive power loads and, in some cases, of unwanted phase separation, a process that can take place in most of, and perhaps all, the chalcogenide glasses typically employed. This deterioration can be minimized or avoided, however, and the basic switching process is believed to be entirely nondestructive. Amorphous threshold switches have been known to withstand more than 10^{13} operations under low-power pulse tests without any change of characteristics. When premature failure does occur, it arises most frequently from misuse or from some aspect of faulty preparation.

Amorphous memory switches are similar in design and construction to amorphous threshold switches. The principal difference is that the low resistance achieved in the course of threshold breakdown remains low, even after withdrawal of the current [*see bottom illustration on page 35*]. Memory switches are thus bistable devices. They are switched "on" by the power dissipated in the course of threshold breakdown and during a short period afterward, lasting some two milliseconds. The breakdown is achieved by the application of an extra-high-voltage pulse. Here there is no doubt that the energy dissipated leads to a thermal transformation from a highly disordered phase that is almost nonconducting to a relatively ordered phase that conducts well.

Evidence in support of this picture has been obtained by differential thermal analysis of the amorphous materials in bulk. Time is needed for the phase transformation, which explains why the power must be kept on for a while after breakdown, to complete the process that the breakdown itself can only initiate. The total power that has to be dissipated is generally of the order of one to 10 microjoules. Once the switch is in its "on" state it can be removed from the circuit with impunity; its low resistance is maintained. Why this resistance should be quite as low as it is observed to be in practice is still a problem, but not one that is linked specifically to the switching phenomenon, since it has its origin essentially in the bulk characteristics of the material. One suggestion is that the switching action observed in the new glass memory switches may resemble the rotation of magnetic domains that occurs in ferroelectric materials.

To regain the high-resistance state a second power pulse is dissipated in the specimen. At first this seems absurd, but the paradox is only apparent. Because the process begins when the specimen is in its low-resistance state, the reset ("off") power treatment takes the form of a high-current pulse. The total power involved is of the same order as the power used during "on" switching, but it is dissipated in a much smaller volume and stops abruptly after about a microsecond. The localized temperature rise is therefore considerably higher than the rise at which order was previously established. Under these conditions the disorder is restored, a fact that can also be demonstrated by measurements on bulk specimens. The rapid power withdrawal causes the region to be thermally quenched. In the course of quenching, the formation of the conducting (ordered) phase is bypassed, because there is not enough time for the conversion. Exactly what degree of softening the glass undergoes before quenching is still uncertain. To perform as a memory the switch has to be made from a material that is close to the edge of the region where quaternary chalcogenide mixtures form glasses.

By combining monostable and bistable switches a variety of logic, control and shaping circuits can be designed, some involving single units and others



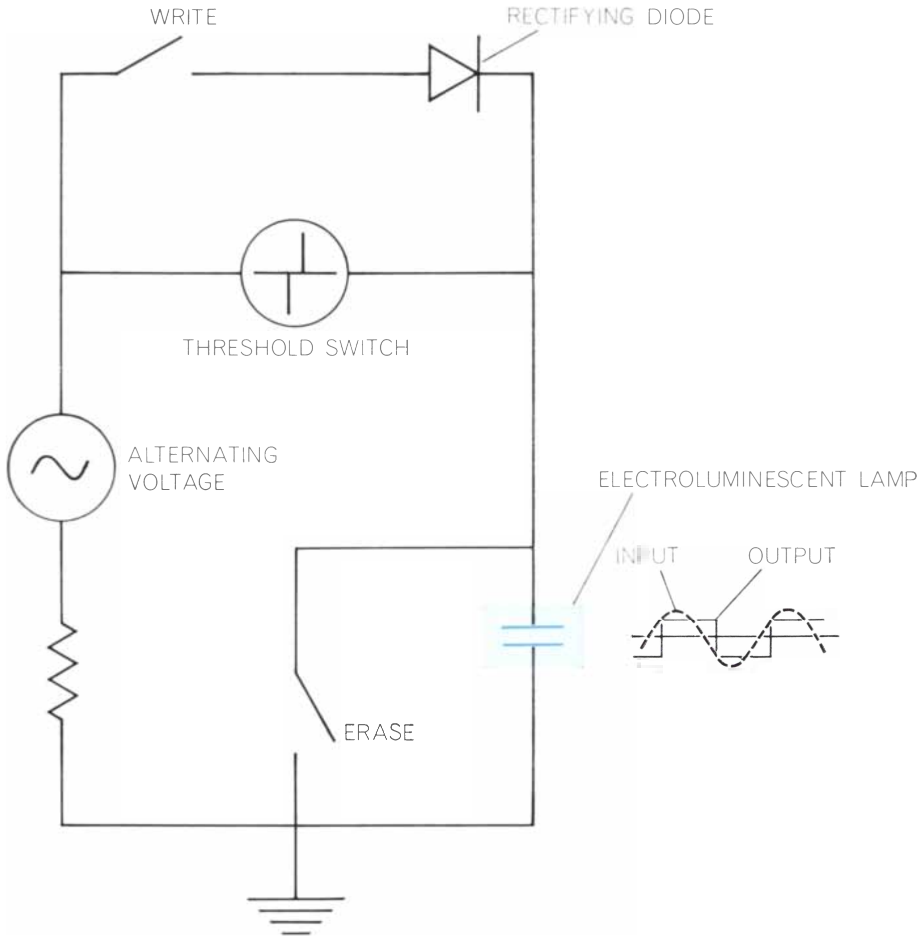
threshold switch. The adjacent memory switch, being "on," allows a substantial current to reach the readout detector, which responds by indicating "on." The memory switch at X_2Y_2 can be reset to "off" by closing the X_2 switch in combination with the "reset" switch leading to Y_2 (c). The same voltage as that required for reading again fires the threshold switch, but since the resistance in the "reset" line is lower than in the "read" line, a high-current pulse reaches the memory switch and returns it to the "off" state. The next time the X_2Y_2 position is read (d), the memory switch, being "off," allows much less current to reach the readout detector than it had when it was in the "on" state; thus the detector reads "off." The memory array was designed by Neale and D. Nelson.

multiple ones. The devices can also be used directly as surge-limiters for the protection of other circuits. Because the switches have temperature-dependent threshold voltages they can be used as temperature-control sensors. Early units were mechanically assembled and packaged, but both threshold and memory switches can be made in the form of thin films, either in the "pore" configuration or in planar form [*see illustration on page 34*]. If desired, the switches can be deposited simultaneously in linear arrays or area matrices of several thousand elements. There is also the possibility of combining them with other solid-state components on integrated-circuit substrates. Small size, high operational speed, ease of manufacture and, in principle, very low cost should make these components attractive for computer applications and a wide range of electronic control systems. Moreover, the amorphous devices are extraordinarily resistant to radiation damage, perhaps because they are already highly disordered before exposure. Thus they are well suited for use in systems designed for service in space.

The illustration on these two pages depicts the operation of a random-access memory array based on amorphous threshold and memory devices. Its performance depends on the response of the threshold and memory switches connected in series at each crossover point. Any one of these pairs can be "addressed" by applying a voltage to the appropriate X and Y conductors. The system is designed so that the memory elements have a somewhat higher switching voltage (and a lower prebreakdown resistance) than the threshold elements. Before firing, any applied voltage is therefore shared unequally between the two components. Application of an increasing voltage in pulse form will cause the threshold element to fire first. The entire voltage will then appear across the memory element and will fire it if high enough. When the firing pulse is over, the memory element remains in its "on" state, but the threshold element reverts to its "off" state, thereby isolating the memory element from events at other crossover points. This sequence of events describes the "set" operation.

For the "read" operation it is necessary only to close the appropriate X and Y switches. This applies half of the voltage required for the "set" operation to the selected crossover point. The lower voltage is sufficient to fire the threshold element but insufficient to fire the memory element if it happens to be in the "off" state. In that case the current reaching the readout detector will be close to zero. If, however, the memory element is in the "on" state, a much larger current will flow through the readout detector. The "reset" operation requires less voltage than the "set" operation. It is necessary only to fire the threshold element (as in the "read" operation) and thereafter to dissipate sufficient energy in the memory element to return it to the "off" state. To ensure this the resistance in the "reset" line is kept considerably lower than that in the "read" line. For successful operation such arrays would have to be made to very close tolerances, and in the present state of the art this is easier for planar memory switches than for planar threshold switches. The problems involved do not appear to be unsolvable, however. In any event, arrays of this type could be made by combining memory switches with conventional p-n junctions (single-crystal components containing appropriate impurity distributions) in place of the amorphous threshold elements. Corresponding circuit modifications would of course have to be made to allow for the fact that the "on" and "off" states of p-n junctions are selected by applying voltages of opposite polarity.

Arrays are only one example of what can in principle be done with amorphous switching devices. Other logic circuits, address-decoders and array line drivers can be designed in forms that appear fully compatible with conventional integrated-circuit technologies. Many new



OPERATION OF ELECTROLUMINESCENT LAMP can be controlled by an amorphous threshold switch in a simple circuit. Such a lamp is, by the nature of its structure, also a capacitor, and it is shown as such in the circuit. The diagram indicates how the sine wave of a normal alternating voltage is converted to a square wave by this combination of components. Over a certain range of applied voltages the circuit can be shown to be bistable. This means that the lamp is permanently on or permanently off, depending on purely transient signals reaching it through the brief operation of the "write" and "erase" switches.

possibilities are thereby opened up for exploration. For instance, Ovonic threshold and memory switches are peculiarly advantageous for the control of electroluminescent light sources, because over a certain range of applied voltages the circuit is inherently bistable. It can easily be shown that the alternating voltage applied to the electroluminescent lamp can be switched from "on" to "off" and back again at will, by purely transient control impulses. In one state the element will emit light; in the other it will remain dark [*see illustration above*]. This operation arises from the fact that each electroluminescent element is also a capacitor that can be given a bias charge by one type of applied control signal or can be discharged by another.

It is evidently feasible to make such circuits and device combinations in the form of a flat panel for the display of pictorial information, with integrated electroluminescent and switching elements. Many such attempts have been made in the past, but all have suffered in one way or another from low light output and from what has become known as the Maltese-cross effect. This is a type of disturbance that destroys contrast. It arises because electroluminescent lamps ordinarily have a "soft" brightness-voltage characteristic: the brightness increases gradually as the voltage is increased. What one really wants is a lamp that remains completely dark unless the applied voltage exceeds a certain predetermined limit. When that happens, the lamp should be on with constant brightness. If threshold switches are used in conjunction with electroluminescent lamps, this form of operation can be achieved, and the duty cycle of each picture element in an array can be controlled externally. This means that each picture element of an array is in its "on" state not only during the instant when it is "addressed" by a control pulse but also for any length of time afterward until it is switched off by a second pulse. For the design of panels that are addressed by scanning this is a matter of great importance.

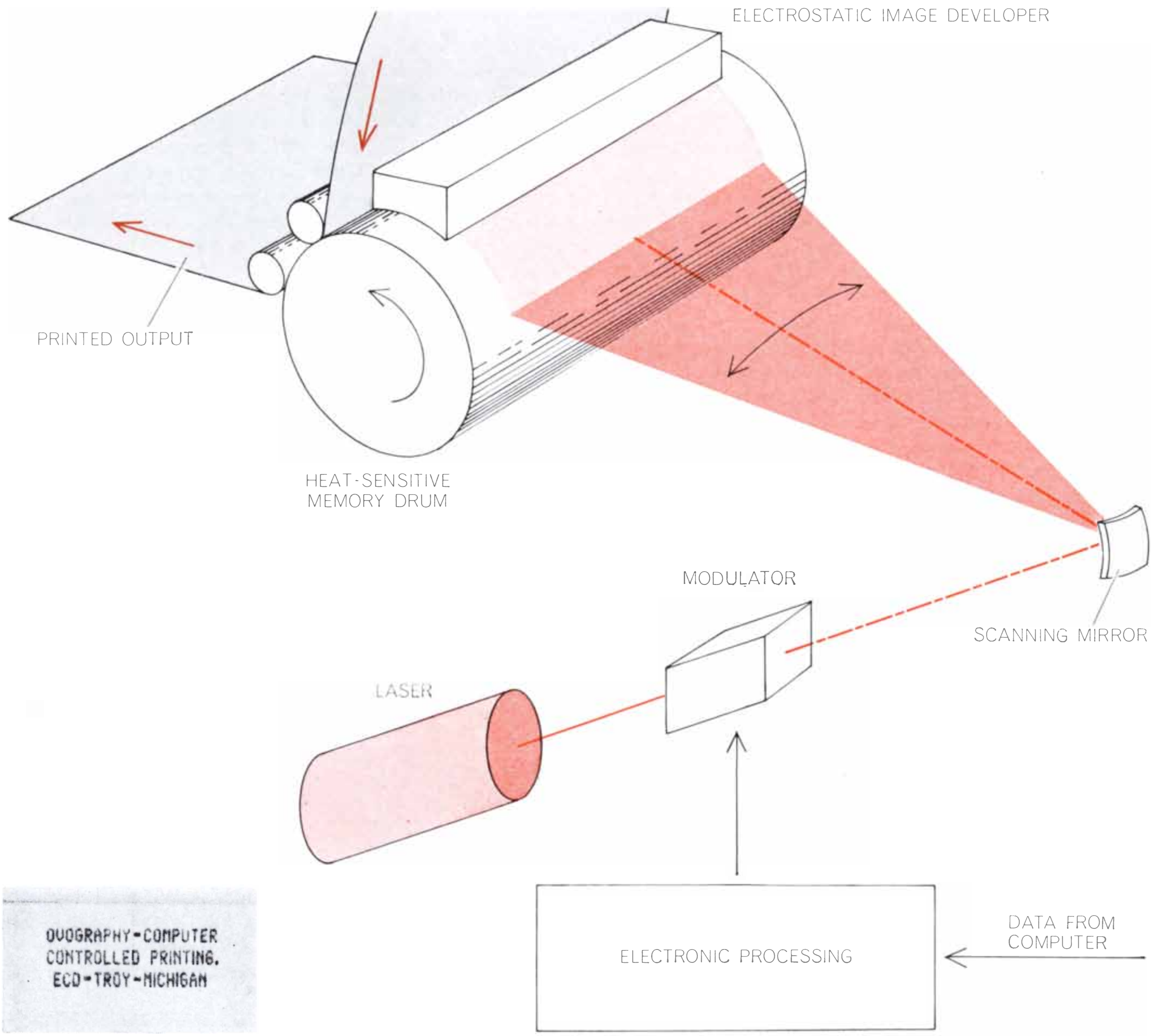
In order to achieve uniform brightness from element to element, it is more advantageous to energize the devices by square waves with a fast rise time than by sinusoidal waves. Moreover, because the light output of electroluminescent lamps increases with increasing frequency, the high-frequency components inherent in square-wave excitation have a useful function. As a result of all these factors a considerable increase of brightness and contrast can be achieved compared with expectations based on traditional panel techniques.

Of course, the development of such picture panels is not without its problems. One arises from the need to match the voltage requirements of lamps and switches. This calls for high-voltage switches and, at this stage, these can be made more easily and reliably by methods in which the active material is encapsulated in cartridges rather than by thin-film integrated-circuit techniques. There is no basic reason, however, why all-thin-film devices should not be made just as reliably as those with bulk electrodes. The flat-panel television screen based on electroluminescent picture elements controlled by amorphous switches is not yet with us, and a major development effort will still be needed to achieve it. On the other hand, a substantial step forward has been taken.

Another potentially important application of amorphous devices is in the field of printing. It is based on material of the memory type used in the form of a thin, ordered (and hence conducting) glass layer on a supporting plate. By suitable heat treatment the material can be made nonconducting on a highly localized basis. In the memory switches such changes are brought about electrically, but they can be made with any other heat source, for example by writing on the film with a light beam from a laser. In this way black-and-white pattern information can be stored in the layer, where it remains stable until it is erased by a second beam-controlled heat treatment. Erasure may be general, regional or local. Once a latent nonconducting image has been produced on the memory plate (which is in the form of a cylinder) printing proceeds in a manner similar to other electrostatic reproduction processes currently in use. The stored information is developed by exposing it to pigmented powders that adhere electrostatically to the nonconducting regions; the powders are subsequently transferred to paper.

The information to be stored is impressed on the memory layer by a laser beam whose intensity is controlled by the output of a computer. A laser of 1.5 watts continuous output power is required. This is still somewhat costly, but one can expect the price of lasers to drop enough in due course to make the equipment as a whole economically as well as technically attractive. Pictorial information could be handled by recording it with a high-resolution television camera and using the camera output to modulate the laser beam. The advantage of this printing process over present electrostatic systems is its ability to accept digital information directly from a computer (or analogue information from a television camera) without the necessity of re-forming the image after each revolution of the printing cylinder. The feasibility of such a printer has been demonstrated by Energy Conversion Devices; the process has been named Ovography.

It is clear that amorphous switching devices have a number of fascinating properties and important potential applications; it is also true that they present solid-state theorists with deeply challenging new problems. This year two scientific symposia have been devoted entirely to the behavior of electrons in amorphous materials; the first took place in New York City in May, the second was held at the University of Cambridge in September. At the Cambridge symposium papers were presented by speakers from 75 universities and 73 industrial laboratories in 21 countries. The meeting took place in an atmosphere of enthusiasm and excitement rarely seen on such occasions since the early days of the transistor.



OVOGRAPHIC PRINTING is an electrostatic system in which a laser beam under computer control "writes" on a drum coated with a glass that is normally a semiconductor but becomes nonconducting when heated. The glass is related to that used in amorphous memory switches. As in other electrostatic printing processes, pigmented powders adhere to the nonconducting areas on the drum (where the laser beam has written) but do not adhere to the conducting areas. The powders are then transferred to paper. A sample of type printed by Ovography appears at the lower left. The new printing system is being developed by Energy Conversion Devices.

EARLY MAN IN THE WEST INDIES

Until recently it seemed that the islands of the Caribbean were uninhabited up to about the time of Christ. Now it appears that men may have arrived 5,000 years earlier. How did they get there?

by José M. Cruxent and Irving Rouse

When and by whom the islands of the West Indies were first settled is a matter of debate among archaeologists. The debate may soon be intensified: evidence recently discovered suggests that people of the Paleo-Indian age, the earliest period in New World prehistory, reached one of the main islands 7,000 years ago. If this is so, it means that New World hunters managed to cross the Caribbean when their level of culture was no more advanced than that of the Paleolithic age in the Old World.

It is our intention in this article to present the new evidence and to relate it to older knowledge and conjecture about the aboriginal settling of the West Indies. Let us first briefly review the geography of the area, with special reference to the various island chains that provide natural migration routes outward from the mainland.

The West Indies consist of two major island chains and three minor ones. The southernmost major chain, the Lesser Antilles, consists of many small islands forming an arc along the eastern side of the Caribbean and separating it from the Atlantic. Starting near the mouth of the Orinoco River in Venezuela, the chain extends northward and then curves around to the northwest, so that its last members (the Virgin Islands) almost touch the first of the Greater Antilles.

The island next to the Virgin Islands is Puerto Rico. Starting there, the Greater Antilles extend westward along the north side of the Caribbean. The next islands in the chain are Hispaniola (politically divided between Haiti and the Dominican Republic) and Jamaica, which lies somewhat to the south. The chain ends with Cuba, whose western tip nearly fills the gap between the peninsulas of Yucatán and Florida at the mouth of the Gulf of Mexico. The Greater Antilles are much larger and better endowed with natural resources than the other islands of the West Indies, and today they are also much more populous.

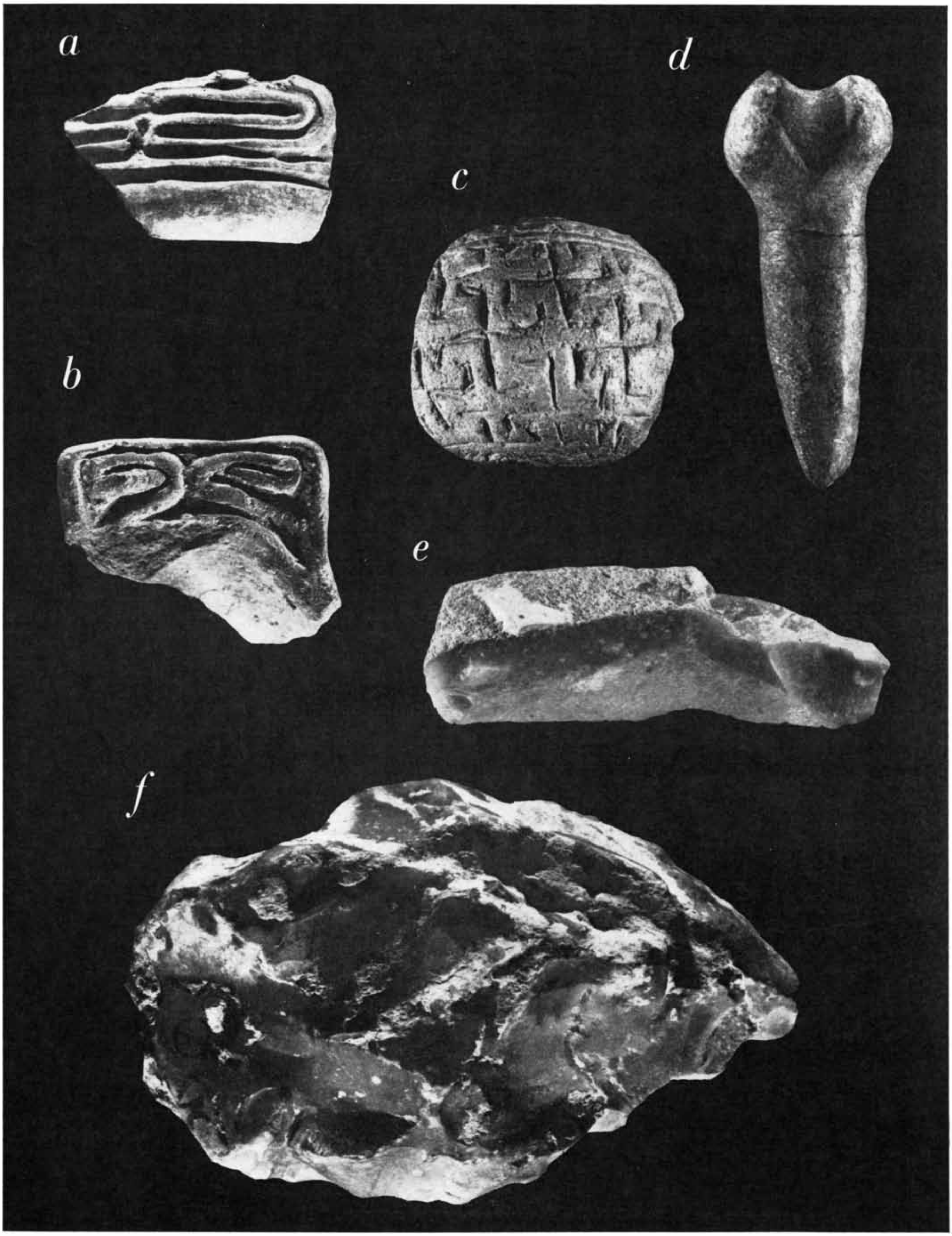
Of the three lesser chains the southernmost stretches along the coast of Venezuela from the mouth of the Orinoco on the east to Lake Maracaibo on the west. Its main islands are Trinidad, which is the easternmost, and two smaller clusters: the Venezuelan islands of Margarita, Coche and Cubagua in the middle and the Dutch islands of Bonaire, Curaçao and Aruba in the west. The chain does not have a generally accepted name. In Spanish it is called the Leeward Islands, but the same name is used in English for a group of former British possessions in the Lesser Antilles. Here we shall call the chain the South Caribbean islands.

The next of the lesser chains crosses the Caribbean from Central America almost to Jamaica, following a submerged mountain ridge. Today this chain consists mainly of banks, reefs and cays, but two of its islands (Providencia and San Andrés, which are part of Colombia) are of fair size. When the sea level was lower a few thousand years ago, the chain formed a nearly continuous series of stepping-stones leading to the Greater Antilles. This group also lacks a generally accepted name; here we shall call them the Mid-Caribbean islands.

The last of the lesser chains consists of the Bahamas, which lie to the north and east of Hispaniola and Cuba and extend northward for some distance along the east coast of Florida. A series of small coral islets, they provide the best stepping-stone route between the mainland to the north and eastern Cuba and Hispaniola to the south.

Prevailing winds and currents in the Caribbean strongly favor some directions of travel over others. The trade winds blow from the northeast, making voyages in a westerly direction easy in both the Greater Antilles and the South Caribbean chain. The currents favor movement to the west and also movement from South America north to the major island chains. The South Equatorial Current flows across the Atlantic to the coast of Guiana and is deflected northwestward through the South Caribbean chain and the southernmost Lesser Antilles. The North Equatorial Current follows a similar course through the northern Lesser Antilles to the Greater Antilles. Still farther north, the Canaries Current skirts the north coast of Hispaniola and flows through the Bahamas. The effect of the currents is reinforced by the flow of the two major rivers of northern South America: the Orinoco and the Magdalena. Both discharge into the sea with such force that debris is carried offshore for miles to be picked up by the South Equatorial Current and, in the case of the Orinoco, carried some distance into the Lesser Antilles.

Given these conditions, one would expect that the West Indies were populated from South America. This was true at the time of the first European voyages to America. When the Spaniards reached the area, they found that both the Lesser and the Greater Antilles were inhabited by Indians who spoke Cariban and Arawakan, languages that are widespread in eastern South America. The inhabitants' material culture belonged to the final pre-Columbian age, or period of development, in the Caribbean area; that age is known as the Neo-Indian. This means that the Caribs and Arawaks made pottery. Coincidentally they also



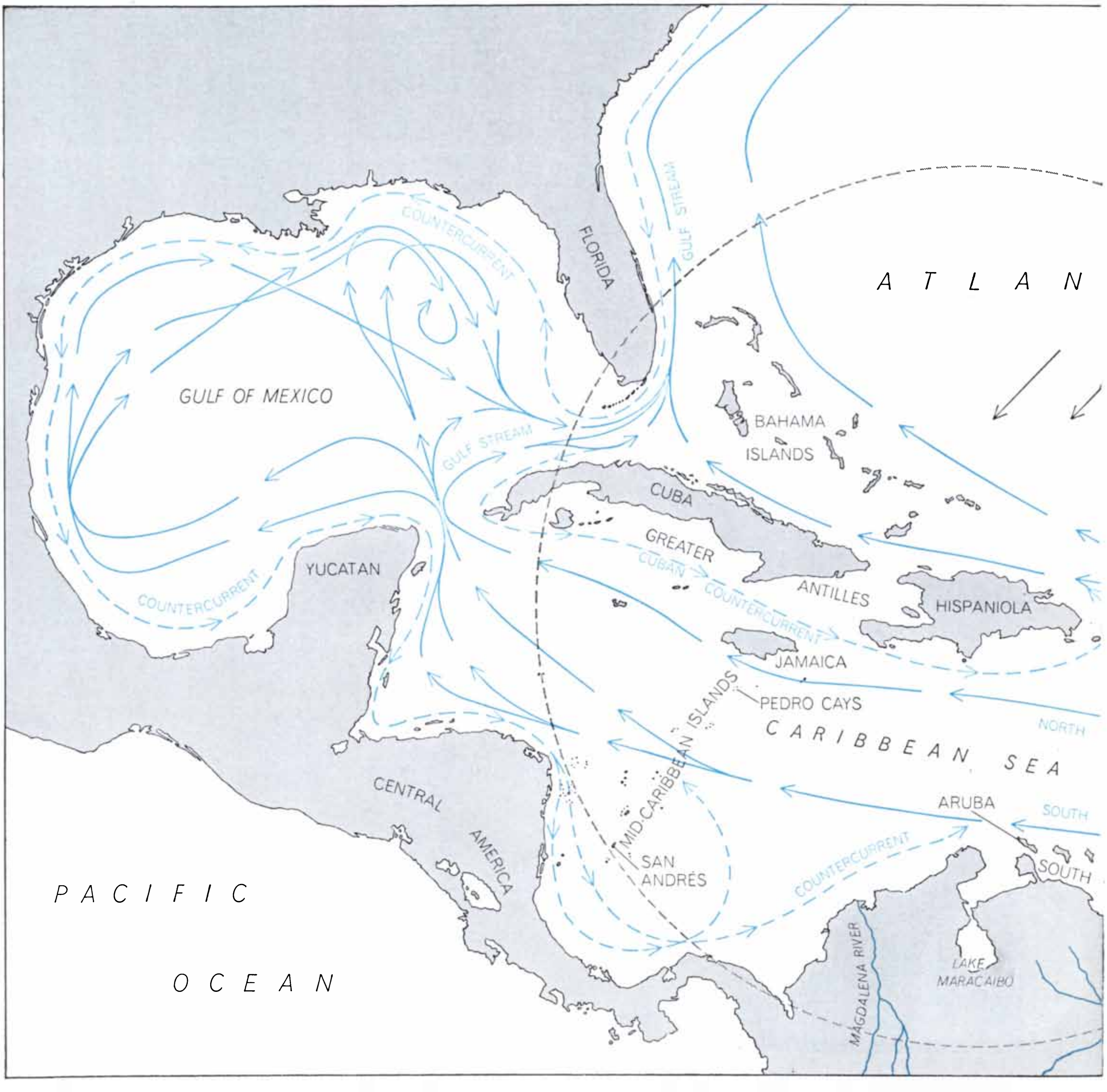
FOUR SUCCESSIVE AGES of prehistory on Hispaniola, a major island of the West Indies, are represented by these artifacts. They are (*a*, *b*) two fragments of Arawak pottery, typical of the most recent, Neo-Indian age; (*c*) an engraved stone bowl from Île à Vache and (*d*) a stone "pin" from a Couri culture site, both shaped by grinding and representative of the Meso-Indian age, whose people preceded and later coexisted with the Neo-Indians; (*e*) a flint projectile point from a Cabaret culture site, typical of the preceding Late Paleo-Indian age, and finally (*f*) a flint chopper of the more ancient Early Paleo-Indian age discovered at Rancho Casimira.

knew the art of farming and were skilled mariners, so that it is easy to understand their successful expansion through most of the West Indies.

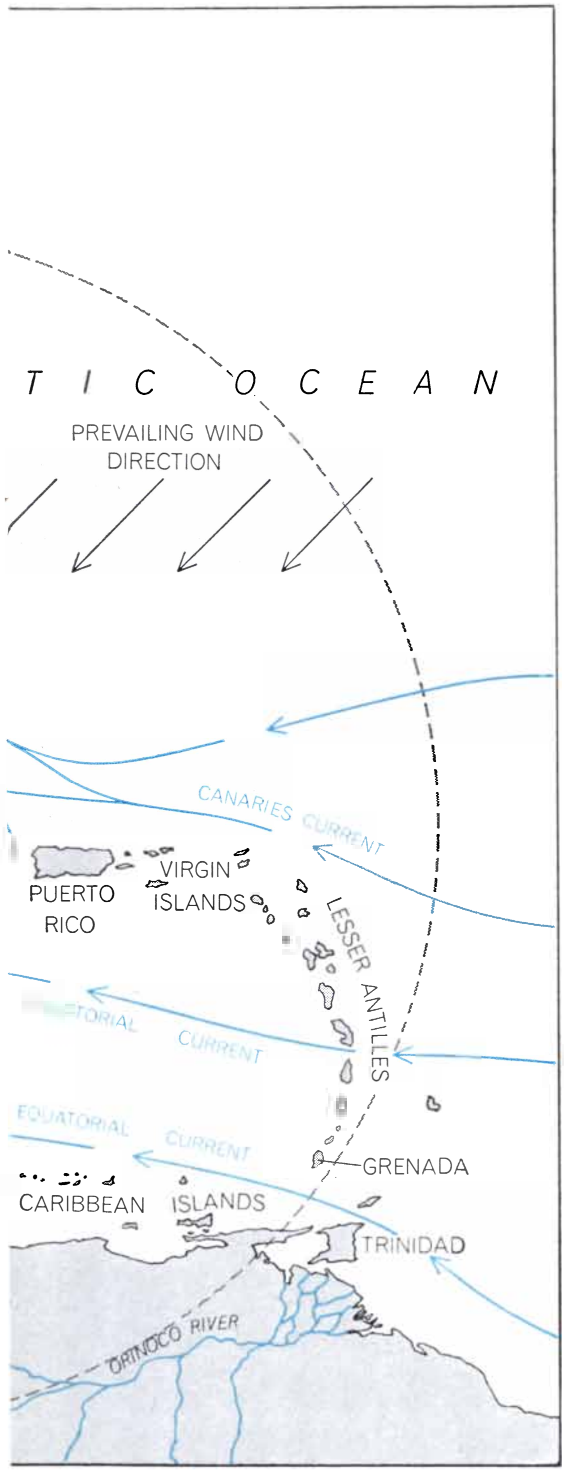
The Neo-Indians, however, were relative newcomers to the islands. We now know that they migrated from South America at about the time of Christ, entered the Greater Antilles about A.D. 250 and did not populate their most northerly territory, the Bahamas, until A.D. 1000. As they moved along they overran an earlier West Indian population. When the Spaniards arrived, the earlier people existed only as remnants in western Cuba, in a few small Cuban offshore islands and in southwestern Hispaniola.

The earlier inhabitants' material culture belonged to the preceding Meso-Indian age. They knew nothing of pottery; they made their distinctive artifacts by grinding stone and by chipping flakes of flint. They did not know farming and fed themselves instead by fishing and gathering shellfish and wild vegetable foods. After a few decades under Spanish administration the remnant Meso-Indians, together with nearly all the Neo-Indians, had become extinct.

Perhaps because the Meso-Indians



ISLAND CHAINS of the West Indies include the Greater Antilles (consisting of Cuba, Jamaica, Hispaniola and Puerto Rico), the Lesser Antilles (beginning with the Virgin Islands and ending with Grenada), the South Caribbean islands (from Trinidad, the most easterly, to Aruba, the most westerly), the Mid-Caribbean islands (extending from San Andrés to Pedro Cays) and the Bahama Islands (extending from Hispaniola to Florida). Hispaniola is geographically central (*broken circle*). The route to it from Venewere fishermen and presumably at home on the water the question of how and when they reached their island homes remained of limited interest to scholars until evidence of an even earlier West Indian population was discovered in recent years. These discoveries, both in the islands and on the adjacent mainland, now make it clear that, no matter how early the Meso-Indian occupation of the West Indies may have been, even earlier Paleo-Indian occupations preceded it. The sites—in Florida, Cuba and Hispaniola—were identified as Paleo-Indian by the absence of ground-stone artifacts; the only stone implements were made of flaked flint. How and when these primitive hunting peoples reached the islands are questions that have also stimulated increased interest in Meso-Indian origins.



zuela is by way of the Lesser Antilles, from Florida by way of the Bahamas, from Yucatán by way of Cuba and from Central America via the islands of the Mid-Caribbean chain.

Meso-Indian sites are found in both the Greater Antilles and the islands of the South Caribbean chain. They are not known anywhere in the Lesser Antilles, with the exception of a site on St. Thomas, the member of the Virgin Islands group that is closest to Puerto Rico. How, then, did the Meso-Indians reach the Greater Antilles? The Meso-Indians who settled in the South Caribbean chain must have come from the Venezuelan mainland. There is no convincing evidence, however, that the Meso-Indians who settled in the Greater Antilles had moved to their final destination through the Lesser Antilles, as the Caribs and Arawaks had. Nor is there any conclusive evidence that the Meso-Indians followed either of the alternative routes: from Florida by way of the Bahamas or from Central America by way of the Mid-Caribbean chain.

There is still another alternative. The Meso-Indians in the Greater Antilles may simply have been descendants of the Paleo-Indians, and their more advanced material culture may have evolved on the spot. If that is the case, there are two further possibilities. The ancestral Paleo-Indians may have reached the islands during their initial period of development: the Early Paleo-Indian age, when their artifacts consisted of simple, unhafted flint flakes and crude flint pounders and choppers. Or they may have arrived during the subsequent period: the Late Paleo-Indian age, when some artifacts (particularly projectile points, which were unknown in the Early Paleo-Indian age) were made in shapes suitable for hafting.

Hispaniola lies near the geographic center of the West Indies; it is nearly equidistant from the other three islands in its chain. It is also roughly equidistant from Venezuela and Colombia to the south, Central America and Yucatán to the west and Florida to the north. This pivotal position made Hispaniola preeminent in West Indian cultural development during the Neo-Indian age. Was it also preeminent in the preceding ages? As we shall see, the answer appears to be yes.

The first evidence that Hispaniola harbored any archaeological remains earlier than Neo-Indian ones was unearthed a century ago, but it was not until 1933 that an unquestioned Meso-Indian site was excavated. The digging was done for the Museum of the American Indian by Godfrey J. Olsen at Île à Vache, an island near the southwest tip of Haiti. The site was a large mound of discarded shells from which Olsen excavated single-bitted and double-bitted ground-stone axes, ground-stone mortars and pestles and a number of other objects. Some of the stones, evidently ceremonial, were engraved with elaborate designs. The designs are reminiscent of motifs seen on Hispaniolan pottery of the Neo-Indian age. It is quite possible, in fact, that the Île à Vache culture complex was contemporaneous with the Neo-Indian occupation of Hispaniola. The site is in a part of the island where Meso-Indians managed to survive until historical times.

Two years after Olsen's work five pre-pottery sites were excavated near Fort Liberté in northeastern Haiti by Froelich G. Rainey, then at Yale University's Peabody Museum, and one of us (Rouse). All five sites represented a single Meso-Indian complex that was given the name Couri. The Couri complex is clearly earlier than the Île à Vache one. Its ground-stone artifacts include vessels, milling stones, balls, pegs and double-bitted axes, but the decoration is unsophisticated [*see illustration on page 48*]. Only one stone vessel and a pendant made of shell bore simple rectilinear designs. The Couri people also made some tools out of flint, including large knives, scrapers and projectile points that have central stems for hafting. There are even some "backed blades," with one edge blunted by chipping, as in the Upper Paleolithic age of the Old World.

The Yale excavations stimulated Haitian archaeologists in their own search for pre-pottery sites. In 1940 Edward Mangonès of the Haitian Bureau of Ethnology found two deposits of flint tools in an area north of Port-au-Prince. Other investigators subsequently located six more sites in the same general area. All of them can apparently be grouped in the same complex, named Cabaret after the best-known of the eight localities. The absence of ground-stone artifacts means that the Cabaret complex belongs to one of the Paleo-Indian ages. Included among its flint artifacts are projectile points, showing that the Cabaret people were of the Late Paleo-Indian age; the points have stems for hafting. Among the other artifacts are long, narrow flakes of flint with ends that have been retouched [*see illustration on page 49*]. In their retouching the Cabaret workers took pains to reduce or eliminate the "bulbs of percussion" or the "striking platforms" that are present on the long flakes because of the way they are made. This presumably facilitated setting the flakes in hafts for use as scrapers or as knives.

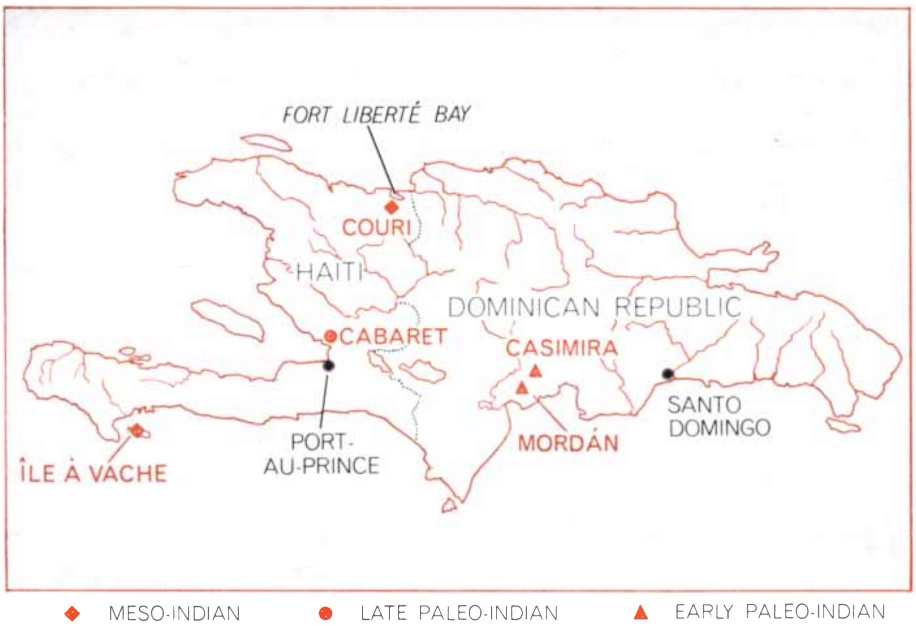
The most recently discovered pre-pottery sites on Hispaniola also appear to be the island's oldest. They are located some 50 miles west of Santo Domingo on the island's south coast and were excavated in 1963 by the Dominican archaeologist Luis Chanlatte and one of us (Cruxent). In the 1920's the Danish archaeologist Gudmund Hatt had collected flint artifacts there, near the village of Barreras, and had reported that the region was rich in flint. At Mordán, a village near Barreras, Chanlatte and Cruxent excavated the first of their sites, a thick deposit of prehistoric refuse, primarily marine shells that had been brought from the seashore less than a mile away.

A trench two meters wide and four meters long was dug in the Mordán refuse deposit to a depth of one meter. In the top 25 centimeters of the trench fragments of pottery were common, but they became scarcer in the next 25 centimeters and were absent in the two lowest 25-centimeter sections. The lowest sections contained a number of flint artifacts; no projectile points were among them. All four sections contained small fragments of animal bone; none of the bones could be identified by species, but some were fishbones. In the bottom section a hearth yielded charcoal samples for carbon-14 analysis.

For a distance of at least three miles in all directions from Mordán the ground is littered with flint boulders and angular flint fragments embedded in the soil and half-exposed by erosion. These provided the raw material from which the Mordán workers fashioned their tools. The long flakes were apparently struck from cores with prismatic cross sections; we say "apparently" because no cores were recovered. The flakes, however, show the bulbs of percussion and prepared striking platforms characteristic of flints struck from such cores.

The Mordán workers did not follow the Cabaret practice of retouching their flakes to facilitate hafting. The absence of hafting and of projectile points shows that the complex belongs to the Early Paleo-Indian age. A few of the flakes show crude retouching on the edges, but most of them were apparently used just as they came from the core. The only secondary flaking visible on them is due to wear. They seem to have served as knives and scrapers, possibly for the manufacture of baskets.

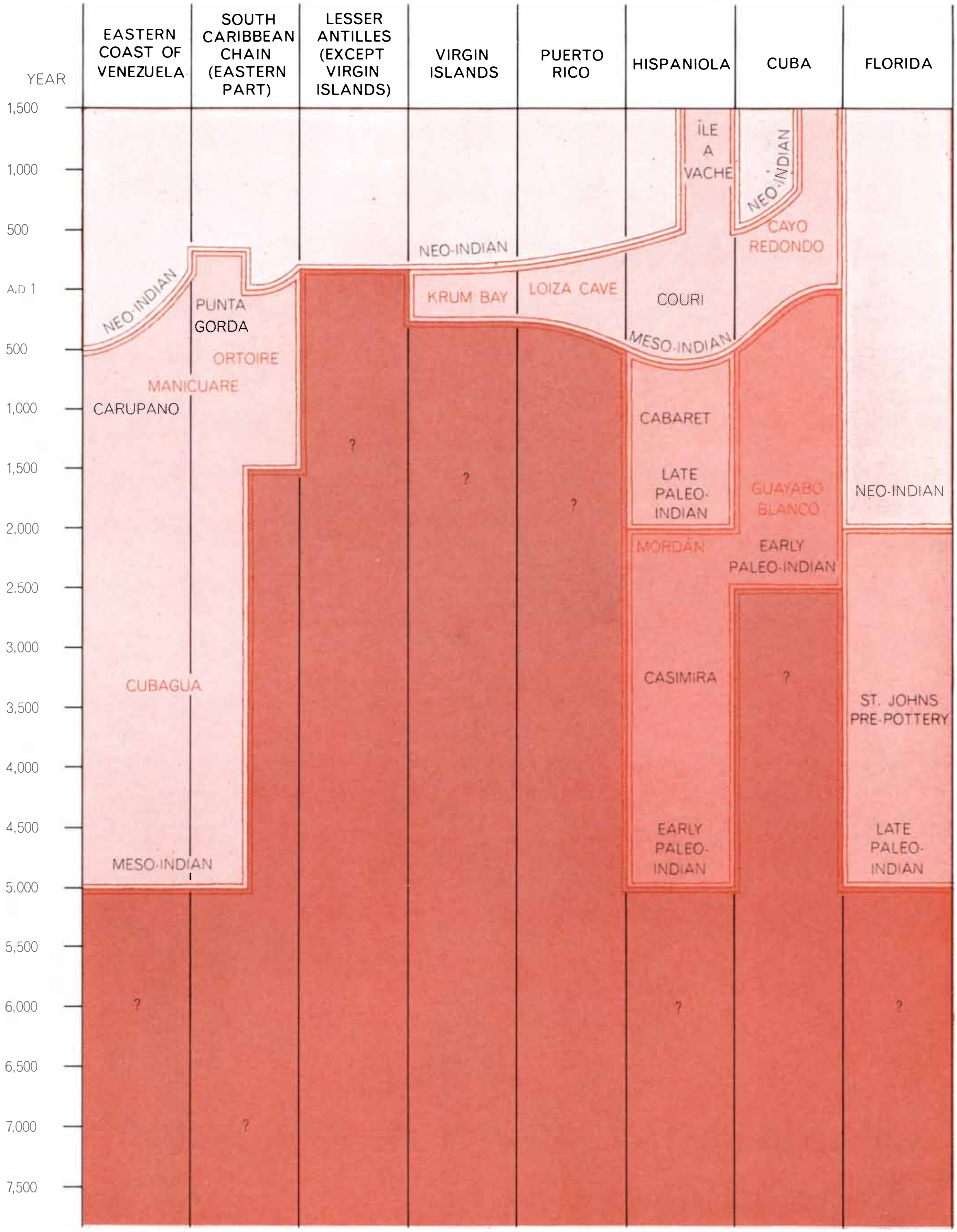
Two other sites were found a short distance inland from Mordán. Both contained flint artifacts but no shell refuse or pottery. At one of them, Rancho Casimira, a two-by-two-meter trench was dug to a depth of 50 centimeters. No charcoal or hearths were found; the contents consisted exclusively of crudely made artifacts and waste flint of poorer quality than the Mordán flint. The upper 25-centimeter section at Casimira bore flint of a quality slightly superior to that in the lower section, as if the Casimira workers had used any flint at hand in the beginning but had later succeeded in finding a source of better material. Nonetheless, only a few of the flakes seem to be from prismatic cores, all the artifacts are strikingly larger and heavier than their Mordán counterparts and the flakes have thick, unretouched ends that would certainly have precluded hafting. None of the implements show evidence of retouching. Signs of wear indicate that the blades were used as scrapers and some of the heavier pieces served as pounders and choppers [*see illustration on pages 50 and 51*].



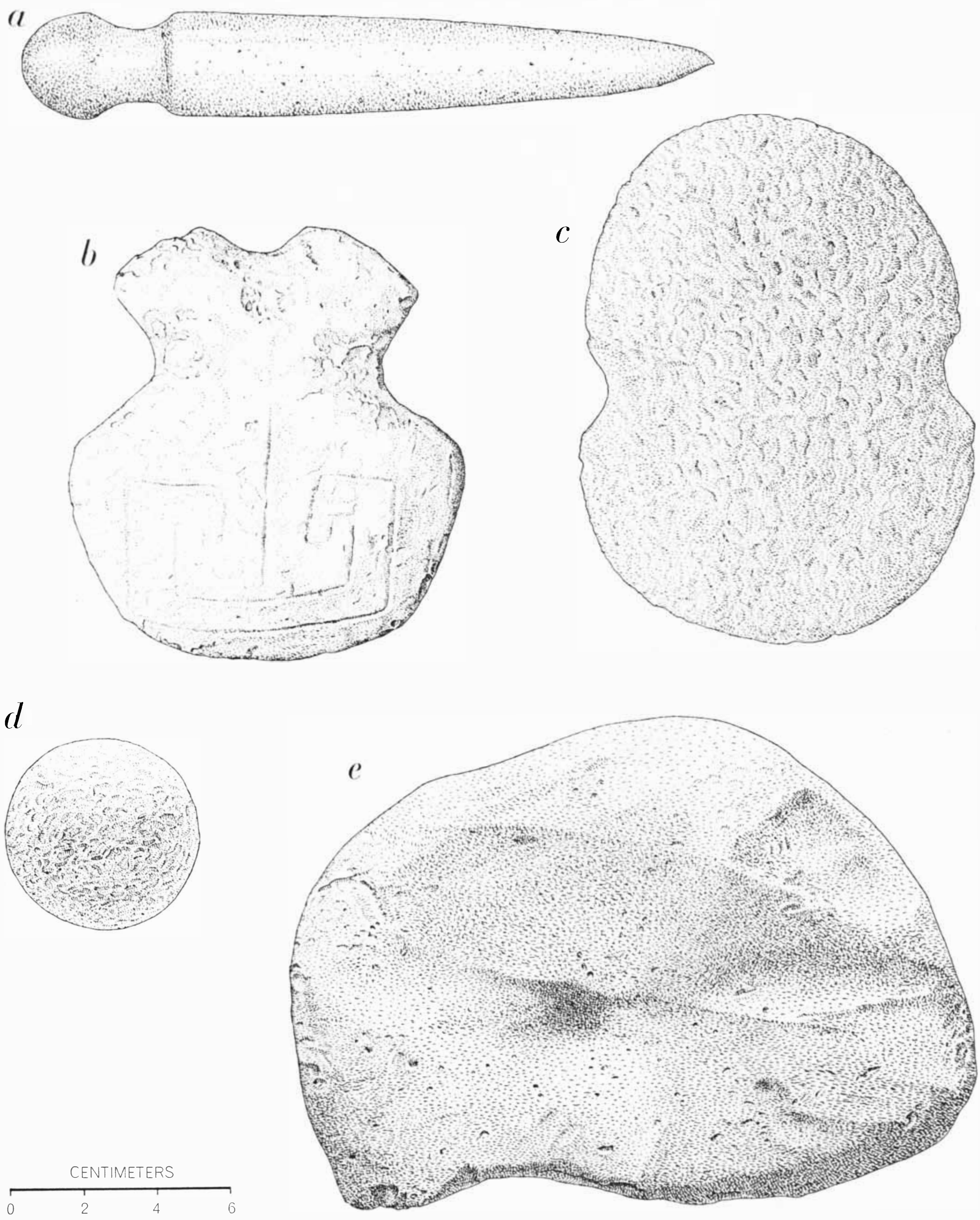
HISPANIOLA, politically divided between Haiti and the Dominican Republic, has a number of archaeological sites representative of the Meso-Indian and Paleo-Indian ages of New World prehistory. The artifacts unearthed at Île à Vache and Couri in Haiti are made out of stone that was shaped by grinding; they are therefore representative of the Meso-Indian age. At sites near Cabaret, also in Haiti, the artifacts are made from chipped flint. They include projectile points and thus belong to the preceding Late Paleo-Indian age. The sites at Mordán and Rancho Casimira in the Dominican Republic yielded Early Paleo-Indian artifacts. Carbon-14 dates show that the Mordán site is at least 4,000 years old. The Casimira site is apparently even older; it may have been first occupied as long as 7,000 years ago.

With the exception of Mordán and Casimira, all the Hispaniola sites described here were excavated before carbon-14 dating had been developed. In recent years work at a number of sites in the West Indies and on the mainland has produced a substantial list of carbon-14 dates. In the case of the Meso-Indian complexes of the Greater Antilles these dates have tended to cluster around the beginning of the Christian era. We had assumed that the same was probably true of the Cabaret complex in Hispaniola. The absence of projectile points and hafting preparations from the Mordán and Casimira complexes placed them in the Early Paleo-Indian age, however, and we eagerly awaited the carbon-14 analysis of the Mordán charcoal.

The samples yielded three dates: 2190 (±130) B.C., 2450 (±170) B.C. and 2610 (±80) B.C. This meant that we had been at least 2,000 years off in our estimate of when man first reached Hispaniola. Indeed, in view of the virtual certainty that the Casimira complex substantially predates the Mordán, the Casimira complex could date back as far as 5000 B.C. In any case the unexpected antiquity of



PREHISTORIC AGES were entered at different times in different parts of the West Indies. The Neo-Indian age, characterized by a knowledge of pottery, began in Florida, for example, about 2000 B.C. and in Venezuela late in the first millennium B.C. It had not yet begun in parts of Cuba and Hispaniola, however, at the time of the first European contact. On Hispaniola the Early Paleo-Indian age may have been introduced by migrants as long ago as 5000 B.C. By that time the inhabitants of Florida were representative of the Late Paleo-Indian age and the inhabitants of Venezuela had entered the subsequent prehistoric period, the Meso-Indian age. Many West Indian complexes are of uncertain age, but the dates indicated for the names shown in color are based on carbon-14 findings.



MESO-INDIAN ARTIFACTS from two sites in Haiti were made by grinding the desired shapes out of stone. The uses of the stone pin (*a*) and the engraved stone (*b*) from Île à Vache are not known; the Meso-Indians who made them may have been contemporaneous with the pottery-making Neo-Indians who reached Hispaniola about A.D. 500. The other artifacts are from sites of the Couri complex. They are a double-bitted axe (*c*), a stone sphere of unknown function (*d*) and a large stone with a shallow depression (*e*), possibly used as a kind of mortar for preparing wild vegetables. All but the large stone are shown two-thirds actual size. The Couri sites also belong to the Meso-Indian age but are clearly earlier in date than the more sophisticated complex unearthed at Île à Vache.

the Mordán complex has made us revise our views concerning the chronology of the Paleo- and Meso-Indian complexes throughout the West Indies [*see illustration on page 47*].

Where do these new findings lead us? Let us put forward the hypothesis that the Casimira complex did appear on Hispaniola about 5000 B.C. What are the implications of this early date? One implication stems from the improbability that lower material cultures are derived from higher ones. For this reason the Casimira people are not likely to have been emigrants from either the northern or the southern mainland; they could scarcely have come from Florida because Florida's inhabitants were then in the Late Paleo-Indian age, and they could scarcely have come from Venezuela because by that time Venezuela's inhabitants were Meso-Indians.

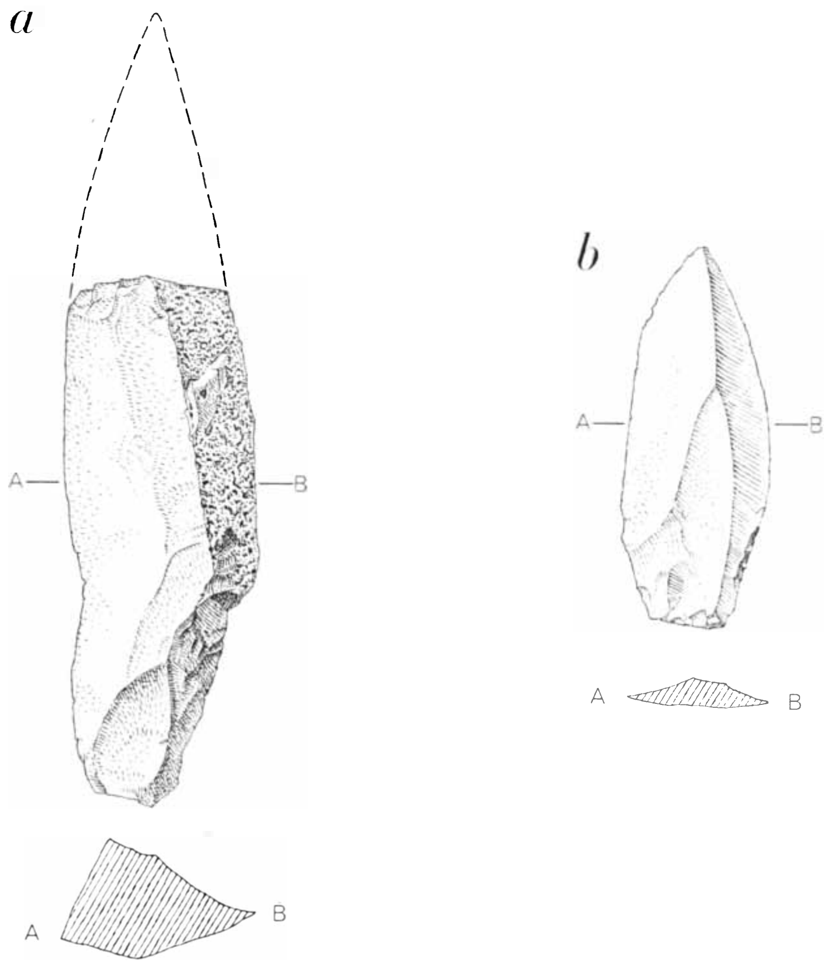
If, on the other hand, Casimira is only slightly older than Mordán, there would seem to be no obstacle to attributing the complex to migrants from Cuba, where the Early Paleo-Indian age did not end until the time of Christ. The assumption of a later date for the Casimira complex would therefore appear to be fatal to our hypothesis, and the reader may view with suspicion our selection of the date that appears to serve the hypothesis best. Nonetheless, we can offer considerable evidence in support of this choice.

Let us first dispose of the weakest alternative: a migration from the south. One would have to assign Casimira a date hundreds and perhaps thousands of years earlier than 5000 B.C. before one could find a well-substantiated Early Paleo-Indian population available to emigrate from either Venezuela or the South Caribbean islands. The carbon-14 dates for the Meso-Indian culture sequence in that region run from 750 B.C. back to 3800 B.C., and it appears that a safe initial date for the Meso-Indian age there would be about 5000 B.C. We have no secure carbon-14 dates for the two Paleo-Indian ages in Venezuela, but we have reason to believe the transition from the Early to the Late Paleo-Indian age in that region took place about 12,000 years ago.

What about the possible migration from Florida? We know that the Neo-Indian age began in Florida about 2000 B.C. Near the headwaters of the St. Johns River a number of sites of an earlier age have been excavated. One of the principal artifacts, unearthed in large numbers, is a kind of gouge made by breaking a triangular section from the outer whorl of a conch shell and grinding one edge of it. These gouges are found in association with pins made of bone and with flint projectile points that are flaked on both sides. There are no carbon-14 dates for the sites, but the absence of ground-stone artifacts and the presence of projectile points suggest that St. Johns is a complex of the Late Paleo-Indian age. This complex is dated between 5000 and 2000 B.C. If Casimira is somewhat more recent than 5000 B.C., one could envision a migration from Florida as its source, except for two facts: neither the shell gouge, the characteristic St. Johns artifact, nor the St. Johns type of projectile point has ever been found on Hispaniola.

There remains the possibility of migration from Cuba. If the date of Casimira is closer to 2500 B.C. than to 5000 B.C., some of the Early Paleo-Indians of a complex known as Guayabo Blanco might have migrated to Casimira. Carbon-14 analysis indicates that the Guayabo Blanco complex was in existence as early as 2050 B.C., and it is our opinion that the complex or something quite like it existed in Cuba at least a few hundred years earlier. Could this complex be the source of Casimira? We think not: the artifacts of the Guayabo Blanco complex are made chiefly of shell and not of flint. They include shell vessels and, as in the St. Johns complex, the principal tool is the shell gouge. Objects made of flint are rare, which suggests that they may have been trade items. Unlike the flint artifacts of the St. Johns complex, none are projectile points.

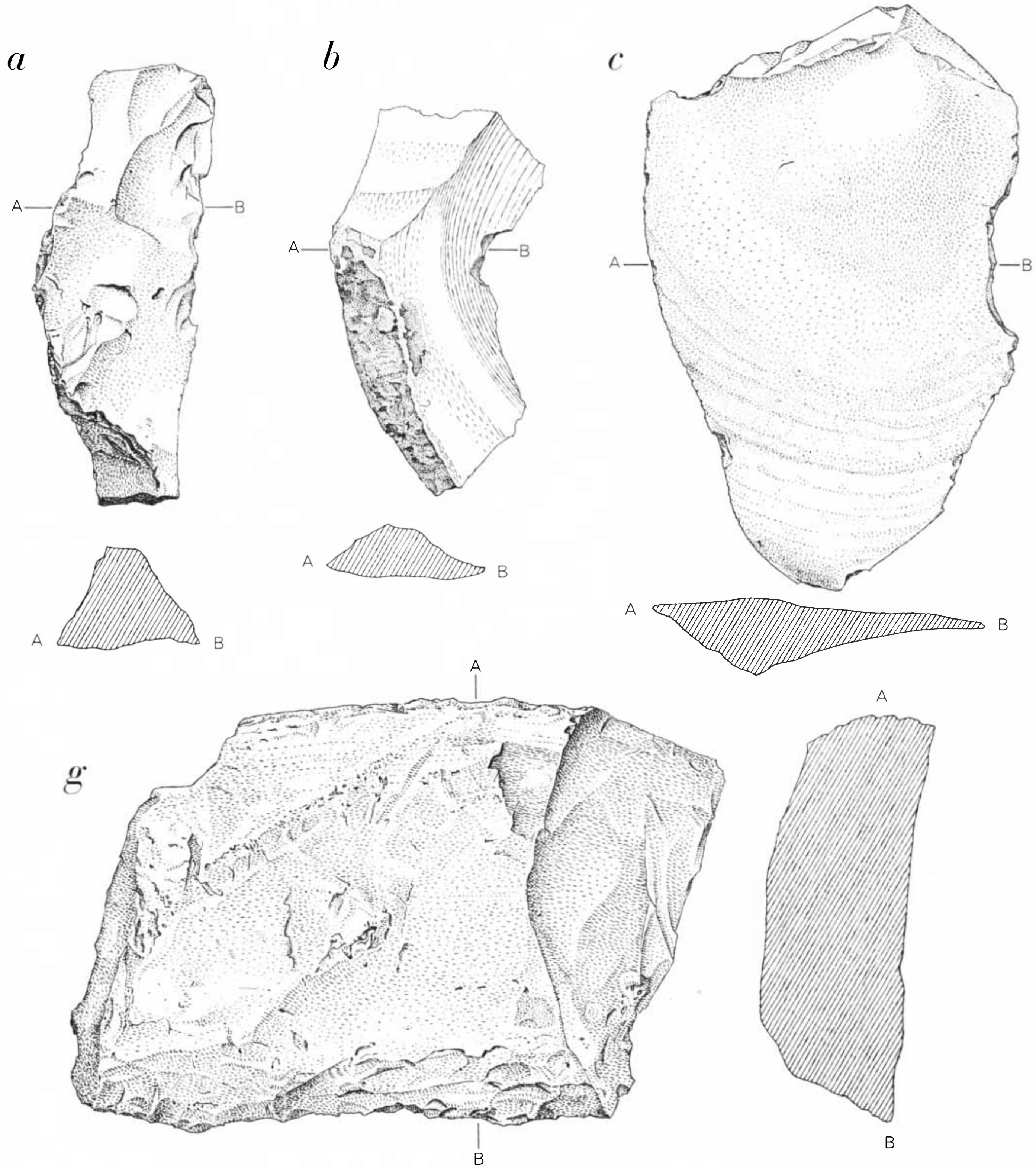
The nearly contemporary age of the Mordán and Guayabo Blanco complexes and the absence of projectile points at both sites point to the closeness of the Hispaniolan and Cuban complexes of the Early Paleo-Indian age. Still, two facts—the paucity of flint tools at the Cuban site and the total absence of shell gouges on Hispaniola—seem to rule out intimate cultural ties between the two complexes. It seems more likely that Cuba's Early Paleo-Indian complex was



FLINT TOOLS of Paleo-Indian workmanship are from sites of the Cabaret complex in Haiti. The projectile point (*a*) has been given a lateral stem to allow hafting. Its tip is missing; the broken line suggests its original appearance. The flake (*b*) has also been reworked at one end for hafting; it was probably a scraper. Both belong to the Late Paleo-Indian age.

derived from Florida, although we know too little about the pre-pottery cultures of Florida to state this as a certainty.

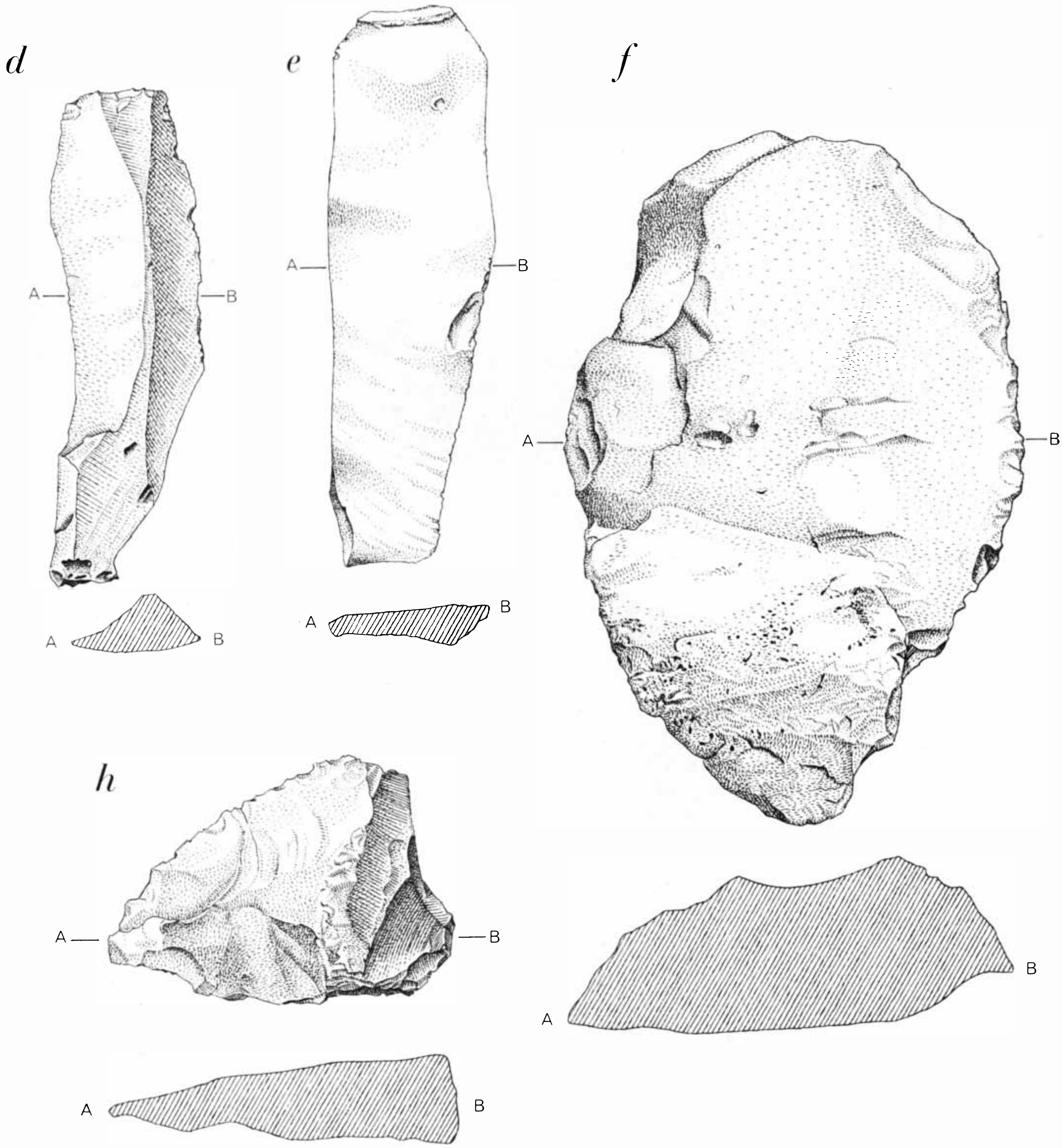
With Venezuela, Florida and Cuba evidently ruled out as sources of the Casimira complex, only one mainland area remains: Central America. It is a further part of our hypothesis that the Casimira people pushed out from Central America into the Caribbean, traveling by way of the Mid-Caribbean island chain first to Jamaica and then to Hispaniola. This part of our hypothesis has weaknesses, at least for the present. For example, although a flint-working tradition like that at Casimira and Mordán is known in Central America, the mainland version is considerably more developed and its artifacts have been found only in a Neo-Indian context. We suggest



OLDEST-KNOWN TOOLS on Hispaniola, made from an inferior grade of flint, were excavated at Rancho Casimira in the Dominican Republic. Among them are narrow scrapers that are often quite long (*a*, *b*, *d*, *e*), broad scrapers, both large (*c*) and small (*h*),

that the Central American tradition has Paleo- and Meso-Indian antecedents that remain to be discovered. Another weakness is that no prehistoric habitation sites of any kind have been reported on the Mid-Caribbean islands and no Paleo- and Meso-Indian ones have been found on Jamaica. No one has systematically looked for such sites, however, and the absence of reports does not necessarily mean an absence of sites.

Two further questions remain to be answered with respect to our hypothesis: Why did the Casimira people leave the mainland and how? We believe the answer to both questions is implicit in the nature of the flint implements at Casimira. The scrapers, pounders and choppers are typical of the tools made by the big-game hunters of the Late Paleo-Indian age on the mainland, who used



and massive chopping tools (*f*, *g*). Lacking projectile points, the Casimira people lived in the Early Paleo-Indian age. Their tools are nevertheless typical of mainland big-game hunters; they may have hunted large sea mammals such as the seal and the manatee.

BOULDERS OF FLINT and smaller flint fragments litter the ground for several miles around the village of Mordán, where the first Early Paleo-Indian site on Hispaniola was found.

SCREEN SIFTER, suspended by ropes, was used to separate the seashell debris at the Mordán site from the animal bones, flint tools and other artifacts that the deposit contained.

them to butcher and process their quarry. The implication that the Casimira people were hunters of big game does not mean that they did not also kill smaller animals and collect various vegetable foods, but large mammals would have been their main prey. Furthermore, we must assume that the Casimira people were at home on the water; wherever they came from, they did, after all, reach Hispaniola. Therefore their principal prey could easily have been large sea mammals. If we may judge by early Spanish accounts, manatees and seals once ranged among the islands in large numbers. Finally, we can logically expect that the Casimira people used rafts in pursuit of their quarry. So much for the "why"; we would assume that a mainland people followed sea mammals out to Hispaniola.

The "how" is far more conjectural. Rafts are largely at the mercy of winds and currents. Even assuming that the Casimira people could have equipped their rafts with mat sails, they would still have only been able to travel before the wind, and the prevailing winds are from the northeast. One can always conjure up a chance storm to drive a hunting party far out to sea, thus making at least the first migration accidental. The assembly of a viable breeding population on a remote island through pure chance, however, seems quite unlikely. We prefer to think of the Casimira people's moving purposefully out along the Mid-Caribbean chain (which, because of the lower sea level at the time, must have consisted of many more islands). Presumably they took advantage of winds and currents that would be in their favor at certain times of the year.

Arriving on Hispaniola, the hunters would have found a supply of the flint from which they were accustomed to making their tools. There would have been sea mammals along the shore and quantities of vegetable food and fresh water, which were probably scarce on many of the Mid-Caribbean islets. All in all, it is easy to see why the Casimira people would have settled down once they reached Hispaniola.

Our hypothesis runs counter to an assumption common among New World prehistorians that Paleo-Indians could only travel overland. We suggest that this assumption be reexamined. It seems entirely possible to us that various Paleo-Indians were using rafts both for river crossings and for coastwise travel in very early times. The first Americans need not have been restricted to overland routes for their movements, as many have supposed.

We want to be useful ...and even interesting

Kodak

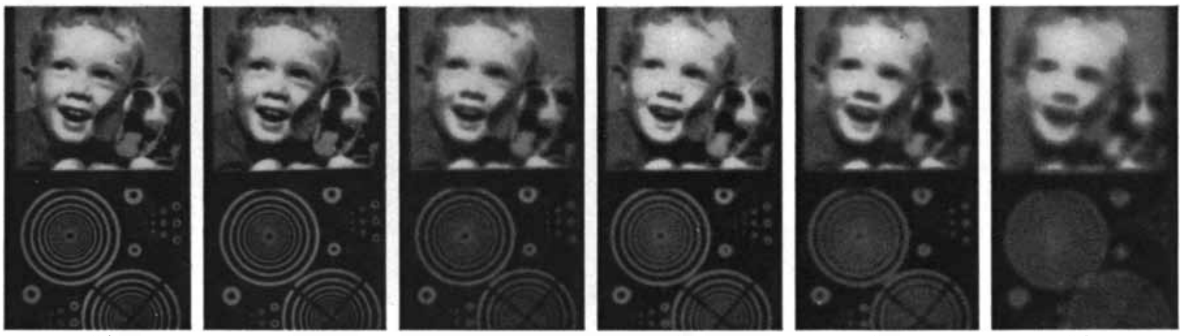
"I took this picture from a fire escape. I think that's the Empire State Building over there."

"My camera takes good pictures"

The picture on the left was taken and captioned by a young teenager in a project to see to what extent a camera can help such kids out of their bind. The picture on the right was taken and captioned by old pros. Any preference?

Now, below is another dimension for subjective evaluation of photography:

"True friends"



#1 #2 #3 #4 #5 #6

"Subjective, my foot," cries a voice. "That's nothing but resolution, or maybe focus, for goodness sake!"

No, sir (or madam, as the case may be). Not entirely. See a difference between #3 and #4? We don't. Yet the resolution target shows #4 resolving a lot more "lines per millimeter." These somewhat puzzling results were prepared for presentation in more meaningful form at the recent fall meeting of the Optical Society of America by K. N. Cupery and E. M. Granger of the Kodak Apparatus Division Research Laboratory.

In more ways than one, our business of photography is a subjective business in much need of highly objective research. What if this research should eventually upset some smug correlations between price paid and satisfaction delivered?

To think of ink

A couple of years ago one of us went to Jordan to help infrared photography of 2,000-year-old ink deliver a message to our times left by the Essenes in a cave above the Dead Sea.

Most modern ink delivers messages of a different tone. Riffle through the pages of this magazine for samples.

We have to think of ink as a link in the chain that starts from professional color film and ends in the brain of a magazine reader. Much depends on it. We must also think of inks that will work with sensitized materials involved in other kinds of presswork.

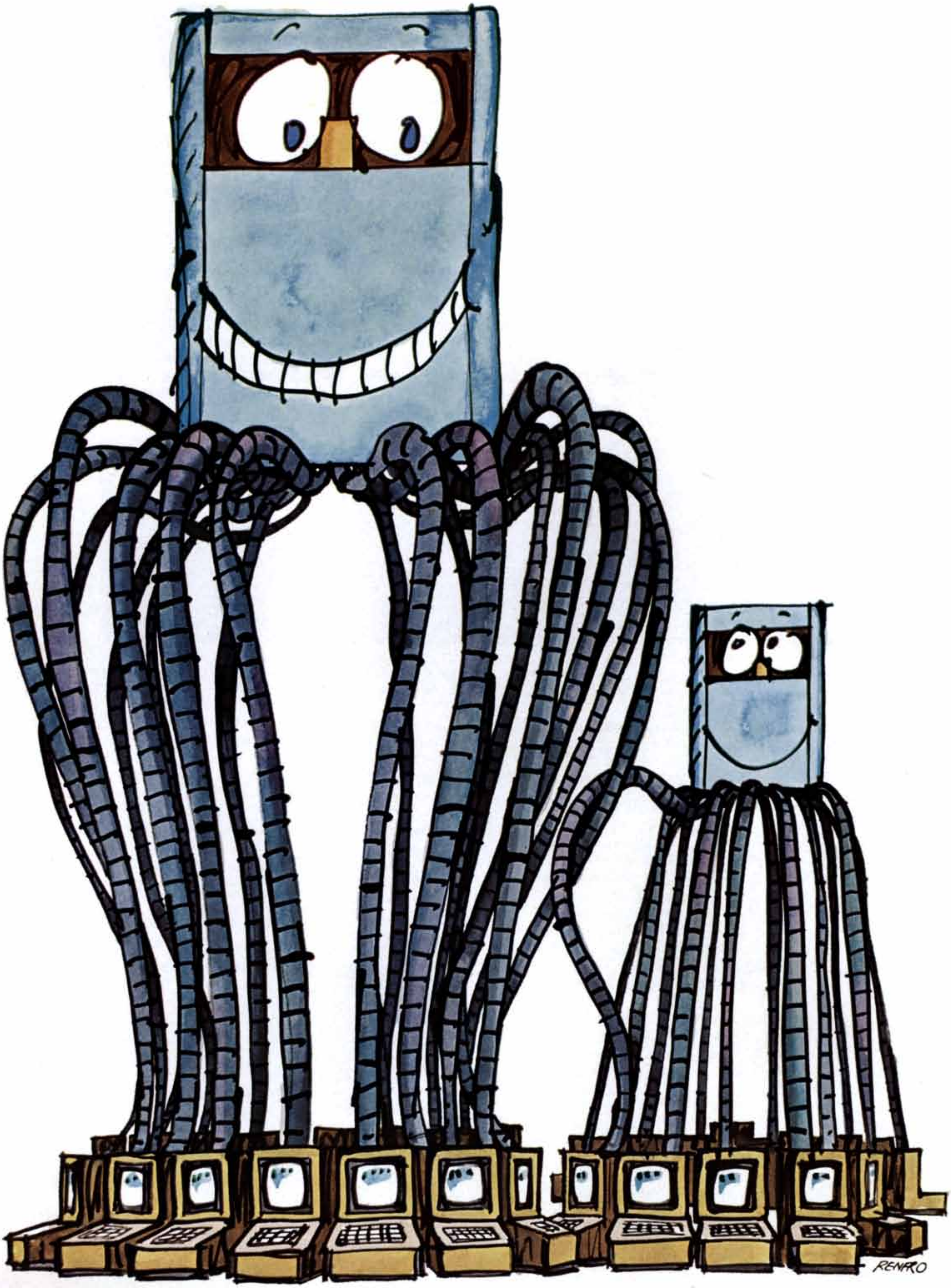
Though we think a lot about ink, we don't manufacture any for sale. This eliminates awkwardness in serving the firms who do compete in ink. There are quite a few of them, some quite small enterprises that earn favor by helping printers solve their problems. Papers come in vast profusion. Not all printing is done on paper. Some printing has to smell good. Some printing stands a chance of being eaten. Some has to stand rain, some sunlight, some strong vapors, some much rubbing by little fingers.

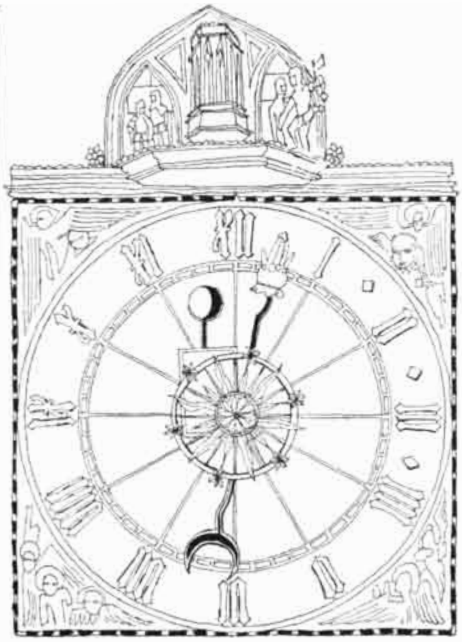
We, in turn, help the inkmakers. We do this by suggesting ingredients which we sell in our role as a tonnage producer of industrial chemicals. Language barriers intervene. Chemical engineers have to learn inkmakers' talk. Inkmakers have to learn printers' talk. Printers have to learn to read their customers' thoughts. The customers may not fully appreciate what printers can do today.

For introductions to firms capable of discussing feasibility of new ideas for using ink to deliver messages, write to Department 923, Eastman Kodak Company, Rochester, N.Y. 14650. Makers of ink who wish to test our direct helpfulness should write Eastman Chemical Products, Inc., Chemicals Division, Kingsport, Tenn. 37662.

KODAK Infrared Scope: A 21-ounce, 1.1X telescope for observing infrared emission or for seeing by infrared illumination. Peak response at 800 nm, 90% of peak at 700-900 nm, total sensitivity range 400 to 1200 nm. Objective can be focused from one foot to infinity. Field of view 26°. Resolves at least 5.5 minutes through focusing eyepiece on viewscreen of its green-emitting image converter. Built-in power supply delivers 12,000 volts from a 1.34-v mercury battery. We buy these metascopes in quantity for resale to photofinishers, who use them for darkroom chores under infrared safelight.

If you too need one, Department 927, Eastman Kodak Company, Rochester, N.Y. 14650, can provide a current estimate of how much change to expect back from $1,000.


RENARO



SCIENCE AND

Cleaner Waters

The U.S. and the U.S.S.R. joined in a sixth finite step along the road to disarmament last month when they agreed on the draft of a treaty to keep the ocean floor free of nuclear weapons. The earlier steps were treaties demilitarizing Antarctica (1959), controlling nuclear-weapons tests (1963), prohibiting nuclear weapons in space (1966), barring nuclear weapons in Latin America (1967) and controlling nuclear proliferation (1968). The new proposal was submitted by the two superpowers to the other members of the continuing disarmament conference in Geneva preparatory to being placed before the UN General Assembly.

Signatories of the treaty would agree not to "emplant or emplace" on or under the ocean floor any nuclear weapons or other weapons of mass destruction or facilities for storing, testing or launching them. Each signatory would have the right to "verify" suspicious activities. The treaty would apply outside a 12-mile territorial limit, the "maximum contiguous zone" defined in the 1958 international convention on territorial seas. The U.S. had proposed exempting only a three-mile zone. It accepted the wider zone in return for a significant narrowing of the substantive provisions of the treaty: the U.S.S.R. had originally proposed the complete demilitarization of the ocean floor. The present draft essentially means that a signatory could not install nuclear mines or nuclear missiles on the deep-ocean floor. It presumably would not affect such ancillary devices as navigation beacons or sensors emplaced on the bottom. Nor would it prohibit missile-carrying submarines, underwater barges or "crawler" vehicles on the ocean floor, whether within or outside the 12-mile limit.

Fundamental Right

A decision by the Supreme Court of California overturning the conviction of a physician under a state abortion law may prove to be a significant step toward the liberalization of abortion laws in the U.S. The case involved Leon Phillip Belous, an obstetrician and gynecologist in Los Angeles who has supported such liberalization. He gave an unmarried woman in an early stage of pregnancy the telephone number of a man (regarded by him as being a competent physician) who, licensed to practice medicine in Mexico but not in California, was performing abortions in Los Angeles. Belous testified that he thought the woman's life was in danger because she had led him to believe she would "do anything" to terminate the pregnancy if he did not help her. He was convicted under a state law, enacted in 1850, that forbade performing or abetting an abortion unless it was "necessary to preserve" the woman's life.

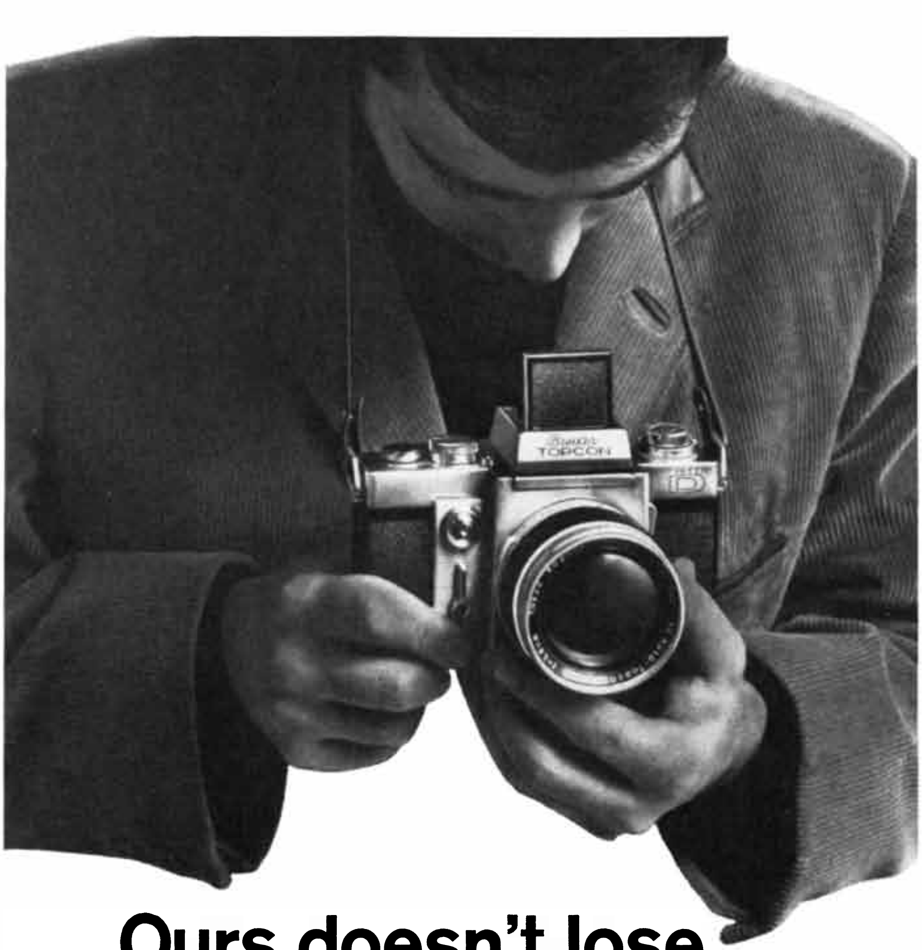
The California Supreme Court reversed the conviction on the ground that the law was invalid because of the vagueness of the term "necessary to preserve." Declaring that a woman has a "fundamental right . . . to choose whether to bear children," the court said the question to be decided was to what extent the state could infringe that right. When the statute was enacted in 1850, the court pointed out, the art of surgery was so poorly developed that an abortion entailed a grave danger to the patient's life. The risk of death or infection is still high in illegal abortions, the court said, but "it is now safer for a woman to have a hospital therapeutic abortion during the first trimester than to bear a child." Hence the 1850 law, which was designed to protect the life and health of women, now tends to have the opposite effect.

The court also attacked the statute because it makes the physician decide whether or not an abortion is justified and exposes him to criminal penalties if he authorizes one that the state holds to have been illegal. Since the pressure is therefore on the doctor to decide against an abortion, the court said, "a woman whose life is at stake may be as effectively condemned to death as if the law flatly prohibited all abortions." Taking note of California's recent Therapeutic Abortion Act, which authorizes an abortion by a licensed physician in an accredited hospital when the abortion is approved in advance by a committee of the hospital's medical staff, the court said: "At least in cases where there has been adherence to the procedural requirements of the statute, physicians may not be held criminally responsible, and a jury may not subsequently determine that the abortion was not authorized by statute." This statement has been interpreted as offering encouragement to physicians to be more liberal in their interpretation of the abortion law.

Gamma Ray Star

The first gamma ray star may have been discovered. On the basis of two high-altitude balloon flights over Australia a team of U.S. and Australian physicists has reported an apparent point source of cosmic gamma rays. Cosmic electromagnetic radiation ranges through the entire electromagnetic spectrum, from long radio waves through visible light to X rays and gamma rays. X-ray astronomy has so far yielded some 40 discrete sources. Until now gamma ray astronomy (which deals generally with photons, or particles of electromagnetic radiation, having energies greater than 10 million electron volts and wavelengths shorter than 10^{-10} centimeter) had given evidence only of a diffuse radiation that seemed to come from the galactic plane, with a maximum at the center of the galaxy.

Last February, G. M. Frye, Jr., J. A. Staib and A. D. Zych of Case Western Reserve University and V. D. Hopper, W. R. Rawlinson and J. A. Thomas of the University of Melbourne collaborated on two balloon flights, in both of which a gamma ray detector was lifted to about 80,000 feet in order to get above most of the earth's atmosphere, which absorbs gamma rays. The results from the two experiments were in good agreement: There was a significant excess of cosmic gamma radiation from a point in



the constellation Sagittarius that could be resolved to within two or three degrees. Although the data from the two flights suggest the existence of a gamma ray star, they also cast doubt on the existence of the diffuse gamma radiation, which had been reported last year by George W. Clark, Gordon P. Garmire and William L. Kraushaar. The Australian results showed no significant excess of radiation along the galactic plane.

The Genetic Code (Reconfirmed)

Investigators at the University of Cambridge have reached the long-sought goal of associating a specific sequence of amino acid units in a protein with the sequence of "letters" in the genetic message—a molecule of RNA—that supplies the instructions for the protein's manufacture. The experiment confirms the correctness of the genetic code dictionary, which was compiled by using the four RNA bases—adenine (A), uracil (U), guanine (G) and cytosine (C)—in the various triplet combinations called codons, to see which of the 20 amino acids they would direct into protein-like molecules in a cell-free system. In this way a specific "meaning," or amino acid equivalent, was assigned to 61 of the 64 possible codons from AAA to UUU. Three of the codons have noncoding functions.

The Cambridge workers (J. M. Adams, P. G. N. Jeppesen, Frederick Sanger and B. G. Barrell) identified a sequence of 57 bases out of a total of 3,300 found in the RNA that carries the hereditary message in the bacterial virus designated R17. The 57-base sequence was found to be part of the cistron, or gene, that specifies one of the three proteins synthesized by the virus: the coat protein. The amino acid sequence of this protein had previously been worked out by Klaus Weber of Harvard University. If one knows this sequence, one can, with the aid of the genetic code dictionary, write down all the sequences of codons that will code it. There is more than one possible sequence because nearly every amino acid can be represented by more than one codon.

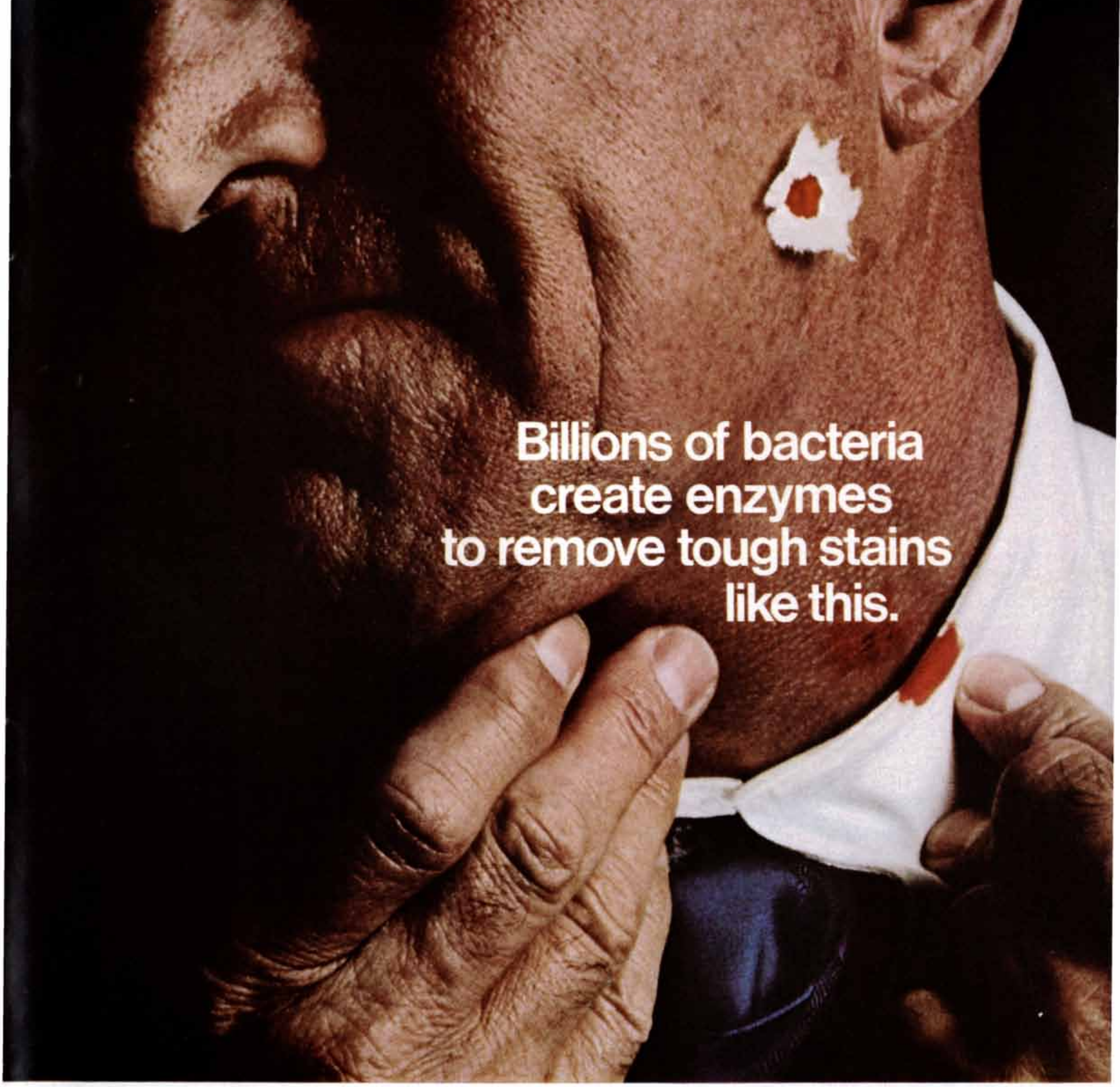
Particular interest centers on the six amino acids that appear twice in the sequence of 57. Would the virus use the same codon for each repeated amino acid or a different one? It turns out that two of the six are specified both times by the same codon. The other four, however, are each specified by two different codons. Thus the amino acid serine is coded for by UCC and UCG; asparagine is represented by CAU and AAC, threonine by CCU and ACG, and alanine by ACU and GCG.

An additional finding of the Cambridge workers is that the sequence of 57 bases can be arranged in a simple loop in such a way that 38 of them could form 19 hydrogen-bonded base pairs similar to those found in the two-strand helical model of DNA, the genetic material of higher organisms. This suggests that the RNA that carries the genetic message in the R17 virus, and in other viruses of its type, has a structure with a double-helix configuration, at least in certain regions.

Protein-rich Gruel

The world's first commercial product made from a new and significantly more nutritious strain of corn has been put on the market in Colombia. It is an infant-feeding formula made from high-lysine corn, and three eight-ounce servings a day, costing less than two cents each, provide 100 percent of a six-month-old infant's minimum protein requirements. The nutritional value of ordinary corn is limited by the fact that its proteins, like most vegetable proteins, almost completely lack two of the essential amino acids that are necessary for growth: lysine and tryptophan. In 1963 Edwin T. Mertz, Oliver E. Nelson and Lynn S. Bates of Purdue University discovered a mutant gene, opaque-2, that controls the synthesis of lysine and tryptophan in the corn kernel, and they developed an opaque-2 strain that contains a good balance of all the essential amino acids.

Dale D. Harpstead of the Rockefeller Foundation saw that the high-lysine corn could have "great and far-reaching impact" in Colombia and in many other underdeveloped areas where corn is a staple of the diet and where protein-deficiency disease is common in children. The Purdue strain was crossed with local hybrids and feeding tests were conducted at the University of Valle hospital in Colombia. The high-lysine corn cured children who were near death from such protein-deficiency diseases as kwashiorkor. For a nutritious new food to be useful, however, it must be accepted by the population. CPC International, Inc., undertook to prepare a high-lysine product in a familiar form: a cornstarch of the kind traditionally used in Colombia to prepare a "colada," a gruel made with sugar and water or milk. The product, called Duryea, is now being manufactured and marketed by a CPC subsidiary in Colombia. In the U.S., meanwhile, CPC is sponsoring large-scale tests of the



Bacteria are playing the key role in the enzyme boom that is revolutionizing the detergent industry.

Coddled in aseptic surroundings at exactly the right temperatures, these bacteria are producing enzymes by the kettleful for Monsanto. These include protease enzymes that dissolve protein stains such as blood, egg and skin oils, and amylase enzymes that attack carbohydrates in gravy, chocolate and ice cream.

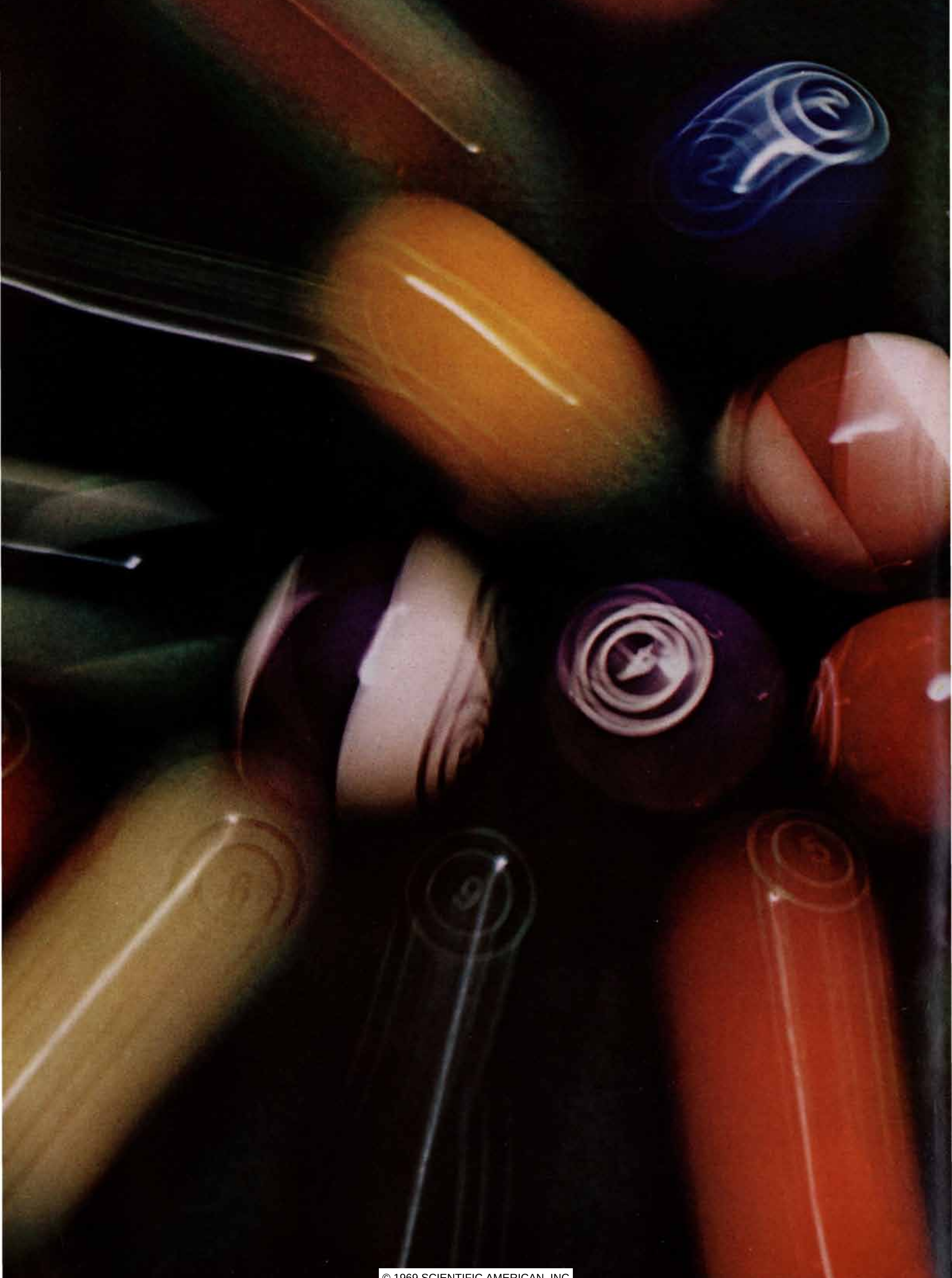
Monsanto has combined these two enzymes to produce a unique product for pre-soaks and detergents. This enzyme product makes home laundering far more efficient than ever before.

Our extensive experience in microbiology and enzymology enables us to produce this unique combination enzyme for the detergent industry. What's more, we have the ability to mass produce many other enzymes as well—for a wide range of industrial processes.

Developments like this are coming from Monsanto—where science and research are working for people.

Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, Missouri 63166

Monsanto



How many can your eye follow at once?

Imagine an eye that can see everything. Everywhere. At once.

Now you have some idea what our new phased array radar is like.

Right now, this new motionless radar is designed primarily for defense. But some day soon, it will help keep track of airline flights, warn of dangerous weather, possibly even help control the flow of highway traffic.

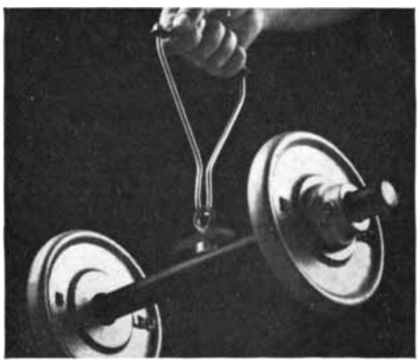
We are continually pushing back the frontiers of technology and putting the results to work in a variety of fields.

Our work in light amplifiers will allow policemen, pilots, mariners, and naturalists to see in the dark.

To improve microwave tubes, Raytheon scientists developed the strongest permanent magnets now in use. These magnets will find many other important uses in communications, medicine, industry, and scientific research.

At Raytheon, common threads of technology run through and weave together our diverse and growing businesses . . . in space, defense, air traffic control, natural resources, education, home appliances, communications and data handling, electronic components, and marine electronics. Raytheon Company, Lexington, Massachusetts 02173.

Raytheon light amplifier strips away cover of darkness.



Strongest magnet in use supports 500 times own weight.

Raytheon makes things happen.

RAYTHEON

high-lysine corn as livestock feed, particularly for hogs, which (unlike ruminant animals) are unable to synthesize certain amino acids that are not included in their feed.

Harappan May Be Dravidian

The urban civilization of the Harappans of the Indus River valley, in what is now West Pakistan, flourished for nearly 1,000 years during the third and second millenniums B.C. Its collapse left a rich legacy of problems for modern scholars. Neither the archaeological record nor the earliest traditions of the succeeding Aryan people, for example, contain useful clues to the identity of the Harappans. Neither have any of the 2,000 or more Harappan inscriptions that survive been intelligibly deciphered. Three linguists and a mathematician in Finland have now opened the way for a meaningful attack on the second problem. If their theory is correct, the first problem has also been solved.

Writing in *Antiquity*, John Chadwick of the University of Cambridge recently summarized the preliminary announcement of a Harappan decipherment made earlier this year by Asko Parpola, Simo Parpola, Seppo Koskenniemi and Pentti Aalto. They conclude that the Harappan script represents proto-Dravidian, a language spoken today in its mature form both in central and southern India and in parts of Baluchistan, to the west of the Indus. In other words, the Harappans were Dravidians. Using computer techniques to speed the isolation of individual words, suffixes, prefixes and the like, the decipherers found that some words were represented by a single sign that provided a quasi-phonetic hint of the spoken sound. Other words were made by combining signs: one would indicate the general class of objects to which the word belonged and the other would provide the phonetic hint.

An example of a single-sign word is the one that acts as a plural suffix; it shows a man with a pole across his shoulders and a load at each end of the pole. In proto-Dravidian the word for a bamboo carrying pole is *kaṛai*. A close homophone, *kar̲rai*, means "collection" or "bundle" and could readily symbolize the concept of plurality. Another single-sign word is "woman." It is a drawing of a five-toothed comb, teeth pointing right. The proto-Dravidian word for comb is *peṇṭika*, a close homophone of *peṇṭi*, "woman."

Compound words include the sign for "man," a simple stick figure, in combination with an unmistakable picture of a scorpion. The proto-Dravidian for "scorpion" is *tēḷ*, a virtual homophone of *teḷ*, meaning "knowledge" or "learned person." The decipherers suggest that the word describes a class of scholars. Other signs used in combination with "man" include a bow-and-arrow pictograph and a drum pictograph. The proto-Dravidian for "bow" (*vil*) is the same as for "sell"; a class of merchants is suggested. The proto-Dravidian for "drum" is *tampala;* for "priest," *tampa*. A class of religious officials is suggested.

Because most Harappan inscriptions are brief (the longest do not exceed 20 signs) it is doubtful whether the effort by the Finnish scholars, even if it is completely successful, will add much to the corpus of written history. It will, however, reduce to six the number of ancient scripts—one in the New World and five in the Old—that are still not understood.

This Winter's Weather

Meteorologists at the Massachusetts Institute of Technology believe they have made a fundamental advance in the technique of long-range weather forecasting. To demonstrate their confidence in their work three of the men involved have formed Statistical Weather Information, Inc., to sell long-range forecasts to clients who can profit from knowing what lies ahead. The firm's initial product, prepared with the help of a high-speed computer, is a forecast (made in September) of what U.S. weather will be like this winter.

The temperature prediction is that a *T*-shaped region running down the center of the U.S. and across the Canadian border will experience temperatures well below normal. The left arm of the *T* curves down to embrace most of California. In the Southeast it will be warmer than usual. It will be dry along the Pacific coast and in the Northeast, meaning a scarcity of snow for skiers. Very heavy precipitation is predicted for the Ohio and Tennessee valleys, with the likelihood of flooding.

The new method is described in a paper prepared for the Environmental Science Services Administration titled "Significant Advances in Statistical Long-Range Forecasting." It was written by Donald B. DeVorkin, a geophysicist and statistician, and John T. Prohaska, a meteorologist, who have been analyzing the problem for the past seven years under the direction of Hurd C. Willett. The three men are also the founders of the new forecasting firm.

In studying climatic fluctuations Willett has found evidence that persistent

Like mother. Like child. Like man.

The potential of nonhuman primate research for answering a host of key biomedical questions appears limitless. Thus, a few years ago, the National Institutes of Health established seven regional primate research centers. Of the seven, the one located at the University of Washington in Seattle is unique in many ways.

Physically connected with the University's medical and dental schools, the Center functions not only as a research tool but as a graduate educational facility. It is also interdisciplinary, for it actively involves some 19 separate fields and 154 professional scientists, post-doctoral fellows and graduate students. Additionally, the Center acts as an information hub. It collects and disseminates on a world-wide scale current primate research findings. Finally, the Center represents literally a new approach in laboratory planning, and has evolved many innovative techniques and some ingenious equipment.

Since the Center opened in 1964, more than 50 individual research studies have been developed. For the first time, trachoma, an eye disease that causes blindness in more than one-half billion people throughout many parts of the world, was produced in monkeys by a Center scientist. A cardiovascular study uses telemetry instrumentation to permit the observation of visceral and neural functioning in unanesthetized, intact simians, often in a behavioral situation. Again, in unanesthetized, behaving monkeys, another study deals with voluntary movement, an area of neurophysiology long neglected in favor of sensory studies. Cross-circulation, extra-corporeal irradiation, bone marrow preservation, transplantation and testing of immuno-suppressive drugs are a few of the other studies currently underway.

A vital phase of the Center's operations is its primate-breeding colony at Medical Lake, in the eastern part of Washington State. This breeding facility permits important perinatal biology studies, studies on oral-facial development and health-related research on birth defects and injuries such as cleft palate, congenital malformation (such as from thalidomide), mental retardation and cerebral palsy. Scientists from three local colleges in the area also benefit from the breeding facility.

In five short years, the Primate Center has achieved much. It promises still more.

The State of Washington offers a fertile field for science and the breakthrough industries of the future. In bio-engineering, aerospace, agriculture, fisheries, oceanography, atomic energy and many other areas, the smart money is betting on Washington State.

Daniel J. Evans
Governor

STATE OF WASHINGTON

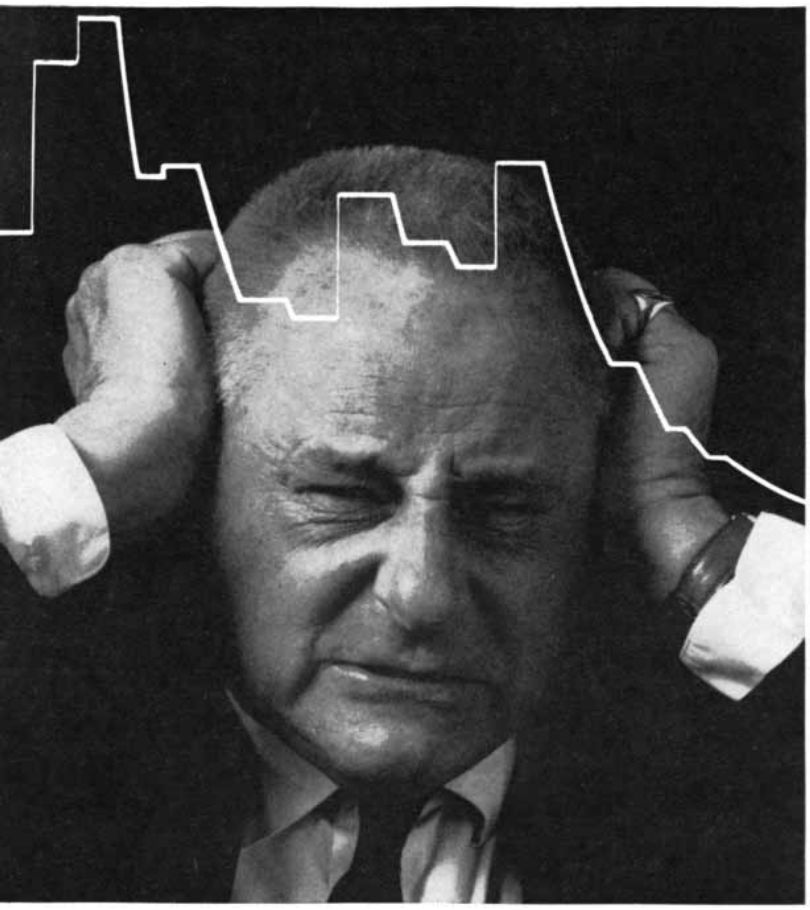
For business location information, write: Daniel B. Ward, Director, Department of Commerce and Economic Development, Olympia, Washington 98501

Computer-aided measurement lets you eavesdrop on airport noise. Make the most of crude oil. And keep tabs on the moon.

A way to quieter airports.

Jet aircraft are not designed to be noisy. The sound level depends more on how and where they're flown. And whether anybody cares. Around airports, people usually do. At Stuttgart, Germany, for example, after Hewlett-Packard installed an airport noise monitoring system, the level of noise dropped immediately. The pilots themselves became more "noise conscious" and exerted their own controls. On the other hand, some residents complained more. Publicity about the system increased their awareness, too.

Actually, a number of things can be done—if you know when and where the noise is building up. Airport officials can change flight patterns to avoid sleeping neighborhoods; they can require new engine settings and approach angles; or redesign the airport! And they can compare present and future performance with actual data logged in the past by the monitoring system.



Besides statistical analysis of airport noise, another boon for the future is that an offending aircraft can be detected in flight and brought to heel immediately. Chronic complainers can be mollified with statistics showing that aircraft are within their proper limits. This is all made possible by the speed and versatility of the HP 80500A Aircraft Noise Monitoring System, a combination of computer, real-time audio spectrum analyzer, and multiple remote monitors. This system can be programmed, with software provided, for any type of analysis the airport needs. The cost of a typical system would run about $70,000.

Airports can be easier to live around. For more information on how an HP airport system could help your community, write for Bulletin Number 80500A. Also yours for the asking is our 116-page Acoustics Handbook.

Analyzing the inconsistencies of oil.

Who knows how much high-grade gasoline you can get from a shipload of crude petroleum? Yet every time the refinery gets another shipment of crude, the chief engineer has to reset his refining towers for maximum efficiency: more heat at this point, less at that. It makes a great difference in the quality, quantity and type of petroleum products extracted, because no two oil shipments are the same.

Previously, analytical chemists had to set up a variety of distilling columns in the lab. Or chemical engineers had to run a sample through a pilot plant in search of the amount in each specified boiling range in the crude oil. Both laborious, time-consuming projects, actual distillations requiring six to one hundred hours per run.

But now there's a way to analyze the component boiling ranges automatically, at less than half the cost. Hewlett-Packard's 7600A Gas Chromatograph System is a meticulous alter-ego for the laboratory analyst. He sets the controls, supplies the samples and lets the 7600A simulate the distillations. (And the material doesn't have to be crude oil: any petroleum fraction will do.)

The chromatograph—a misnomer from blotting-paper days—takes an exact sample, squirts it into a vaporizer section, vaporizes it, separates the components according to their boiling points, drives the vapors past sensors—all at the right temperature—evaluates the components, draws a graph on a built-in recorder, integrates and prints out the quantities, and resets itself for the next sample.

If there's a computer handy, in the lab or by time-share, the chromatograph will prepare a tape with all the data. The computer is programmed with a BASIC language system supplied by Hewlett-Packard. The computer will then print out the initial and final boiling points of the sample and the boiling point at each percent in-between. These data provide a fast and convenient guide to refinery operation. In short, the HP 7600A Chromatograph System saves much time and frustration for chief engineers and analysts alike.

And the system is not limited to petroleum problems. In fact any application for GC analysis—the production of perfumes, alcohols, turpentines or ethical drugs—can be speeded by the HP 7600A. More information is contained in our brochure, to demonstrate the GC's potential and its ability to cut costs.

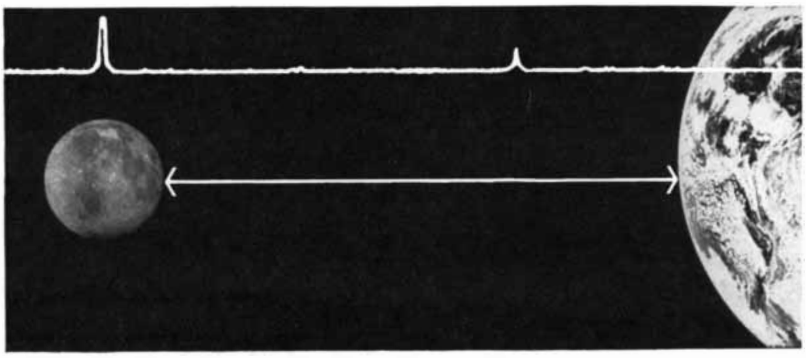
How much did the moon move today?

The moon pulls at the earth. And the earth pulls back. Predicted decay in the lunar gravitational field might be confirmed by precisely measuring changes in the radial distance between the earth and moon. But the perturbations are small indeed. Only with today's laser systems, and space accomplishments, can measurements of distances over a quarter million miles be exact enough to detect them. Accuracy depends on the quality of the system's resolution and the stability of its time base. But now, at least one new instrument can improve on range readings.

If you can get a clear signal to the moon and back, the Hewlett-Packard 5360A Computing Counter can calculate the distance within a foot. With its time-interval plug-in, the counter can resolve the time it takes signals to bounce off the moon within 100 picoseconds, and be accurate to 1 nanosecond.

Computing circuits inside the counter will convert the time to distance—feet or inches—out to 10 or 11 significant figures. The counter itself will decide from the measurement how many figures are significant.

The flow of calculations is directed by a small accessory keyboard, as on a small calculator. Except this one can remember up to 32 successive instructions. For example, if you want the distance to the moon computed even closer, you can crank in today's constant for the speed of light in the atmosphere. Or you can have the counter compare successive measurements, and compute the change in distance.



The 5360A Computing Counter can reduce the period, frequency, or time interval of any series of events to units directly meaningful to the engineer or scientist.

With appropriate plug-ins and a keyboard (available later this year), the 5360A performs arithmetic operations on its basic functions to permit such measurements as velocity, standard deviation, phase angle.

The counter sells for $6500, the accessory keyboard for about $1350.

If extremely high resolution measurements in the frequency and time domain are of value to you, write for our brochure on the HP 5360A Computing Counter.

And if you have any special problem you think computer-aided measurement might solve, give us a call. Hewlett-Packard, 1508 Page Mill Road, Palo Alto, California 94304; Europe: 1217 Meyrin-Geneva, Switzerland.

00910

HEWLETT hp PACKARD

Measurement, Analysis and Computation

patterns in the global circulation of the atmosphere are associated with differing levels of solar activity. In particular he finds that periods of low solar activity, such as the one now coming to a close, are associated with a "low-latitude zonal pattern," characterized by an intensification of westerly wind flows at latitudes to the south of the normal maximum flow. This pattern, which dominated the 1960's and persisted until as recently as last summer, tends to raise temperatures in the southern part of the U.S. and lower them in the northern part. Last winter the pattern contributed to one of the harshest winters on record in the U.S.

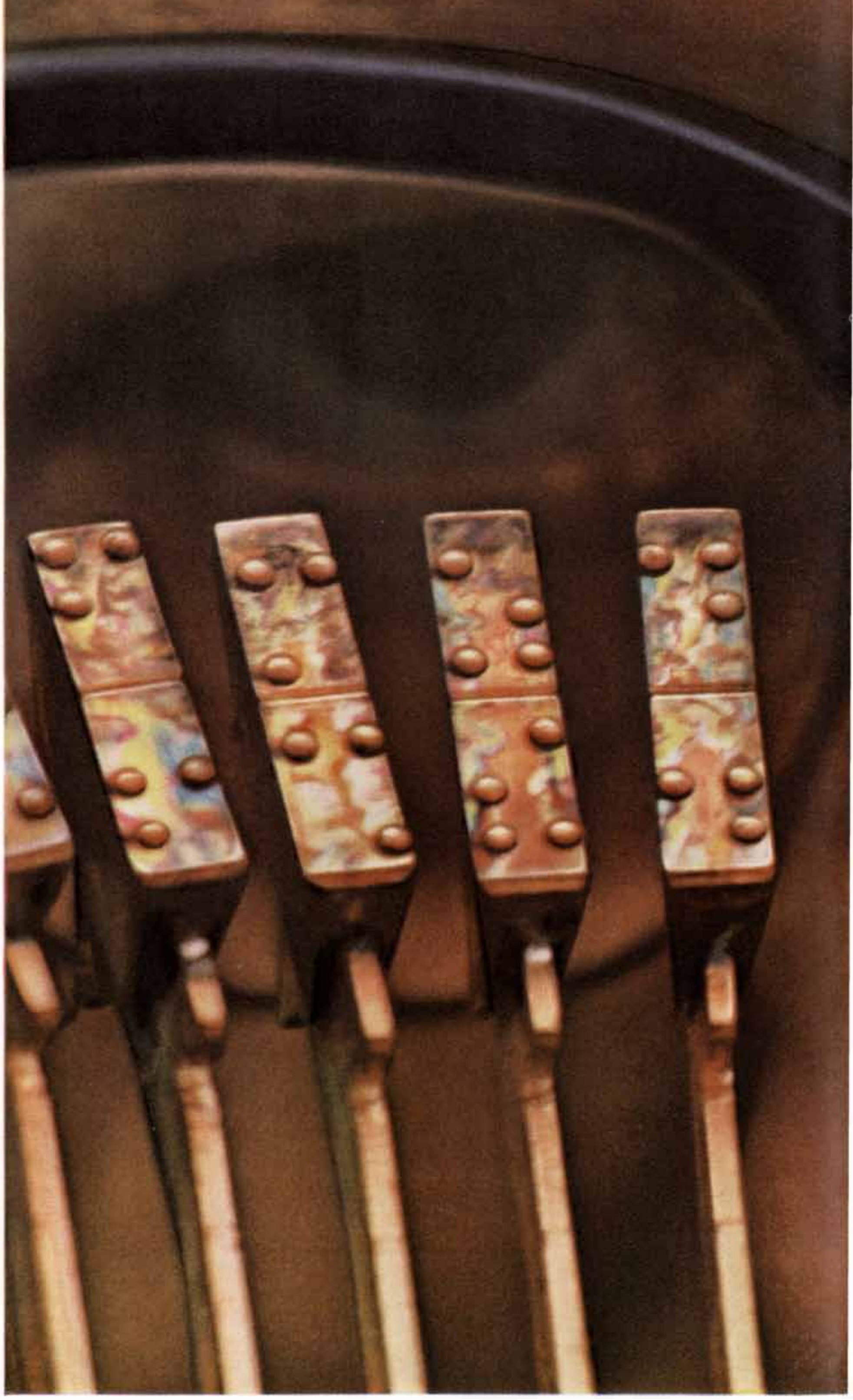
Although there is no evidence that the low-latitude zonal pattern has yet been broken, in response to rising solar activity, the computer prediction published by Willett and his two associates implies that it will break up this winter and be replaced by a "cellular blocking pattern." This is the second of two major patterns identified by Willett. It is associated with strong and persistent high-pressure areas that interrupt the zonal flow, forcing cold air to dip farther south and warm air to be carried northward. This pattern favors the northward movement of storms originating in the Gulf of Mexico and more northeasters along the Atlantic coast in winter; it also guides hurricanes up the Atlantic coast in summer.

The statistical model developed by DeVorkin and Prohaska is based on perhaps the most careful compilation of pressure, temperature and precipitation records yet made. It involved assembling atmospheric pressure records throughout the Northern Hemisphere for the past 70 years and temperature records over the U.S. for the same period. An analysis of these records showed a significant correlation between atmospheric pressure at particular places and times and subsequent temperatures at other locations. This suggests that temperatures at various localities can be predicted for several months in advance by the suitable analysis of recent pressure records at other localities. The study also shows that, if the temperature is known, the precipitation can be predicted. In making a prediction the M.I.T. group uses a computer to analyze some 400,000 items of weather data.

Various methods of making long-range forecasts have been devised in the past, without notable success. For periods beyond 12 days the best of them give results that are no better than random guesses. Willett and his co-workers expect their forecasts to hold up satisfactorily for four to six months.


PAN AM
FOR 1970:
A NEW FLIGHT
OF BIRDS
Soaring into the '70's far ahead of the rest... 1970 Thunderbird. With dramatic new front-end styling, shaped to slice the wind. Longer, lower and wider for '70. Yet, still uniquely Thunderbird. With its impressive list of standard luxury features you'd pay extra for in other cars. Options other cars don't even offer. And standards of quality most others only aspire to. Choose from three distinctive models. The New Flight of Birds is ready for take-off.
Above: Pan Am's® Boeing 747 Jet and the 1970 Thunderbird 2-Door Landau with Special Brougham interior.
THUNDERBIRD
Ford

Marsden Emig's inventiveness widened communication between the seeing and the blind.



Keys that open a new world between the seeing and the blind.

How could people who don't know Braille write to the blind? The problem baffled engineer Marsden Emig for three years. His story tells how people often benefit when IBM meets a challenge.

Close-up of the Braille keys adapted to an IBM electric typewriter.

"It didn't make sense to me," says Marsden Emig, an IBM engineer, "that a person had to know Braille in order to write to a blind person.

"The answer, I felt, was to adapt a standard electric typewriter so that anyone who could type, could type Braille. In this way, the advantages of Braille would be available to many more people.

"To print Braille, the type has to strike the back of the paper, and raise dots on the front. This is almost completely 'backwards' for a typewriter. But the idea fascinated me.

"In 1964, I decided to do something about it. I worked at home, in my spare time. After three years of plugging away at the problem, I finally developed the right Braille typewriter slugs and the necessary typewriter modifications.

"In 1967, I was able to type a letter to Fred Gissoni, a blind friend who was a consultant. The day after I wrote to Fred, I took the prototype to the office and proposed that my company manufacture the machine. Less than a year later, the first commercial IBM Braille typewriters were in use.

"Why put so much spare time into a Braille typewriter? Because a real need existed. An important need. And I was pretty certain that if I could figure out a practical way to do the job, IBM would work out a way to turn it into a reality. Today, the Braille typewriter exists."

IBM®


CONRAC

MAGNETIC RECORDING

The storage of information in magnetic particles is an integral part of communication and data processing. The demands of different kinds of information have given rise to variations on the basic theme

by Victor E. Ragosine

Magnetic recording plays a central role in modern communication and data processing, ranging from the "instant replay" of televised sports events to the memory of computer systems. Its basic principle is straightforward: sound, for example, is transformed into an electric current by a device such as a microphone. The current varies in intensity with variations in the loudness of the sound. The current passes through a coil surrounding an iron core, inducing proportionately varying degrees of magnetism in the core. A magnetizable surface such as a tape that is moved past the core is correspondingly magnetized in varying degrees. Since the magnetic particles in the tape retain their magnetization after they pass the core, the result is a recording. Playback, which re-creates the sound, is the same process in reverse: the magnetization in the tape induces varying electric currents in the coil, and the currents are transformed into a reproduction of the sound.

Two features of the process distinguish it from other forms of communicating and storing information. The first is that the recorded information cannot be read directly by a human being, as books, musical scores and pictures can. Since the information is stored in the form of the orderly orientation of magnetic dipoles in a magnetic surface, it has to be read by a machine, which ultimately transforms it into something that is comprehensible to the human ear or eye. Second, magnetic recording is versatile: it can handle sound and pictures, analogue data and digital data. It is therefore a means of communication not only between human beings but also between man and machine and between machines.

In order to make the magnetic-recording process feasible a number of technological hurdles had to be surmounted. The magnetic surface has to be manufactured with great care; the recording head and the playback head require precise machining to close tolerances, and the speed of movement of the tape must be meticulously controlled. I shall describe the major elements of the present technology and then touch on a promising new development that relies on electro-optical rather than mechanical techniques of recording and playback.

TAPE RECORDER for color television has its recording unit in the cabinet at right in the photograph on the opposite page. The screen at left monitors the recording, which here is a test pattern of color bars. Below the monitor is a circular cathode ray tube that is a color analyzer. The keyboard below the analyzer is for mixing input signals.

The Recording Process

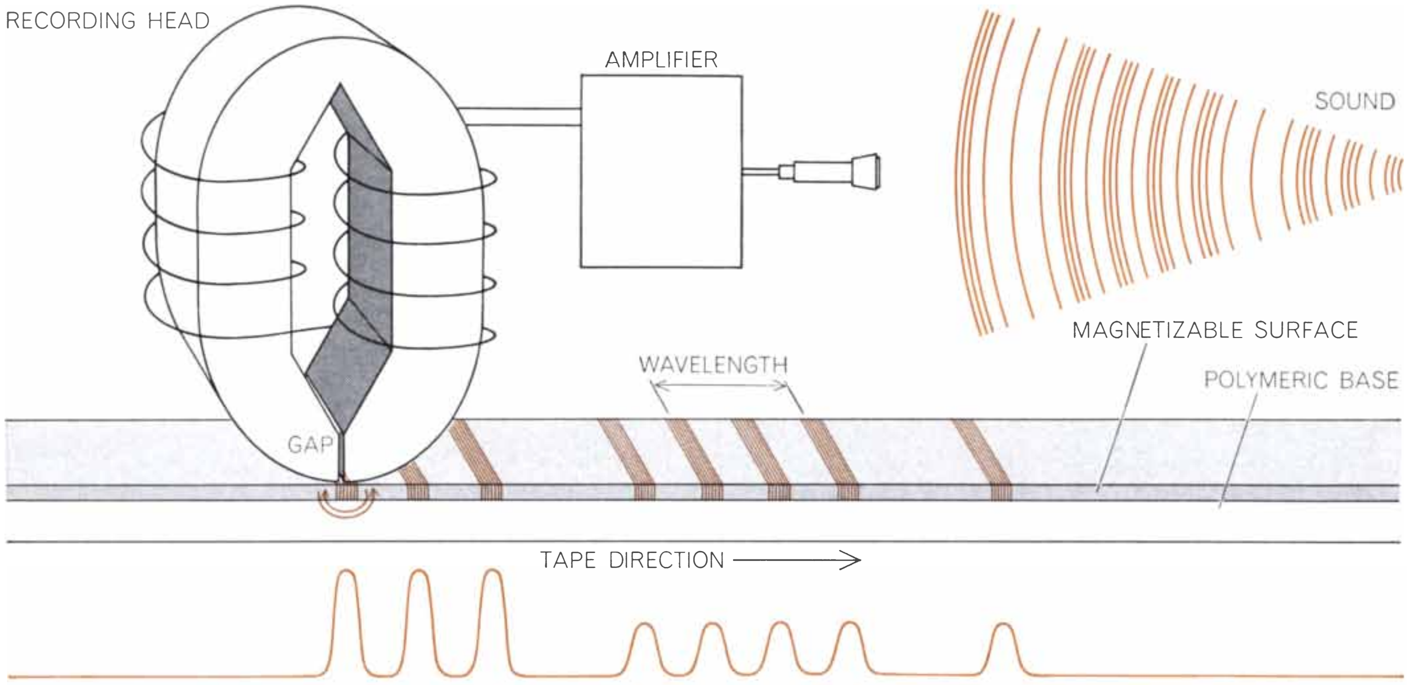
Let us begin by considering a recording process in which a strip of magnetic tape 1/4 inch wide mounted on a supply reel is moved past a recording head at a speed of 7½ inches per second and then rewound on a take-up reel. An arrangement of this kind would be typical of a tape recorder of good quality for amateur use in recording sound.

The iron core in the recording head is an electromagnet, that is, the core is readily magnetized when current flows through the wire wound around it and loses its magnetism just as readily when the current stops. The core is shaped in the form of a circle or rectangle, except that its north and south poles are separated by a small gap that is central to the recording process. Not all the magnetic lines of force extending from one pole to the other are straight; many of them take a looping path and thereby create what is called a fringing field [*see top illustration on next page*]. When the magnetic tape, moving across the gap and adjacent to the recording head, is in contact with the head, it provides an easier route for the lines of force to follow than air does. As a result most of the magnetic force crosses the gap by flowing through the tape.

The tape consists of a thin plastic base coated with an even thinner layer of magnetic oxide particles in a polymeric binder. The behavior of each particle under the influence of an external magnetic field is bipolar, meaning that there is no change in the magnetization of the particle until the strength of the field has reached a certain threshold level, whereupon the magnetization of the particle switches from one direction to the opposite direction. The threshold level is called the magnetic coercivity of the particle. It is determined by the shape of the particle and the magnetic properties of the particle material.

Imagine now a microscopic point on the tape as it approaches the recording head. The point senses an increasing magnetic field as it draws near the gap in the recording head. Nothing happens, however, until the field is strong enough to switch the particles of lowest coercivity. As the field increases, more and more particles are switched. If the field is sufficiently strong, every particle will be switched and the point on the tape will be magnetically saturated. As the point moves past the gap the magnetic field decreases and eventually has no further effect on it.

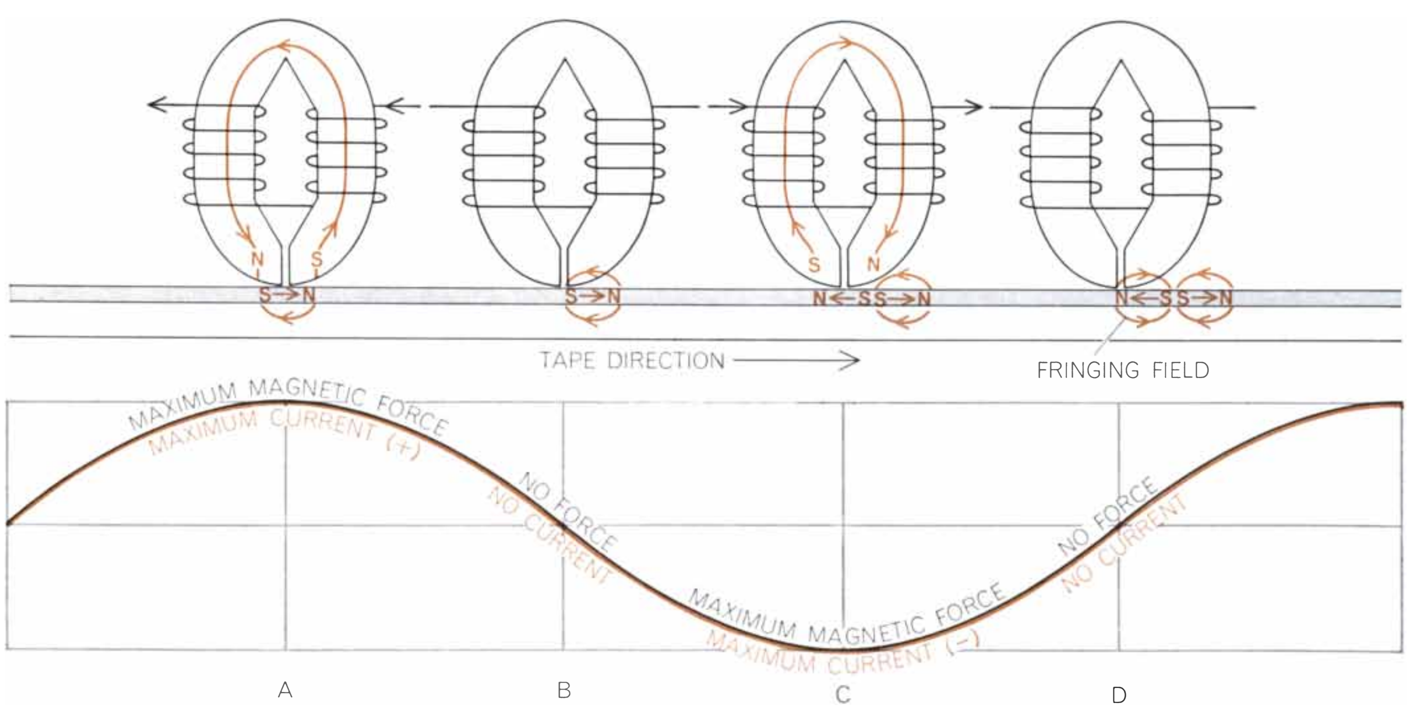
In the recording process the current applied to the recording head varies constantly in reflection of the variations in



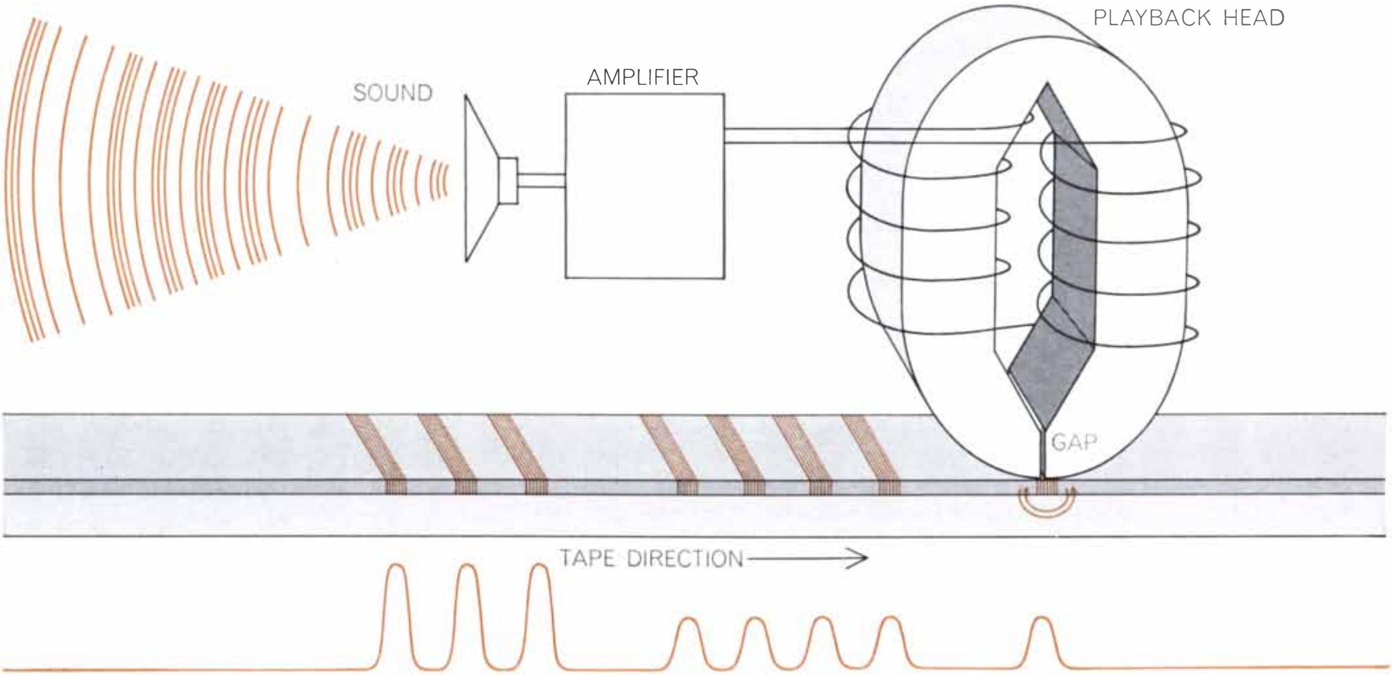
RECORDING PROCESS for sound entails the conversion of the sound into an electric current that flows through the coil surrounding the recording head, inducing magnetism in the core of the head. Magnetic lines of force crossing the small gap between the north and south poles of the head flow through the tape as it is moved past the head, thereby imposing an orderly orientation on the randomly oriented particles of magnetic oxide on the surface of the tape. The degree of magnetization varies in intensity with variations in the strength or frequency of the original sound. Pictures and analogue or digital information can also be recorded.

the strength or frequency of the signal it is receiving. Hence the microscopic point on the tape senses magnetic fields that vary with time. The point is left in a state of magnetization determined by the last field that was of sufficient magnitude to produce a change in magnetization.

Once this point has passed the recording head the magnetic lines of force extending from each dipole on the tape tend to form closed loops through the air. If a playback head—a magnetic structure much like the recording head—is now placed in the path of the lines of force, the lines will be closed through the head rather than through the air, since the head has a much lower magnetic impedance than the air. As long as there is no relative motion between the playback head and the tape there will be no induced current in the coil wound on the head. If there is such motion, a current will be induced to flow in



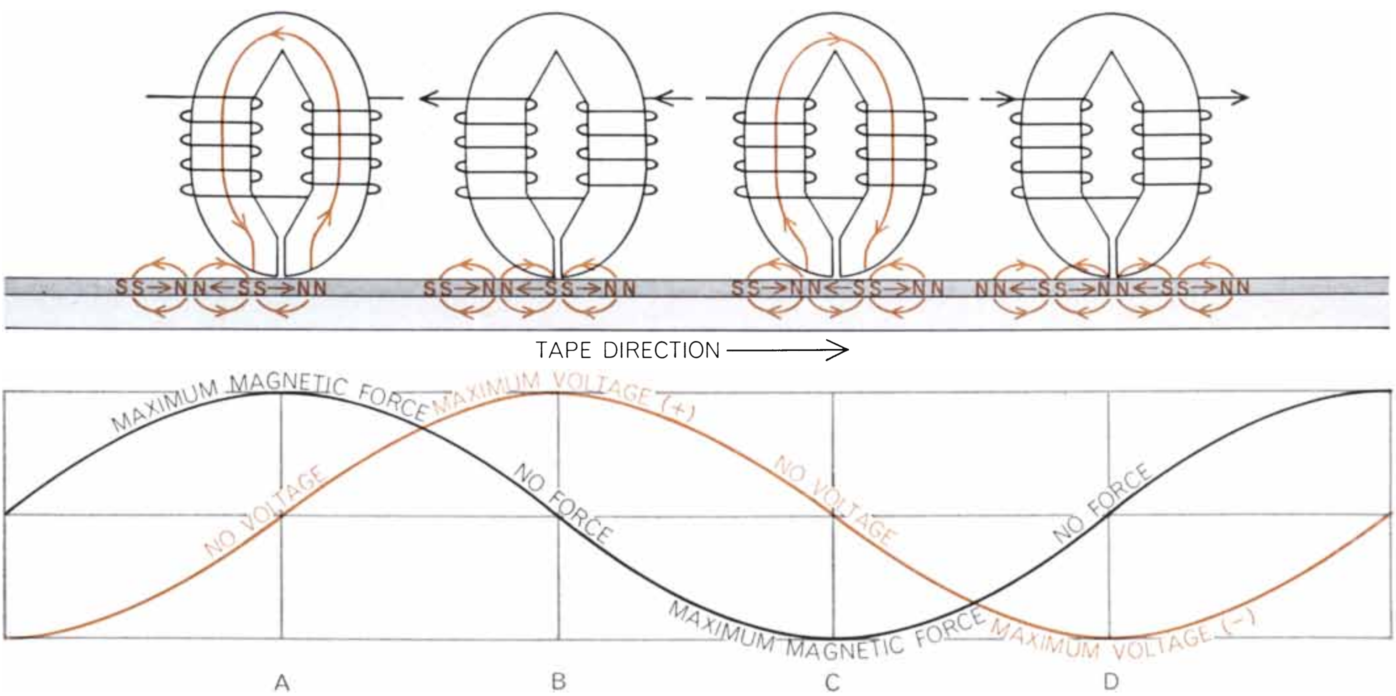
MAGNETIZATION OF TAPE during recording creates a series of small dipole magnets on the tape. They reflect the alternating current passing through the coil. At maximum positive current (*A*) the lines of force point in one direction. At point *B* the current entering the coil is zero at its point of alteration and creates no magnetization, so that a small segment of the moving tape is not magnetized. The surge of maximum negative current at *C* creates a magnetization of polarity opposite to that imposed on the tape at *A*.



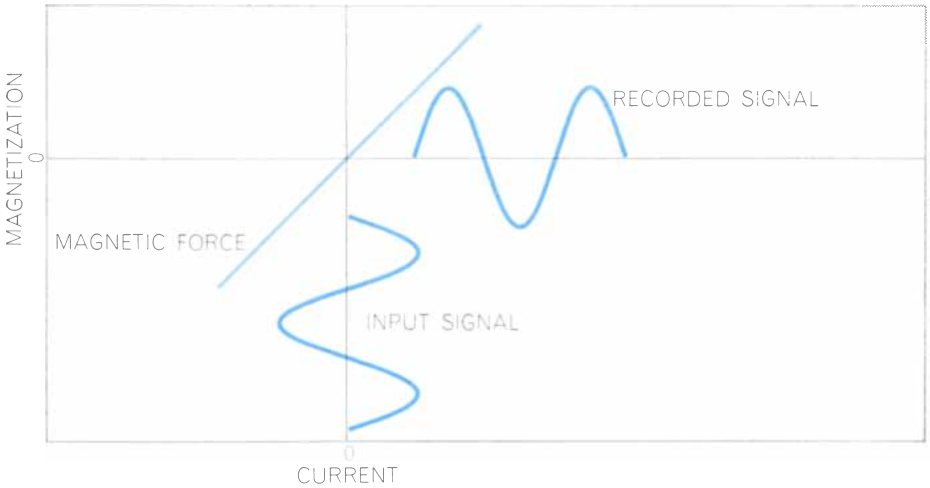
PLAYBACK PROCESS is the reverse of the recording process. The tape, which carries in the orderly orientation of its magnetic particles a record of the original sound, is moved past a playback head that is similar to the recording head. The lines of magnetic force emanating from the segment of tape that is under the gap of the head close through the head. Because the tape is moving, and the motion of a magnetic field through a coil induces an electric current in the coil, the coil around the playback head carries a current reflecting the magnetization of the tape. Amplification of the current yields a reproduction of the sound that was recorded.

the coil. This is the mechanism of playback.

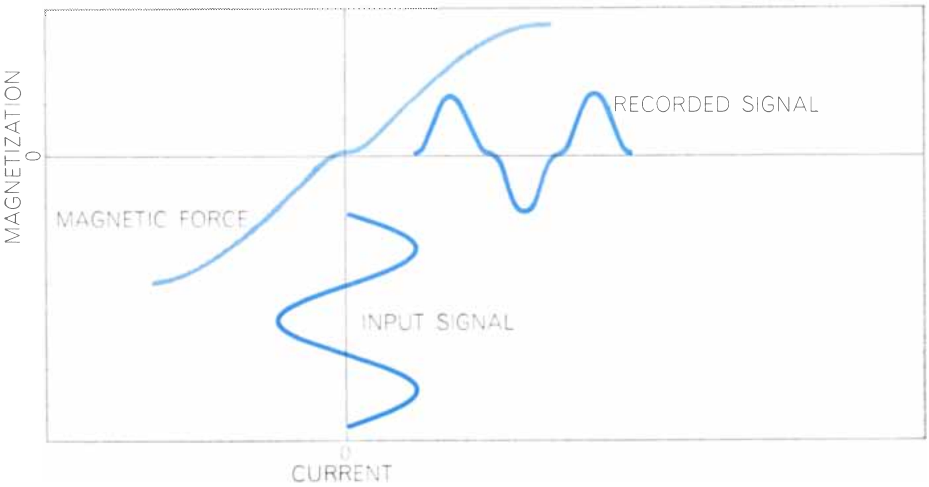
In the recording and reasonably faithful reproduction of sound, and particularly of analogue data, a high degree of linearity is required, that is, the output of the magnetically recorded surface has to be (as nearly as possible) directly proportional to the input. To put it another way, if music is being recorded and the music suddenly doubles in volume, the signal obtained on playback must also double in strength, otherwise the fidelity of the recording will be poor. In order to achieve such linearity it is necessary to employ the technique known as alternating-current bias. The essence of this technique is that at the recording head an alternating current is added to the signal current and imposed on the recording surface. The added current is much higher in amplitude than the signal current and several times higher in frequency than the highest frequency



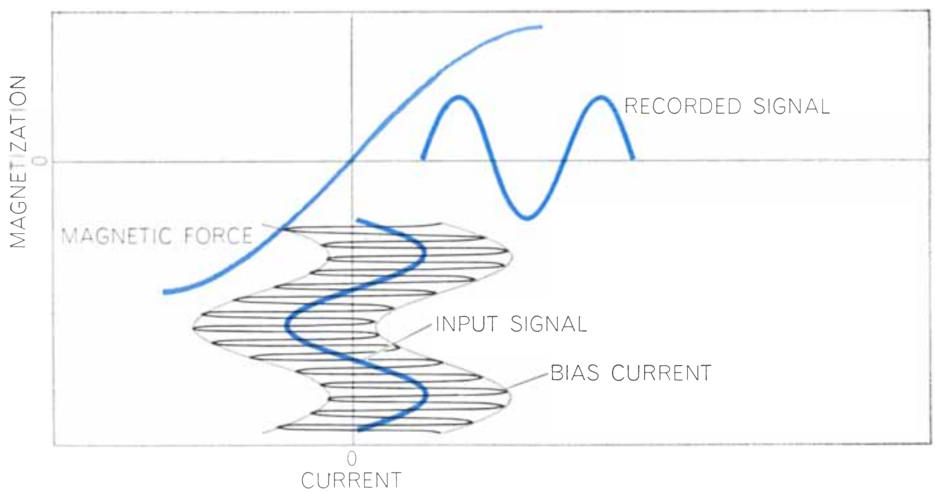
FLOW OF CURRENT during playback arises from the opposite polarities of the magnetic fields on the tape. The recording process produced on the tape a series of short, permanent bar magnets. During playback the motion of the tape has the effect of moving this series of magnetic fields of opposite polarity through the coil, thus producing an alternating electrical voltage. The peak surge of electric current comes at the time when the polarity of the magnetic field is changing most rapidly, as is the case at points *B* and *D*.



IDEAL PATTERN of recording is depicted by sine-wave curves. When an increasing magnetic force, represented by the slanting line at top left, is applied to the tape and then removed, a certain amount of magnetic flux will remain in the tape. In a linear recording the remanent flux would be exactly proportional to the input, so that the sine-wave curve representing the input would be matched by the curve representing the magnetic pattern on the tape. The result would be a recording that reproduced faithfully the original sound.



ACTUAL RECORDING in the absence of special treatment produces distorted reproduction because an increase in magnetic force is not precisely duplicated on the tape. Hence the sine-wave curve representing the recording does not match the input-signal curve.



BIAS CURRENT is put through the recording head with the input signal to reduce distortion. The bias is an alternating-current signal much higher in frequency and amplitude than the basic input signal. In effect the bias current stimulates the magnetic particles in the tape so that they are more responsive to the signal produced by the sound being recorded.

of the signal that is being recorded.

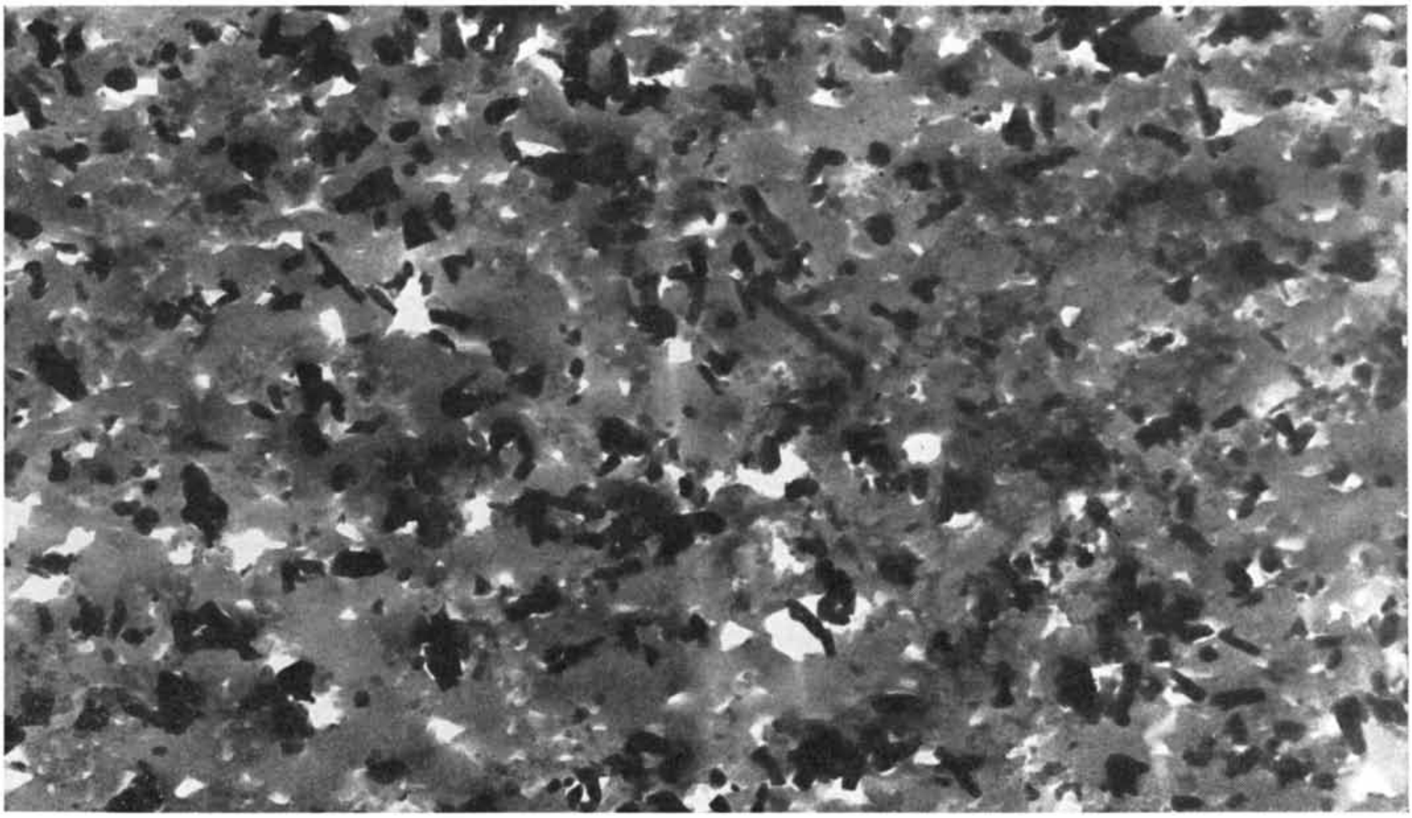
One can depict the problem with a graph that has a vertical axis indicating the strength of the recording signal and a horizontal axis showing the amount of magnetic field remaining in the tape at any given strength of signal. The graph reflects the fact that a certain amount of magnetism remains in the tape after the magnetic field that created it has been removed. This remanent flux of magnetism is what creates the playback signal. In the ideal linear response a recording signal producing a given magnetic force would result in an equal remanent flux. If the recording signal were doubled, the remanent flux would be doubled. The fact is that without alternating-current bias the response is not linear, and one hears distortion on playback.

When a bias current is applied, the magnetic particles in a spot of tape are subjected to many cycles of the current as they cross the gap of the recording head. In effect they are shaken up, so that they respond more readily to being magnetized by the signal one wants to record. They leave the head with a magnetization that is much more nearly proportional to the signal [*see illustrations at left*]. Alternating-current bias thus provides low distortion.

Linearity is required in the recording of signals such as music because the information being recorded is contained not only in the frequency of the signal but also in the changes in its amplitude. Linearity is not a requirement in other kinds of modulation and coding. A case in point would be the recording of a signal that was frequency-modulated on a carrier signal; the carrier signal might have a frequency of a million cycles per second and the modulations might cause it to deviate by 200,000 cycles, that is, to a high of 1.2 million and a low of 800,000. The information is contained in the deviation of the frequency from the carrier signal and in the rate of change of the deviation. One need only be able to detect the number of times the signal crosses the zero axis (the carrier signal) in a given period of time, and that requirement does not call for linearity.

The same is true of digital coding and other codes that take advantage of such phenomena as changes in output level and changes in polarity. A simple digital code, for example, might use a reversal of magnetic flux to indicate 1 and a lack of reversal to indicate 0. In this case it is necessary only to distinguish the difference between two states.

The simplest mechanical system for



MAGNETIC TAPE appears in this electron micrograph at an enlargement of 20,000 diameters. What one sees is an end view of a piece of tape, showing part of the surface containing the particles of magnetic oxide that are magnetized by the recording process. The randomly oriented black objects are the oxide particles, the gray areas are the polymeric binder and the white areas are voids.

magnetic recording consists of two reels, together with guides that position the tape with respect to the recording head. The speed of the reels must be amenable to individual control, because a reel that is almost empty will necessarily have a higher rotational speed than one that is almost full. This control can be achieved by providing a holdback tension on the supply reel while pulling tape with the take-up reel.

A simple reel-to-reel system is unsatisfactory for any but the most primitive recorders. The reason is that any perturbation in the motion of a reel will be transmitted immediately through the tape, and since the playback head is sensitive to rates of change of flux, any variation in tape velocity as the tape moves through the head gap will produce a variation in the signal.

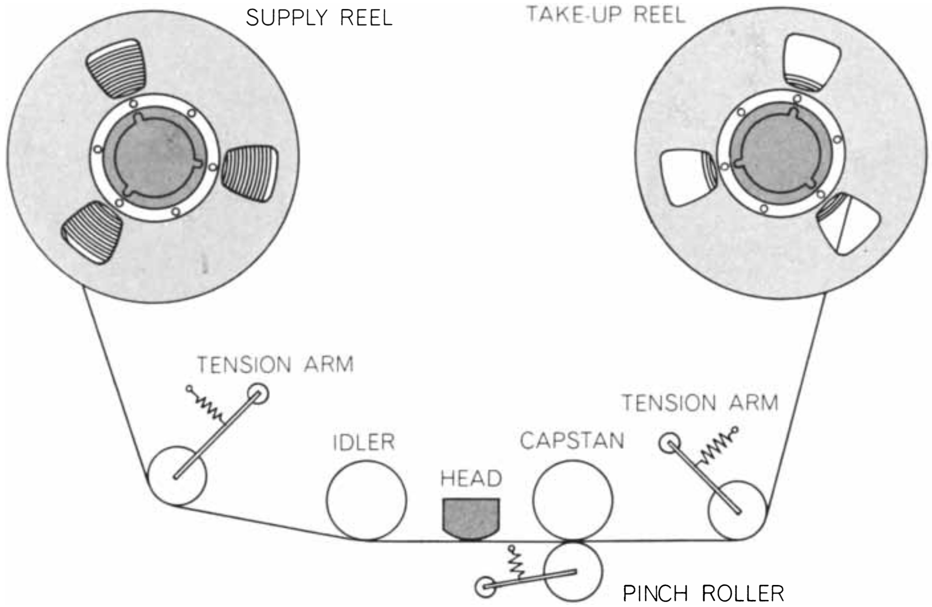
The Control of Tape Speed

In order to provide a measure of control over the velocity of the tape in the vicinity of the head, capstans, or rotating cylinders, were introduced near the head. In early machines a pinch roller pressed the tape against each capstan and constrained the tape to move at the circumferential velocity of the capstan. In order to prevent high-frequency oscillations capstans were loaded with a high-inertia drum, acting as a flywheel, that would not respond readily to small perturbations.

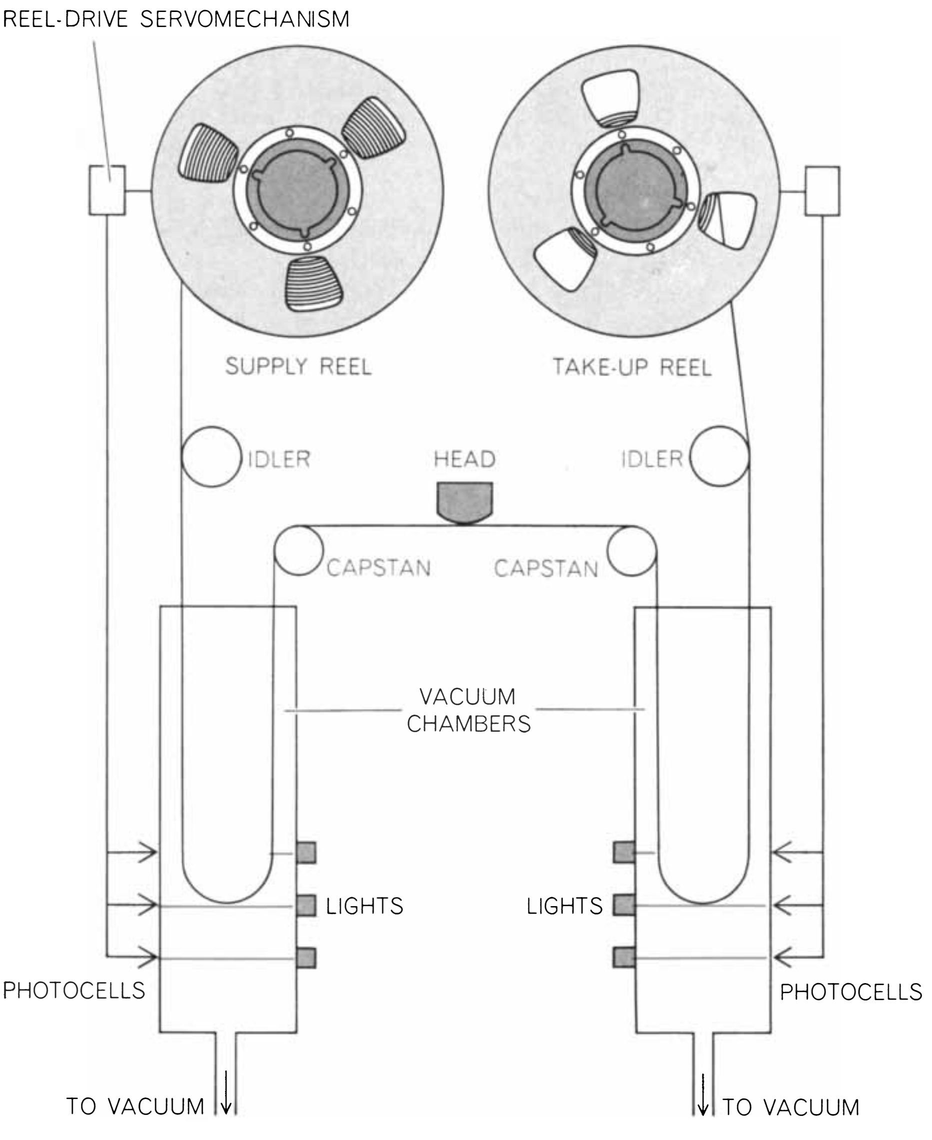
High-inertia systems with the velocity of the tape controlled in the vicinity of the heads by means of single or double capstans were quite satisfactory for sound recorders when the signals being recorded had fairly long wavelengths and correspondingly low frequencies. If one wanted good quality in the recording of signals of short wavelength and high frequency, however, or if one sought to record on one machine and play back on another, thereby requiring a recording that was compatible between machines, the control of tape speed with high-inertia capstans was inadequate. The solution to these difficulties was optical control. Light passing through one or more holes through the capstan produced a signal showing the speed of the capstan. That signal was compared with a reference signal to provide a means of regulating the speed of the capstan. Another technique is to record a control track on the tape. The velocity of the tape during playback is governed by the frequency recorded on the control track. In this way it is possible to compensate for variations in the speed of recording by introducing comparable variations in the speed of playback.

Even with these techniques of control it is possible for fluctuations in the velocity of a reel to be transmitted through the tape and to result in variations of velocity at the heads. It is therefore necessary to isolate a reel and its servomechanism from the corresponding capstan and its servomechanism. Modern machines employ a vacuum chamber below each reel. A segment of tape hangs down from the reel, forming a loop in the chamber [*see bottom illustration on next page*]. The partial vacuum provides an environment of constant pressure so that changes in the position of the loop can be sensed accurately. The position of the loop is sensed by an arrangement consisting of an array of lights on one side of the vacuum chamber and an array of photocells on the other. The signal from the photocells is proportional to the length of tape in the vacuum chamber. This signal controls the speed of the reel.

Vacuum chambers and tension arms can serve an additional purpose in tape-recording systems that require rapid starting and stopping. A case in point would be a system of digital recording serving as the memory for a computer. Typical starting and stopping times are



MECHANICAL SYSTEM for a tape recorder typical of the kind that an amateur might use for recording voice or music has devices to control the speed of the tape past the recording and playback head. In particular the pinch roller presses the tape against the rotating capstan and forces it to move at the controlled circumferential velocity of the capstan.



VACUUM CHAMBERS provide control of tape speed in advanced recorders. The partial vacuum in each chamber makes it possible to sense accurately the position of the loop of tape hanging down from the reel. Tape position is sensed by photocells, which provide signals that control the servomechanism driving the corresponding reel. Reels must be isolated to prevent perturbations in their velocity from affecting speed of the tape at the head.

measured in milliseconds. A reel motor, which has a large amount of inertia, cannot be brought to a high speed in such a short period of time. Rapid starting is achieved by energizing a low-inertia capstan, which thereupon begins to pull tape out of a vacuum chamber. The reduced length of tape in the vacuum chamber then produces a signal that energizes the reel motor.

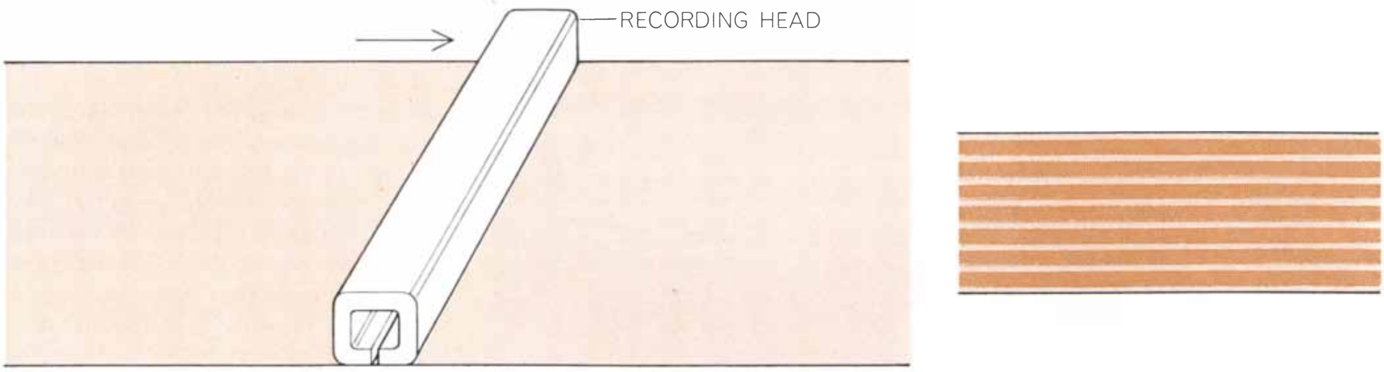
Formats for Recording

Most systems for recording sound or signals from measuring instruments employ a longitudinal format: the recording is made in one or more tracks on tape that moves from one reel to another past the recording head. Track widths vary with the requirements of the system. A typical sound-recording system will involve two or four tracks on 1/4-inch tape, whereas digital or instrument-signal recorders may use seven or nine or more tracks on 1/2-inch or one-inch tape.

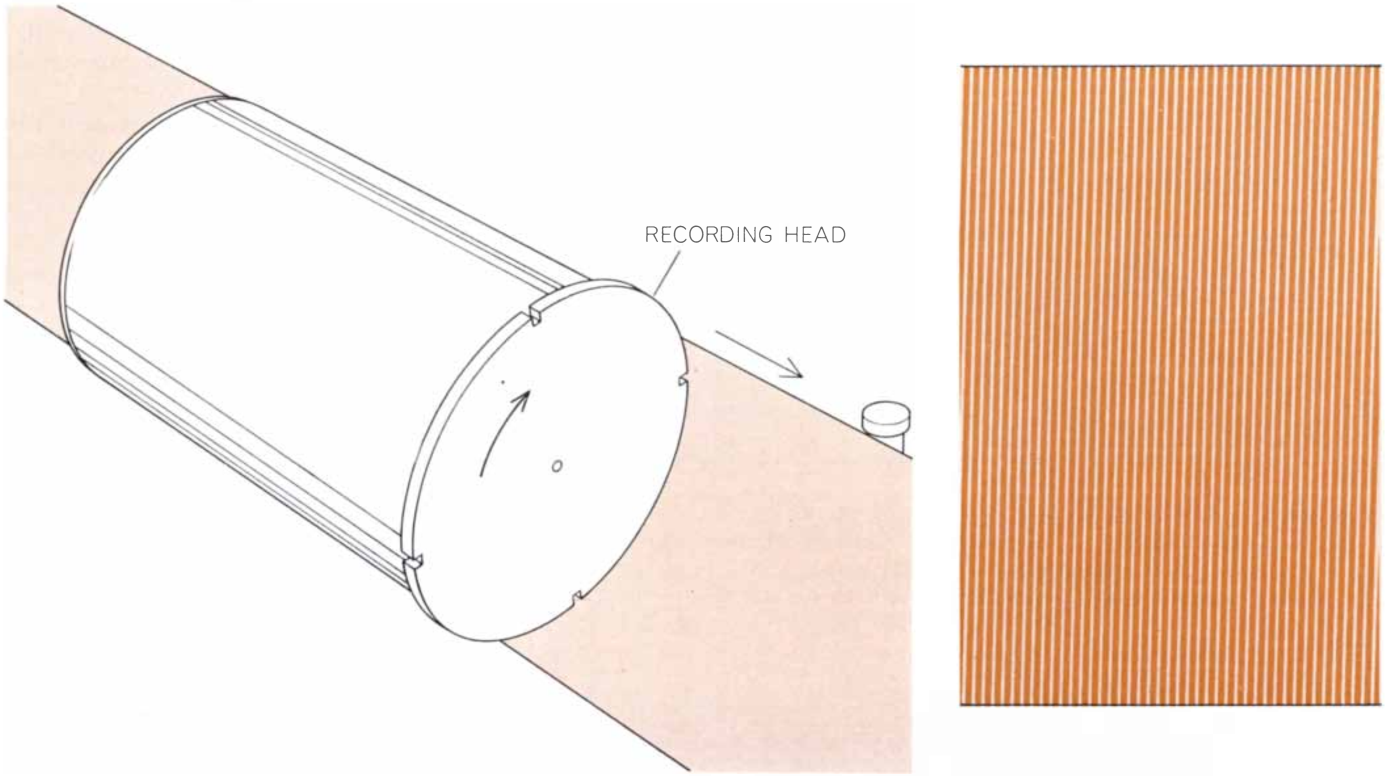
In most longitudinal recorders mechanical considerations and the need to use tape at an economically reasonable rate impose limitations on the speed at which tape can be moved. Apart from novel experimental systems, the highest tape speed in commercially available instrumentation recorders is 120 inches per second. Tape transports for use in computers have been built with speeds of up to 1,000 inches per second.

The consumption of tape is of course related to unidirectional playing time. At a speed of 120 inches per second a 10½-inch reel holding 3,600 feet of tape will provide six minutes of unidirectional playing time. At this speed, if the shortest wavelength to be recorded is 60 microinches, the highest frequency that can be recorded is two megahertz (two million cycles per second).

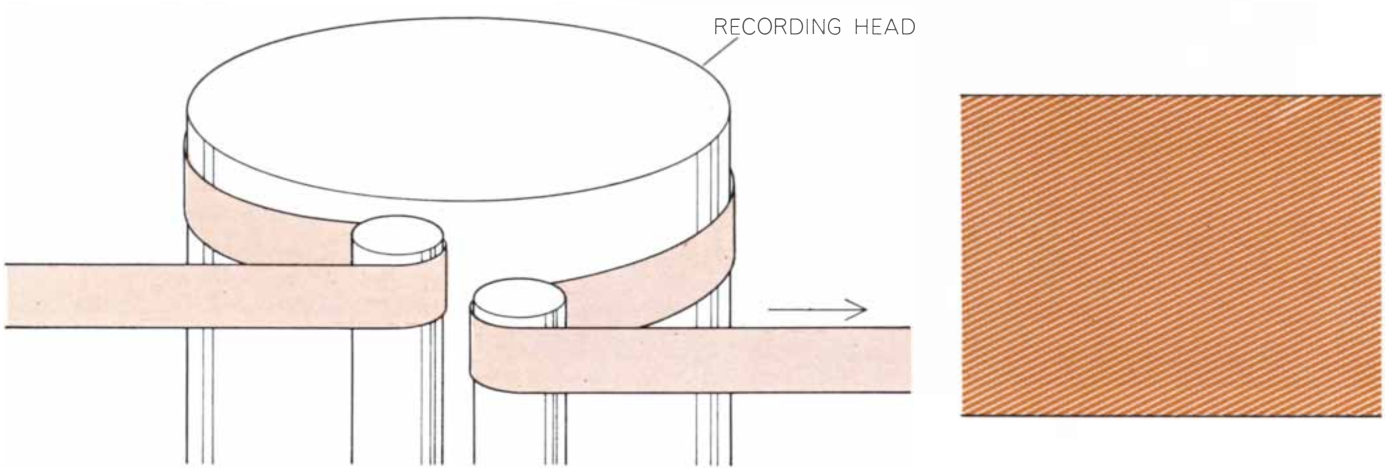
In order to make possible the recording of higher frequencies and to increase the playing time for a given length of tape, other recording formats are employed. One is transverse recording, which entails recording across the width of a tape rather than along its length [*see middle illustration on opposite page*]. In transverse recording the tape is moved at a relatively low rate of speed; in professional television transverse recording is normally done at a speed of 15 inches per second. The heads for the picture portion are mounted on a drum that is rotated in a direction perpendicular to the movement of the tape, while the tape is constrained to follow an arc and maintain contact with the heads. The sound track and the control track



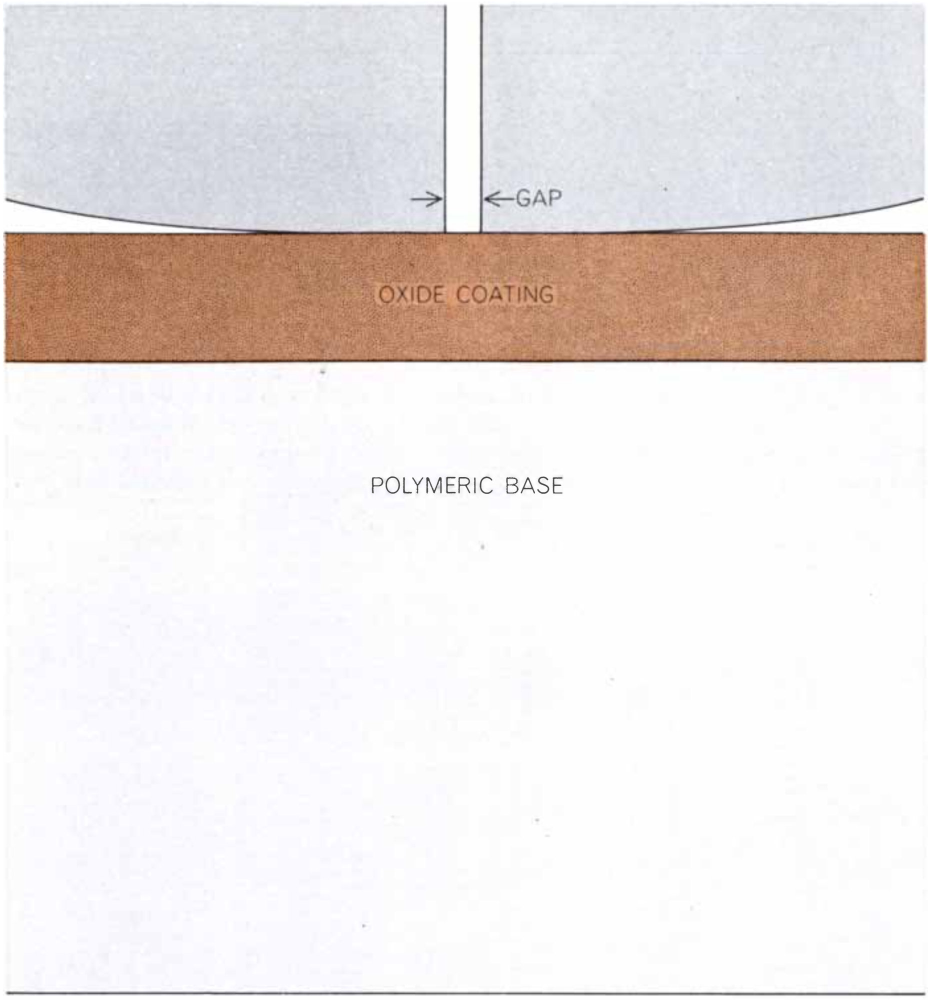
LONGITUDINAL FORMAT is employed in most tape systems for recording sound. The tape, which is kept flat, is moved from one reel to another past the recording or playback head. At right is a typical set of longitudinal tracks, in this case seven of them. On an actual tape 1/2 inch wide each of the recorded tracks would be .05 inch wide and would be separated from an adjoining track by .02 inch. The part of the illustration showing the tracks is at the same scale as the corresponding part of the two illustrations below.



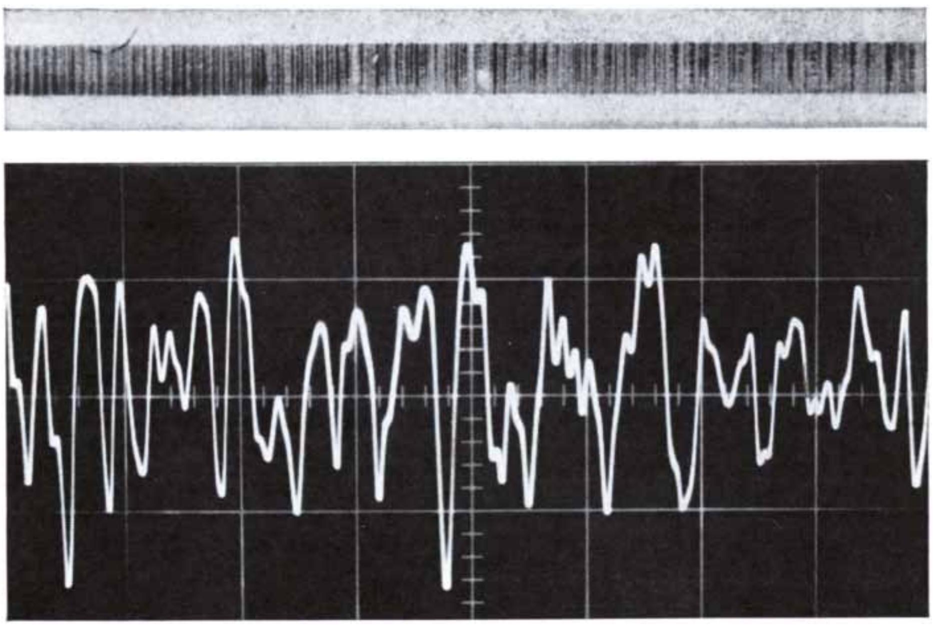
TRANSVERSE RECORDING entails recording across the diameter of a tape instead of along its length. The recording head is a revolving drum. Because of this arrangement it is necessary to hold the tape in the shape of an arc as it passes the recording head. The system is used for recording television pictures. A tape two inches wide with .01-inch tracks and .005-inch spacing is represented.



HELICAL RECORDING employs a revolving drum with the tape wrapped around it. The drum contains the recording head and the playback head. A helical system requires less elaborate controls than a transverse one but yields recordings of somewhat poorer quality. Such systems are used for recording television pictures when the resolution need not be up to commercial standards.



CLOSE VIEW of the relation between a recording or playback head and the tape indicates the dimensions involved. The plastic base of the tape is 1,000 microinches thick and the coating of magnetic oxide particles is 200 microinches thick. The head gap is 60 microinches. On this scale the thin tape actually has the behavior of a flexible structural beam.



MAGNETIC PATTERN on a tape is brought out (*top*) by applying a colloidal solution of magnetic particles to the surface of the tape. The electrical signal corresponding to the pattern appears in the oscilloscope tracing at bottom. These patterns result from the recording of the opening of Beethoven's Quartet No. 12 for two violins, viola and cello.

are recorded longitudinally with stationary heads.

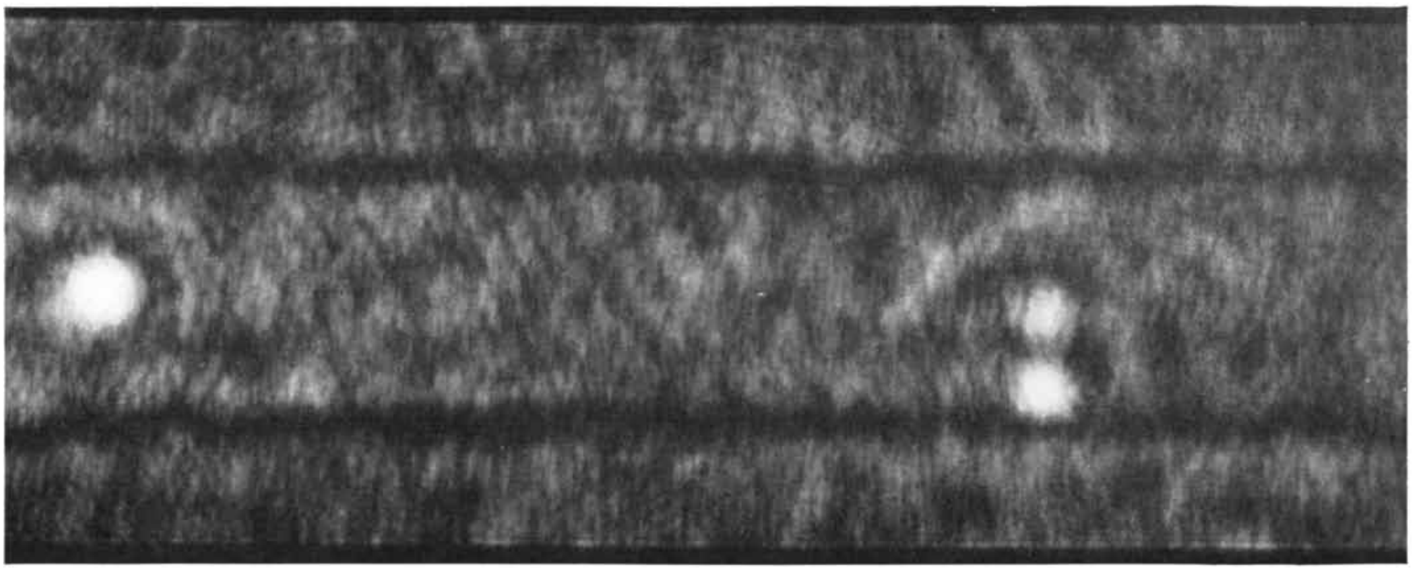
The transverse format involves a discontinuity of recording, inasmuch as the recording is done line by line. For recording television pictures the discontinuity presents no problem; indeed, the transition from line to line on the tape can be done during the "blanking" period when the electron beam in the television receiver, which also operates on a line-by-line basis, is repositioned to begin a new line. If the transverse format is to be used for recording where a continuous stream of information is required, some redundancy or overlap must be provided. The schemes employed for this purpose slowly switch and mix signals from one track to the next so that the recording of information is effectively uninterrupted.

With a transverse arrangement the consumption of tape is determined by the longitudinal speed of the tape, as with longitudinal recorders, except that in transverse recording the head-to-tape speed can be substantially higher than in longitudinal recorders. The result is a corresponding increase in the amount of information that can be recorded with good fidelity. Modern recorders used for color television record signals in excess of 15 megahertz at a head-to-tape speed of 1,560 inches per second. A 14-inch reel of tape can provide an hour of unidirectional playing time.

Helical Recording

Although transverse recording offers the advantages of high frequency and reasonable tape consumption, it also imposes the disadvantage of having to control both the longitudinal speed of the tape and the rotational speed of the recording head. In addition, the complexity of the head assembly increases cost. Helical recording provides a compromise between the longitudinal and the transverse method. In the helical format the tape is wrapped around a drum that contains the recording head and the playback head [*see bottom illustration on preceding page*]. This format has given rise to machines that cost substantially less than machines for transverse recording but do not perform quite as well. Helical recorders have found increasing use in television applications where picture recording can be at a lower resolution than is acceptable in commercial television.

Magnetic recording need not be only on tape; it can also be done on chips, on relatively short strips and on disks. With



MAGNETO-OPTICAL METHOD was used to record two "bits" of digital data, which appear as dark parallel tracks in this photomicrograph of the thin film of metal alloy on which the bits were recorded. Each track is 50 microns wide. The white spots at left and right are imperfections in the metal alloy film. This method of recording takes advantage of the fact that the response of certain materials to a magnetic force can be changed by the absorption of radiant energy such as light, which in this case was light from a laser.

chips and strips the recording is usually longitudinal. After the chips or strips are retrieved from storage they are put on a rapidly rotating drum that provides the necessary motion between the head and the recording surface. With disks the head-to-surface motion is provided by rotating the disk, and information is recorded in circles.

Disk systems are widely employed in digital-computer memories. The disk can be rotated at high speed and the recording head moved quickly across the radius of the disk, providing rapid access to all the information in the recording. Air-bearing heads are used to reduce wear on both the head and the disk; instead of the two being in contact there is a thin film of air (about 100 microinches thick) between them. If simultaneous access to several tracks of data is required, several disks with separate recording heads can be ganged on a common shaft or many heads can be used on one disk.

Recently disks have come into use for the recording of analogue data and television pictures. In video recording one head is ordinarily employed; it is moved on a carriage to selected tracks. This is the basis for the instant replay that has been such a boon to the watchers of sports events on television. Air bearings are not used in television disk recorders because space between the head and the recording surface results in a weaker signal. At the high frequencies and short wavelengths required for television the loss arising from a separation of even 100 microinches is unacceptable.

There are several physical considerations that bear on the efficiency of a magnetic recording process. They include signal-to-noise ratio, bandwidth and harmonics. Noise is any unwanted signal. It is inescapable in a recording process; the aim is of course to minimize it in proportion to the strength of the wanted signal. Bandwidth is the range of frequencies in a particular recording; if the lowest frequency is 10,000 cycles per second and the highest is 12,000, the bandwidth is 2,000. Harmonics are integral multiples of the fundamental frequency of a sound; they can be considered as noise if they are unwanted.

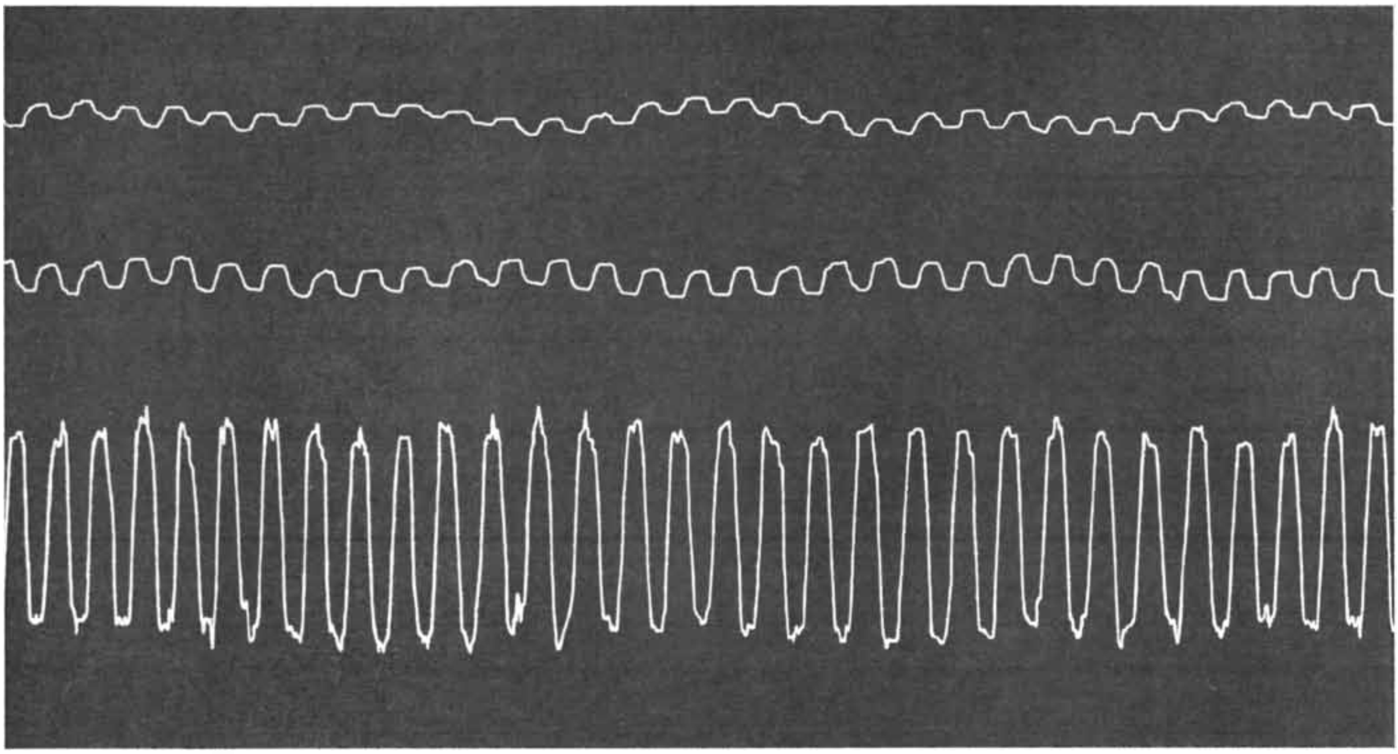
Machine Efficiency

The efficiency of a tape recorder as a communication channel is determined by the signal-to-noise ratio and the bandwidth that can be achieved. The recorder's efficiency in storing information is determined not only by signal-to-noise ratio and bandwidth but also by the density with which the information can be physically packed. In linear systems the maximum signal that can be recorded is determined by the portion of the recording system that approaches linearity. Linearity or deviation from linearity is generally specified by the third-harmonic content of the output signal. A typical specification calls for a distortion of not more than 1 percent in the third harmonic.

The linear range that a tape is capable of recording is determined by the magnetic properties of the recording medium. In tapes made with magnetic powders the limit is determined not only by the intrinsic magnetic properties of the particles but also by the way the particles are packed on the tape and by the magnetic interactions of the particles. In systems where linearity is not required the maximum output signal is determined by the efficiency of the recording process, by the thickness and magnetic properties of the recording medium and by the spacing between the tape and the recording or playback head.

The loss of signal due to spacing between the tape and the head increases rapidly as the spacing increases. It is the principal limit of the shortest wavelengths that can be recorded. In modern recording it is common to encounter signals with wavelengths as small as 60 microinches. A 60-microinch wavelength corresponds to about three wavelengths of blue light. A spacing of as much as 20 microinches between head and medium—equivalent to one wavelength of blue light—can cause a sevenfold loss in the reproduced signal.

In order to minimize the loss due to spacing, the most intimate possible contact between head and tape is required. On the other hand, it would be desirable to introduce a lubricating film between the head and the tape in order to reduce wear and drag, but it is impractical to do so because of the loss that even minuscule spacing produces. Hence every effort is made to keep head and tape in close contact and to try to minimize wear and friction by making tape surfaces smooth and by putting lubricants in the



DIGITAL INFORMATION recorded on a film of metal alloy by magneto-optical methods and read out by means of plane-polarized light is represented by oscilloscope tracings of the readout. The top two lines show the output of split beams before the application of noise-cancellation procedures designed to enhance the signal. The enhanced signal is represented by the bottom tracing.

tape rather than between the tape and the head.

As long as head-to-tape speed is low, reasonably intimate contact between head and tape is limited only by irregularities in the surface of the head and the surface of the tape. The contact can be maintained without applying a strong tension to the tape. At high tape speed, however, air is dragged into the interface between the tape and the head, creating an unwanted air bearing. If tension is increased to squeeze out the air, serious penalties such as wear may be incurred. Great care is therefore taken in shaping heads so that the tendency for an air film to form is minimized.

The limitation that spacing loss imposes on wavelength has important implications for bandwidth and packing density. The implication for bandwidth arises in the following way. The head-to-tape speed is the product of the recorded wavelength and the frequency. If the shortest wavelength yielding an adequate signal is determined by the mechanical spacing of the head and the tape, the recording of high frequencies requires an increase in the relative head-to-tape speed. The increase in speed creates new mechanical problems in accurately moving and positioning the tape with respect to the head.

Linear packing density is determined by the shortest wavelength that can be recorded and played back. Area packing density is determined not only by the wavelength but also by the width of the track and by the spacing between tracks. Here the tolerable values of signal-to-noise ratio impose limitations. Since the magnitude of the signal is determined by the rate of change of magnetic flux, wide tracks produce a larger signal. Because a wider track has a larger population of particles, the noise is reduced.

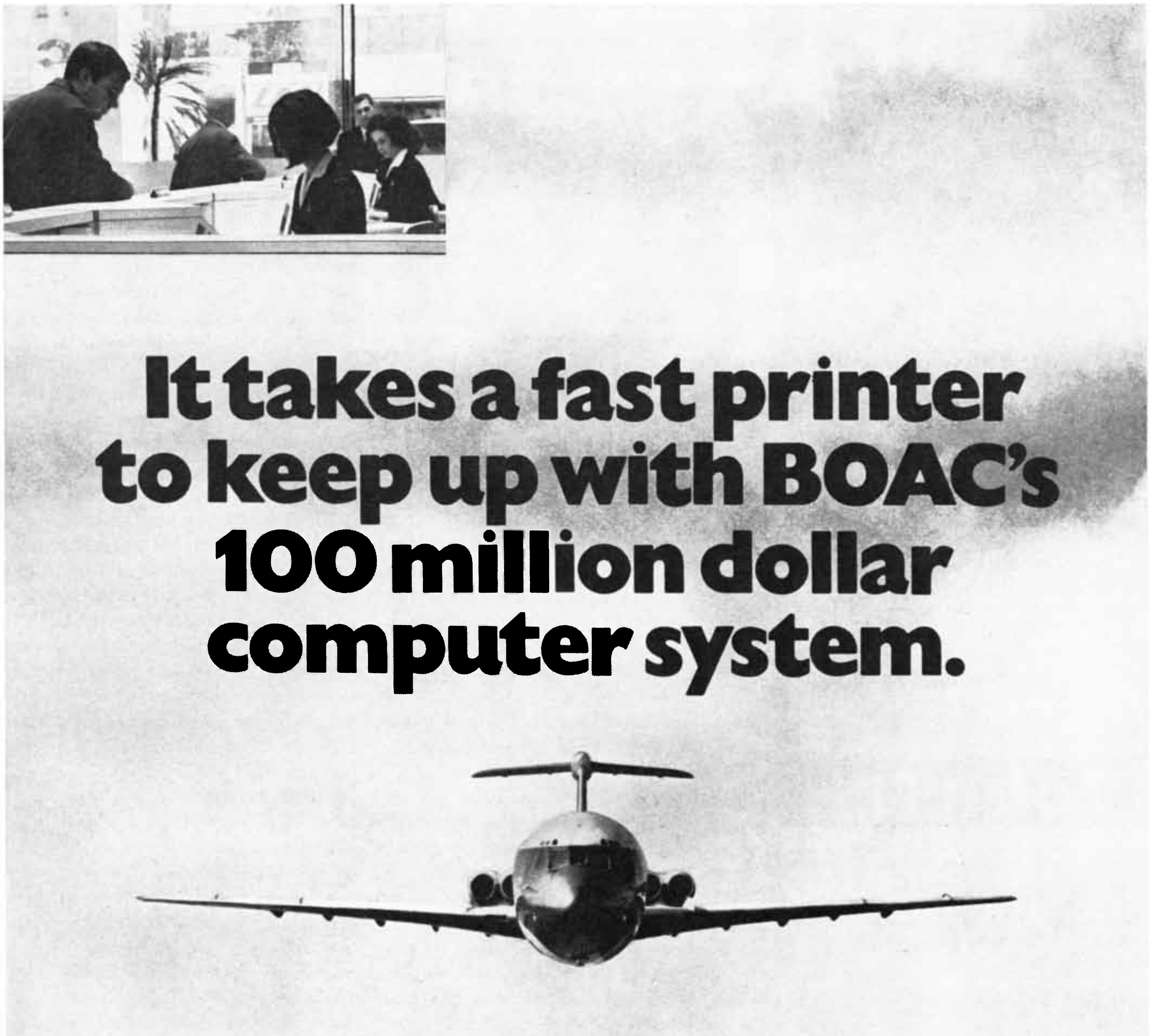
Tracks that are too close together will give rise to cross talk between adjacent tracks. Cross talk of course represents noise. Hence the avoidance of cross talk imposes limitations on the spacing between tracks.

The volumetric information-storage efficiency of a recording medium is also determined by the total thickness of the medium. The thickness of the magnetic surface can be very low: typical coatings on magnetic tapes range from a hundred-millionth of an inch to 400 millionths of an inch. Plated coatings on disks for digital recordings can be as thin as five millionths of an inch. Volumetric efficiency is further affected by the thickness of the base holding the recording medium: the thicker it is, the less tape can be stored in a given space. The manner of storage also makes a difference.

The most efficient volumetric packing density is obtained with magnetic tape rolled on a reel. Chips and strips, which must be stored in a pile with spacers interleaved, are less efficient. Disks with thick substrates are still less efficient. On the other hand, reels of tape embody the disadvantage that access to a given block of information can be obtained only by moving past the playback head all the unwanted information that intervenes between the starting point and the desired point. Chips and strips offer improved access time because they can be picked out individually from their storage locations. Disk systems, particularly the ones that employ a number of playback heads, offer a further improvement in access time.

New Techniques

The difficulties I have described—mechanical problems presented by the sliding contact between head and tape, failures experienced with air-bearing heads and the limitation of packing density and signal output arising from spacing—have led to the investigation of new methods that do not require a conventional head for recording or playback. The most promising techniques take advantage of the fact that the coercivity of certain magnetic materials can be changed by the heat produced by the absorption of radiant energy such as light. Although such techniques have been under investigation for many years,



Kleinschmidt.

The amazing Kleinschmidt 311™ Teleprinter is part of BOAC's 100 million dollar computer system called Boadicea. This system puts all the BOAC offices in the United States right next door to the home office in London.

For example, if the BOAC office in Miami needs a passenger list for any BOAC flight, Kleinschmidt gets the answer from London in seconds. At 400 words a minute.

The Kleinschmidt 311 Teleprinter gives BOAC the ultimate in speed and reliability. Furthermore, the 311 Teleprinter can accept parallel information on demand at an average speed of 40 characters a second, which eliminates the formatting problems of serial printers.

What Kleinschmidt is doing for BOAC it can do for any business that requires quick reception of the printed word. If you would like to learn more, please write Kleinschmidt Division of SCM Corporation, Deerfield, Ill. 60015. Or call us at 312-945-1000.

SCM®

the arrival of the remarkably intense light source represented by the laser has greatly improved the performance of magneto-optical systems.

In each such system the recording technique involves the coincidence of a beam of light and a magnetic field at the spot where information is to be recorded. Among the materials used as recording mediums are garnets, europium oxide and metal alloys. All are in the form of thin films.

Garnets are ferrimagnetic materials. Their crystals embody two sublattices that are oppositely magnetized. At a temperature termed the compensation temperature the oppositely directed magnetization of the two sublattices is equal, and the coercivity of the garnet is high. If the temperature is raised by a few degrees, however, the equilibrium of the sublattices is disturbed and the coercivity decreases appreciably. If a constant magnetic field weaker than the field that would be required to switch the garnet is applied in coincidence with a light beam of sufficient strength to heat the garnet a few degrees, the area illuminated by the beam will reverse its direction of magnetization. A "bit" of information is thus recorded. If the time needed to produce the rise in temperature is short, the size of the recorded spot will be approximately equal to the diameter of the incident beam.

Similar methods have been used with europium oxide and metal alloy films. With these materials much larger temperature increases are required to reduce the coercivity. For europium oxide a rise in temperature of 100 degrees Celsius is required to reduce the coercivity enough for recording. Moreover, the material must initially be at a cryogenic temperature, that is, not far above absolute zero. The low specific heat of the europium oxide film at cryogenic temperatures makes recording possible with milliwatts of power in the optical beam.

Temperature increases of the same magnitude are required for metal alloy films. With these films, however, the higher specific heat calls for an optical beam with hundreds of milliwatts of power. On the other hand, the process can be carried out at room temperature.

All the readout techniques take advantage of the same physical phenomenon. It is that if plane-polarized light is incident on a magnetic medium, the plane of polarization will be rotated during transmission through the medium or after reflection by the medium. The direction and magnitude of the rotation of the plane of polarization are determined by the magnetic state of the material. The changes in polarization are detected by an analyzing crystal and a photocell, thereby generating the electrical signal for the readout.

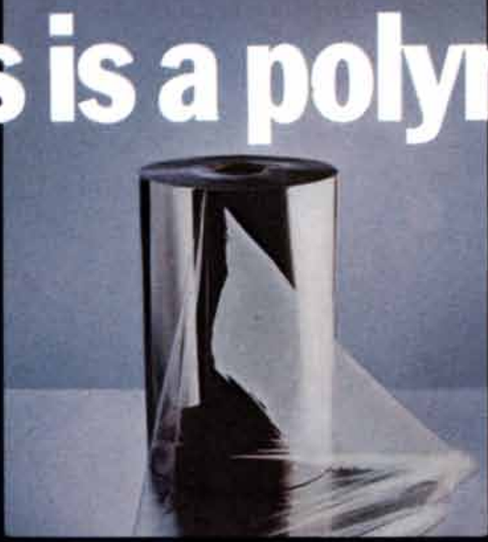
Since ferrimagnetic garnets and europium oxide are transparent to light of an appropriate wavelength, recording systems employing these materials take advantage of the rotation of the plane of polarization during transmission. The amount of rotation is much larger than can be produced by reflection, and usable signal-to-noise ratios are obtained. Unfortunately these materials are difficult to fabricate, and large areas of continuous film have yet to be produced. In addition, the systems must be operated at carefully controlled temperature. The applications that seem most practical are therefore in chip systems, where small areas can be used, rather than in continuous systems such as tape.

Films made of metal alloys are more promising. The readout from them is accomplished by reflection. Although the rotation of the plane of polarization in reflection is small, it has been possible to develop noise-cancellation techniques that achieve acceptable signal-to-noise ratios in the readout of digital information recorded on metal films by heat. In addition, it appears that it may be possible to make metal alloy film with a large surface area, so that the film could be used in tape and disk systems.

Magneto-optical systems offer three major advantages. First, there is no contact between a magnetic head and the recording medium. Wear, frictional drag across the tape and spacing losses are therefore eliminated. Second, the readout is static in that no relative motion of the medium with respect to a playback head is required. The effects of alterations in velocity on the amplitude of the signal are thus eliminated. Third, since signal-to-noise ratio no longer depends on track width, relative head-to-tape speed or spacing, very high packing densities are possible. Indeed, they are limited only by the size of the optical spot that can be produced. Area packing densities of more than 10 million bits per square inch have been achieved.

Notwithstanding the substantial improvement of magneto-optical systems in the past few years, two major problems must be solved before such systems will progress beyond the laboratory stage. Scanning and the deflection of the optical beam are now accomplished mechanically; there is a need for electro-optical deflectors. In addition, it will be necessary to develop methods for accurately positioning the playback beam on the extremely narrow recorded track.

This is a polymer.



It's also a TV film. A computer tape. And a printed circuit.

Linking molecules together into intricate chains, called polymers, helps us develop some very exceptional plastics.

Celanar polyester film, for example. So strong and dependable that it's used in filming TV programs and commercials. Which is a reason you've probably never seen a program interrupted by a film break.

For the special requirements of computer tapes, our polymer technology produced Celanar "Low Profile" film. Three times smoother than other polyester films, it permits packing more data on the tape, with far fewer errors.

Celanar is also used in making printed circuits, because its toughness and great dimensional stability makes them more reliable.

Polymer technology is the science upon which all of Celanese Corporation is based. One in which we have more than tripled our development efforts in the past ten years. And a reason why Celanese plastic films and resins have more to offer you. Celanese Plastics Company is a division of Celanese Corporation.

Celanese®
Celanar®

CELANESE PLASTICS

Portrait of a man headed for a stroke.

The portrait you're looking at is a thermogram, a "heat map" of the patient's skin.

It reveals a blockage in the carotid artery, often signifying an impending stroke. Fortunately, when so forewarned, doctors can prevent such strokes.

Thermography is remote scanning and measurement of infrared radiation which indicates surface temperatures, and their positional translation into black-and-white or color values on film.

It is only recently that the benefits of thermography as a medical diagnostic tool have become apparent. It is totally nondestructive, as opposed to X-ray or operative methods.

And it is of significant, even dramatic, diagnostic value in such areas as trauma, arthritis, and screening for early detection of breast cancer.

Of the growing number of hospitals and diagnostic centers in the United States currently equipped for thermography, virtually all use Polaroid Land film.

In cases where time is critical, the advantage of instant picture development is obvious.

In any case, the Polaroid Land photographic process offers the doctor the opportunity to retake a shot, when necessary, without having to recall the patient and having to re-create a specific temperature environment; and to make diagnostic or prognostic decisions without delay.

Medical authorities are confident that in the near future many more medical centers will be equipped for clinical thermography. Its importance as an early detection device is clear.

And Polaroid Land film is the visual part of the picture.

Polaroid Corporation

Thermogram by Barnes Engineering Company

Polaroid®

Non-Euclidean Geometry before Euclid

Certain works of Aristotle, written 2,000 years before the advent of non-Euclidean geometry, contain references to the possibility of a non-Euclidean approach to the famous problem of parallels

by Imre Tóth

The invention of non-Euclidean geometry is universally recognized as one of the great intellectual achievements of the 19th century. The independent discovery of Carl Friedrich Gauss, Johann Bolyai and Nikolai Ivanovich Lobatchevsky, it exploded the ancient notion that mathematics could be a source of absolute truths by showing that the most venerable mathematical "truths" of all, the geometry of Euclid, could be negated in a consistent deductive system of mathematical theorems.

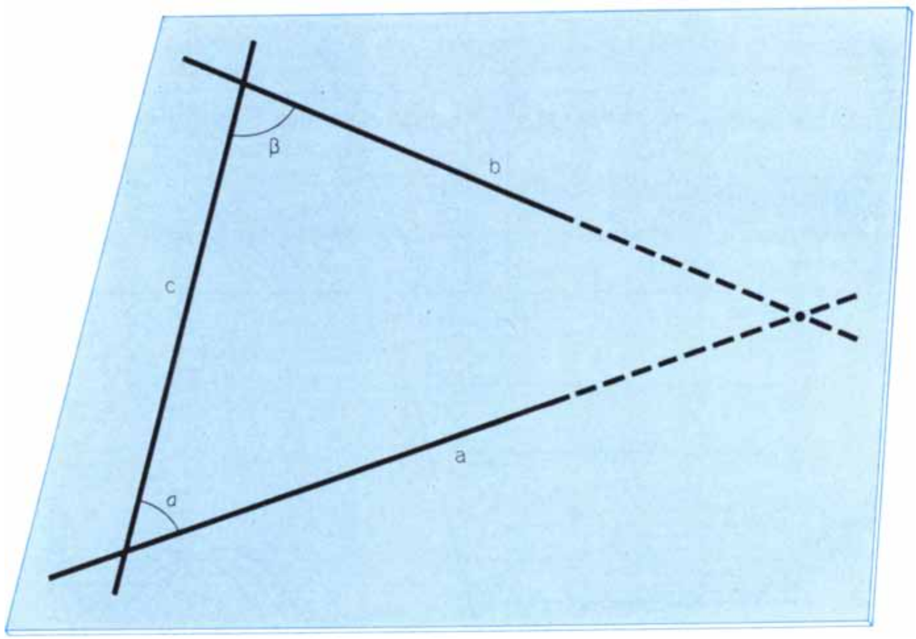
The seed of this revolution was the ancient problem of parallels. The purpose of this article is to describe some research on the problem *before the time of Euclid*—non-Euclidean theorems that hitherto have remained unnoticed in certain pre-Euclidean texts. First, however, it is necessary to describe the problem of parallels as it appeared from Euclid's time until the modern era.

In brief, the problem of parallels is to prove that Euclid's fifth postulate [*see illustration at right*] is in fact not a postulate at all (in the sense of its being an undemonstrable statement) but rather is a theorem that can be proved using only axioms of what Bolyai called "absolute geometry." Bolyai gave this name to those theorems that could be proved without using Postulate 5. A non-Euclidean geometry is obtained if Postulate 5 is denied. Hence absolute geometry consists of those propositions that are true regardless of whether Postulate 5 is assumed to be true or false. In other words, absolute geometry is that which is true in both Euclidean and non-Euclidean geometry. Euclid's first 28 propositions belong to absolute geometry.

As we can now see clearly, it was Euclid's great contribution to recognize that a postulate or an assumption was needed to specify the nature of parallel lines. To many of his followers, however, the presence of this statement in the set of the undemonstrable postulates was a blemish that should be removed from the theory.

If the problem of parallels could be solved, it would be proved that the fifth postulate is a theorem of absolute geometry. It would then follow that absolute geometry includes all of Euclidean geometry. A consistent non-Euclidean geometry would then be impossible. From our superior vantage point, knowing that non-Euclidean geometry *is* possible, it is both amusing and frightening to list the mathematicians of times past who thought they had proved the fifth postulate. The list begins with Claudius Ptolemy (about A.D. 150), author of the definitive Greek work on astronomy. Ptolemy's error was pointed out in the fifth century by Proclus, who gave his own "proof," based on an argument that turned out to be just as circular as Ptolemy's. So it went, through Nasr-ed-din (Persian astronomer and mathematician, 1201–1274) and John Wallis (English mathematician, 1616–1703).

In 1889, well after non-Euclidean geometry had been established, a lost chap-

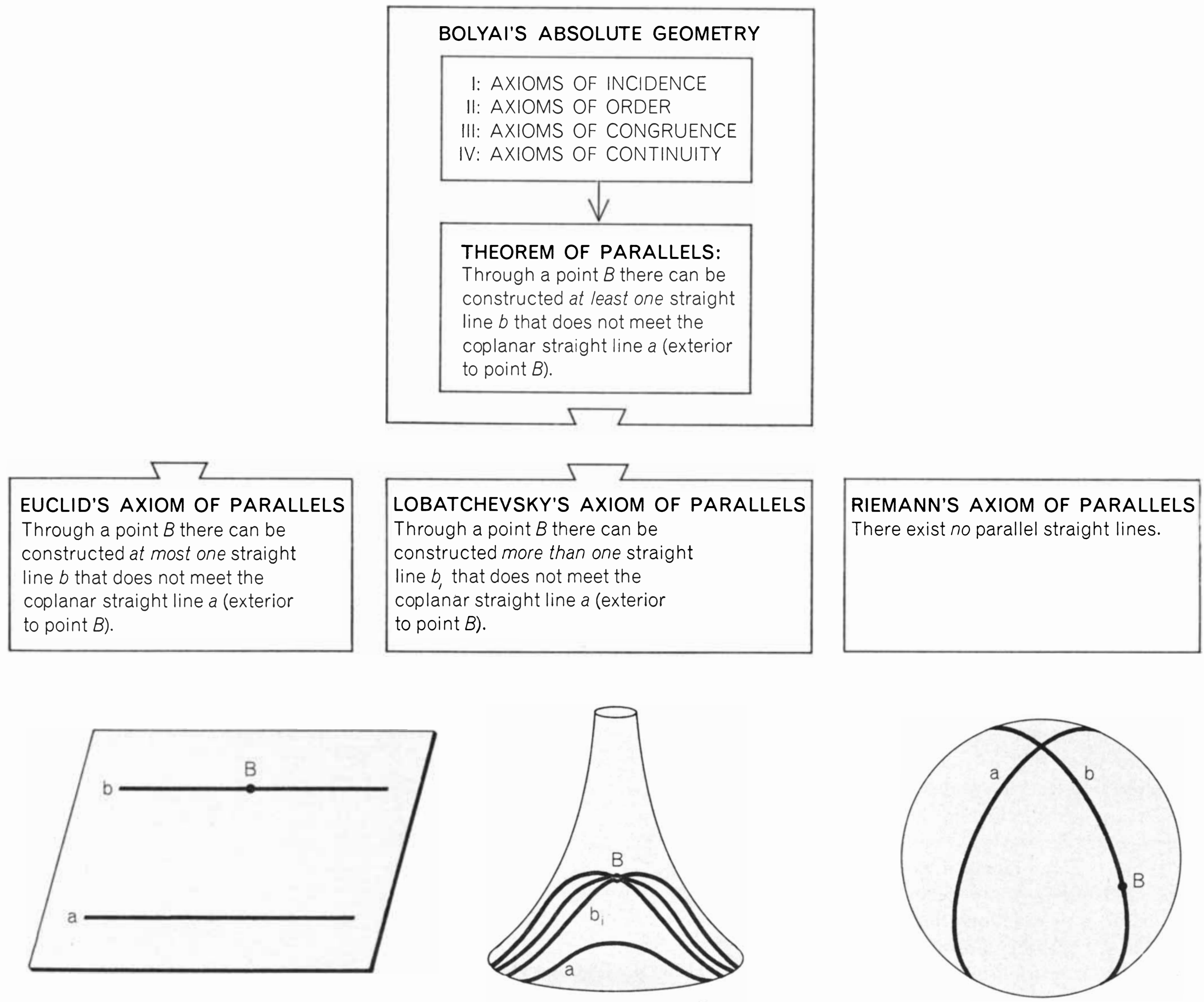


EUCLID'S FIFTH POSTULATE, sometimes called the postulate of parallels, is one of five postulates, or undemonstrable statements, on which Euclid based his system of geometry. The fifth postulate states that if a straight line (c) falling on two coplanar straight lines (a, b) makes the interior angles (α, β) on the same side less than two right angles, then the two straight lines, if produced indefinitely, will meet on that side on which the interior angles are less than two right angles. Mathematicians from Greek antiquity to the 19th century tried in vain to prove that Postulate 5 is in actuality not a postulate at all but rather a theorem that could be proved. The great achievement of the 19th-century mathematicians Carl Friedrich Gauss, Johann Bolyai and Nikolai Ivanovich Lobatchevsky was to recognize that consistent systems of mathematical theorems could be erected on the denial of Postulate 5, thus opening the way for the proliferation of non-Euclidean geometries in the modern era.

ter in the history of the parallel problem was rediscovered. Eugenio Beltrami (Italian mathematician, 1835–1900, and one of the important contributors to non-Euclidean geometry) resurrected the writings of his forgotten compatriot Girolamo Saccheri (1667–1733). Saccheri was a Jesuit professor who, as a logician, became a specialist in *reductio ad absurdum:* the method of proof in which one shows that if the desired conclusion were *not* true, a contradiction (*absurde*) would follow. Saccheri found the fallacies in the attempts of Nasr-ed-din and Wallis to provide a direct proof for the fifth postulate. He then undertook, in a book titled *Euclides ab omni naevo vindicatus* (*Euclid Freed of Every Stain*), to succeed where they had failed. He used the indirect method, and he tried to prove the fifth postulate by showing that its negation led to an absurdity—a contradiction. Of course, as we now know, he could not possibly arrive at a contradiction; what he could and did arrive at (unknowingly) was non-Euclidean geometry.

Saccheri started with the first 28 propositions of Euclid, which, as I have mentioned, do not use Postulate 5. He then considered the following figure [*see illustration on opposite page*]. Sides *AC* and *BD* are parallel, and angles *A* and *B* are right angles. It follows easily, using some of Euclid's first 28 propositions, that angle *C* equals angle *D*. But is it *equal to, greater than* or *smaller than* a right angle?

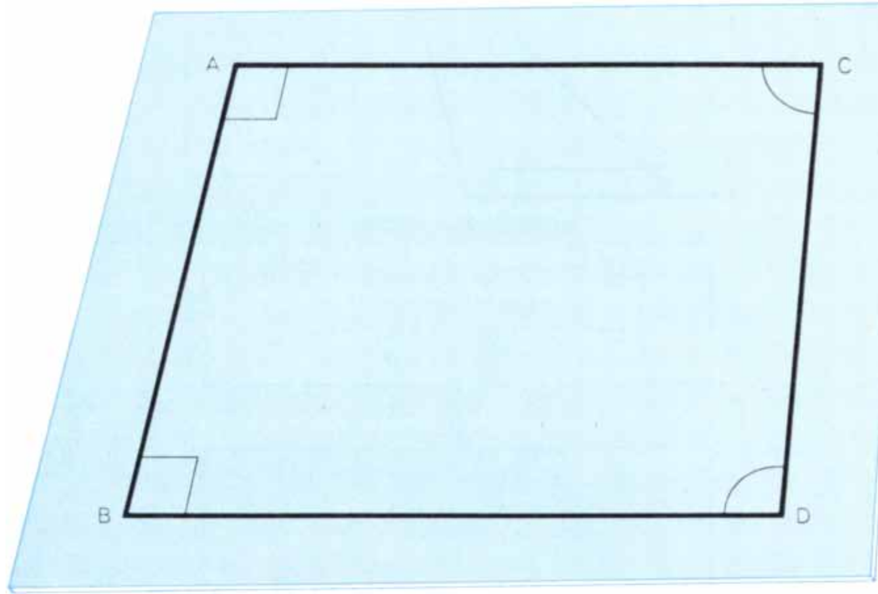
The first case, the "hypothesis of the right angle," or the Euclidean hypothesis, is equivalent to the fifth postulate, and also to the Euclidean proposition that the sum of the angles of every triangle equals two right angles. It is this case that Saccheri wished to establish. To do so he had to reduce to absurdity the opposite case, that the angle is *not* equal to a right angle. Clearly this "counter-Euclidean" hypothesis consists of two cases: the "hypothesis of the acute angle" (less than 90 degrees) and the "hypothesis of the obtuse angle" (more than 90 degrees).



ABSOLUTE GEOMETRY, the name given by Bolyai to those theorems that could be proved without using Postulate 5, is actually an incomplete system; it consists of those propositions that are true in both Euclidean and the "hyperbolic" variants of non-Euclidean geometry. Thus absolute geometry can be combined either with Euclid's axiom of parallels to form Euclidean geometry (*left*) or with Lobatchevsky's axiom of parallels to form hyperbolic non-Euclidean geometry (*center*). In hyperbolic geometry the straight line drawn on a plane can be interpreted on a Euclidean model as a "geodesic" drawn on a surface of constant negative curvature, such as the "pseudosphere." The axiom of parallels put forward in 1854 by Bernhard Riemann, on the other hand, is inconsistent with Bolyai's concept of absolute geometry (*right*). Hence the elaboration of Riemannian, or "elliptic," non-Euclidean geometry, in which straight lines correspond to great circles on the surface of a sphere, required another definition of absolute geometry.

I use the term "counter-Euclidean" for a hypothesis founded on the philosophy that a single geometry alone is possible, either Euclidean geometry or the opposite. The term "non-Euclidean geometry" applies only to a system founded on the view that simultaneously accepts both systems.

Saccheri actually did prove that the hypothesis of the obtuse angle is impossible. To do so he had to use the tacit assumption that a line has infinite length. (In Bernhard Riemann's "elliptic" geometry the obtuse-angle hypothesis prevails, but straight lines have a finite total length.) In actuality what Saccheri demonstrated was that the "hypothesis of the obtuse angle" is inconsistent with the absolute geometry of Bolyai. This theorem of Saccheri's is equivalent to the absolute theorem that the sum of the angles of a triangle cannot be greater than two right angles. In this form it was rediscovered by Adrien Marie Legendre (French mathematician, 1752–1833), and it was long known as Legendre's theorem, since Saccheri's book, printed in Milan in 1733 a few months before its author's death, had been forgotten until it was rediscovered by Beltrami.



INDIRECT PROOF of Euclid's fifth postulate was attempted by the 18th-century Italian geometer Girolamo Saccheri. Given the figure shown here, in which sides *AC* and *BD* are parallel and angles *A* and *B* are right angles, it follows, using some of Euclid's first 28 propositions (which are not based on Postulate 5), that angle *C* equals angle *D*. Saccheri tried to show that a contradiction would follow from the "counter-Euclidean" assumptions that either *C* or *D* is greater than or smaller than a right angle. In the process Saccheri actually proved a number of important theorems in non-Euclidean geometry, but he unwittingly regarded these results as "absurd." References to strikingly similar mathematical investigations have been found recently by the author in the philosophical writings of Aristotle.

Saccheri did not succeed in eliminating the third possibility, the acute-angle hypothesis. His asserted absurdity in this case is an unconvincing argument involving vague concepts about elements at infinity. The acute-angle hypothesis is precisely the hypothesis of Lobatchevsky and Bolyai. Beltrami, the same man who rediscovered Saccheri, showed that a certain surface in Euclidean three-space, the pseudosphere, actually satisfies the non-Euclidean geometry of the acute-angle hypothesis. This showed that if non-Euclidean geometry were self-contradictory, so would be Euclidean geometry. If Saccheri had succeeded in removing the "blemish" from Euclid, it would have meant blowing up all Euclidean geometry at the same time! This news would have puzzled Saccheri rather unpleasantly.

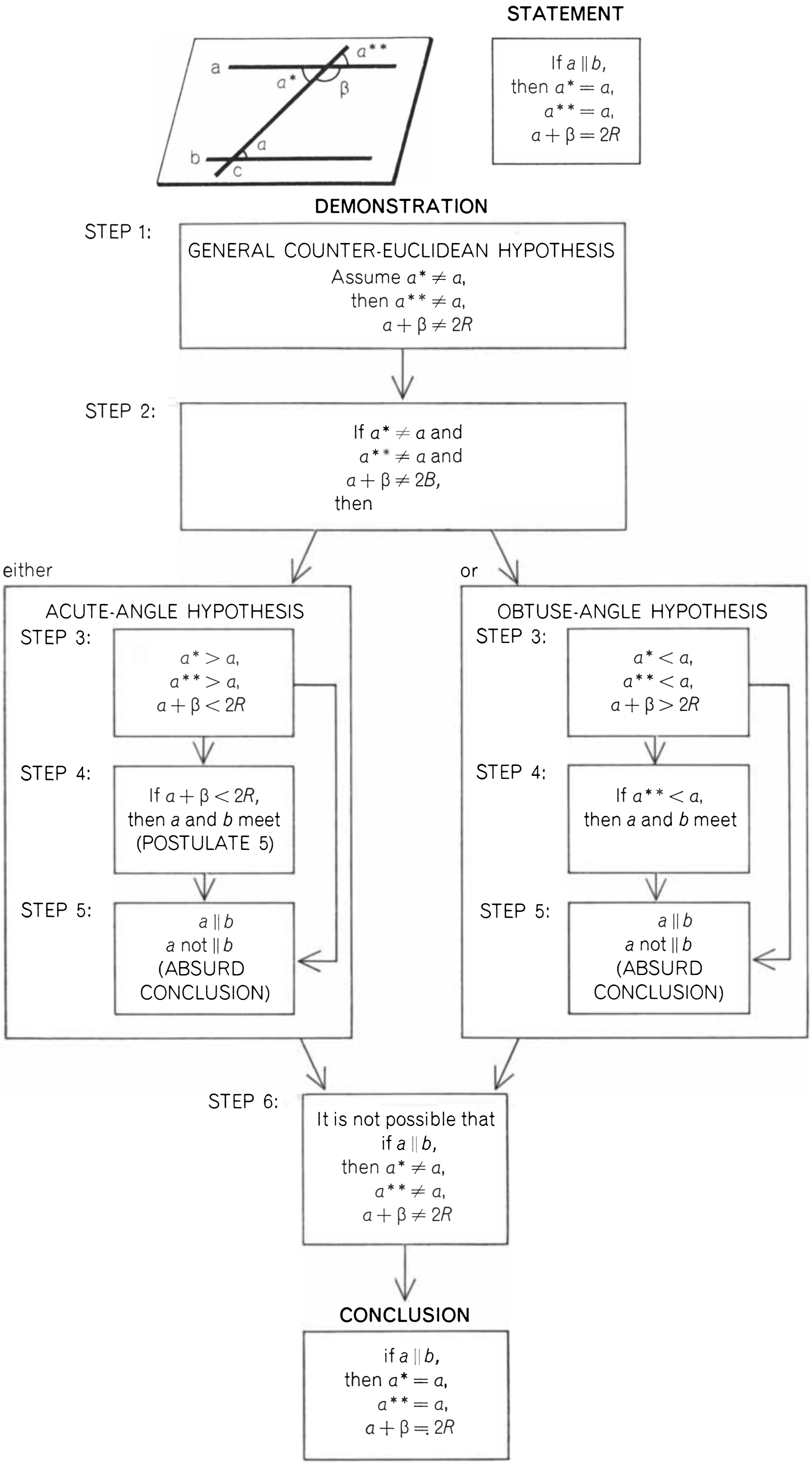
From today's viewpoint Saccheri's real achievement was to prove a number of important theorems in non-Euclidean geometry. His great shortcoming was to insist on regarding them as absurd. In trying to absolve Euclid, however, he became an unwitting forerunner of the non-Euclideans.

It is commonly thought that Saccheri was the first mathematician to achieve this distinction. (The German mathematician Johann Heinrich Lambert later independently repeated and went beyond Saccheri's work.) The amazing fact is that similar investigations were carried out by the Greeks a generation before Euclid! Clear references to these investigations occur in Aristotle's philosophical writings, where they have lain, seemingly unnoticed by mathematicians, for more than 2,000 years.

Tracing back the notion of an indirect approach to the parallel problem, we can find indications of it in Gersonides (14th century), Nasr-ed-din (13th century), Omar Khayyám (11th century), Al Hazan (10th century) and even in Ptolemy (second century), as his work is described in Proclus' *Commentary* of the fifth century (the earliest source concerning the problem). Of course, only Saccheri and Lambert established an entire connected system of counter-Euclidean propositions; the others rapidly eliminated the counter-Euclidean hypothesis by some more or less trivial error. The source of the idea is to be sought much earlier. Indeed, the acute-angle hypothesis was explicitly formulated by Euclid himself, in the demonstration of Proposition 29, Book I, of his *Elements* [*see illustration on next page*].

Euclid starts the demonstration of this theorem using the indirect method. Against the statement of Proposition 29, Book I, he formulates at first the general counter-Euclidean hypothesis: "If *a* and *b* are parallels, then the interior angle α is unequal to the alternate angle α^*, consequently to the exterior angle α^{**} too, and the sum of the two interior angles α and β is also unequal to $2R$" (two right angles). This is obviously an attempt to reduce the hypothesis to absurdity. "But if α is unequal to α^* [continues Euclid], then one of them is greater. Let the alternate angle α^* be the greater and consequently the exterior angle α^{**} too; then $\alpha + \beta < 2R$." This last conclusion can be explicitly formulated as follows: "Two straight lines, *a* and *b*, intersected by a secant *c*, forming the interior angles α and β on the same side of *c*, do not meet; specifically, this nonincidence occurs on the same side of the secant where the sum of the angles, $\alpha + \beta$, is less than $2R$." This, however, is nothing other than the acute-angle hypothesis. And Euclid opposes to it the fifth postulate, which says: "If $\alpha + \beta < 2R$, then the straight lines *a* and *b* meet." But they are also parallel! This is a formal contradiction, and from it Euclid immediately draws this conclusion: "Angle α cannot be therefore unequal to angle α^*; then they are equal, and we have also $\alpha = \alpha^{**}$ and $\alpha + \beta = 2R$; Q.E.D."

Ptolemy rightly criticized the formulation Euclid gave this postulate, observing that the half-plane within which the straight lines *a* and *b* are to meet is being specified, which is unnecessary because it constitutes in itself an absolute theo-



EUCLID'S PROOF of Proposition 29 in the first book of his *Elements*, presented in symbolic form in this illustration, uses an indirect method of demonstration. Euclid begins by formulating the general counter-Euclidean hypothesis: If *a* and *b* are parallels, then the interior angle α is unequal ($\neq$) to the alternate angle α^*, consequently to the exterior angle α^{**} too, and the sum of the two interior angles α and β is also unequal to two right angles ($2R$). Relying on the fifth postulate, Euclid then shows that the acute-angle hypothesis (*middle left*) results in an absurd conclusion (*a* and *b* meet and are also parallel). He immediately concludes that "Angle α cannot be therefore unequal to angle α^*; then they are equal, and we have also $\alpha = \alpha^{**}$ and $\alpha + \beta$ equal to $2R$; Q.E.D." Curiously Euclid made no attempt to eliminate the obtuse-angle hypothesis (*middle right*), which can be done without using Postulate 5. The missing parts of this proof appear, however, both in Aristotle's *Analytics* and—2,000 years later—in Saccheri's *Euclid Freed of Every Stain.*

rem. If the proposition appears as the consequence of a chain of arguments with its source in the acute-angle hypothesis, however, then it *is* necessary to specify that the intersection occurs in the same half-plane within which nonintersection was initially assumed. Otherwise there is no formal contradiction.

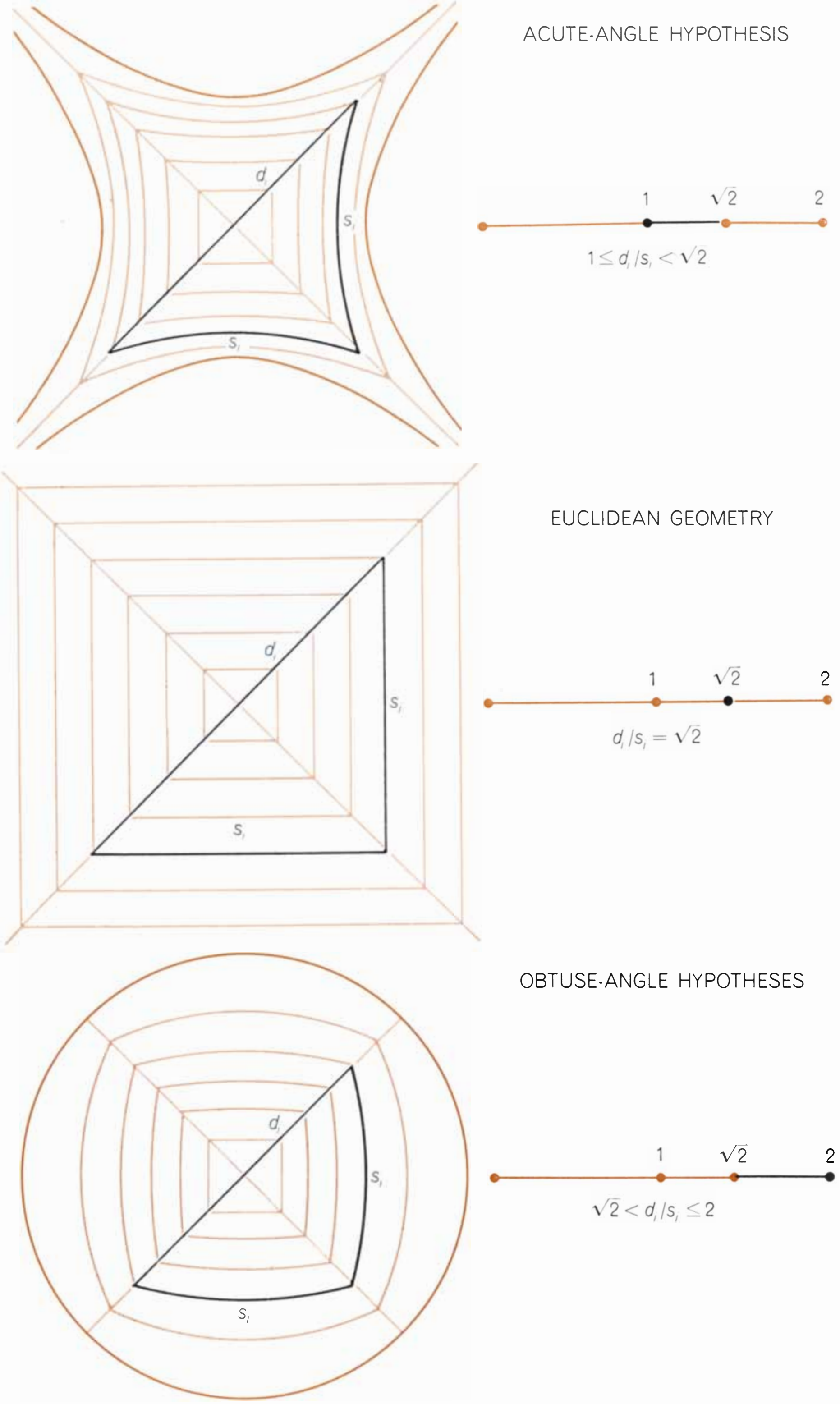
All of this suggests that the postulate of parallels was originally a proposition that forerunners of Euclid attempted to deduce from the acute-angle hypothesis, reducing this hypothesis too to the impossible. The most striking thing in Euclid's argument, however, is the fact that, from a strict formal standpoint, it is entirely wrong. Indeed, Euclid sets forth the general counter-Euclidean hypothesis that one of the angles α and α^* should be greater. He demonstrates that angle α^* (and, of course, α^{**}) cannot be greater than α, but this by no means demonstrates that α^* may not be less than α. In order to attain this end he should have demonstrated that neither could the interior angle α be greater than the exterior angle α^{**}. This hypothesis—obviously the hypothesis of the obtuse angle—is mentioned at the beginning, but no attempt to eliminate it can be found in the *Elements*. This is a stain—and a real one—that soils "the splendid body of the *Elements*" (as Sir Henry Savile put it in 1621). Still more remarkable than this error, it seems to me, is the fact that out of a host of ruthless, hypercritical commentators no one should have observed it, even though criticism has always been focused on Proposition 29, Book I.

The missing part of Euclid turns out to be buried a half-century or so deeper in the geological strata of mind—in Aristotle's *Analytics*. "The same absurd conclusion 'Parallels meet,' " writes Aristotle, "may be obtained either from the premise 'The interior angle is greater than the exterior' [he was undoubtedly referring to the angles α and α^{**}] or from the premise 'The sum of the angles of the triangle exceeds $2R$.' " (Aristotle's manner of expression is allusive and elliptical; I am giving an interpretive transcription in which the gaps are filled. The meaning of the passage is nonetheless unmistakable.) This passage reduces to absurdity the second particular case of the general counter-Euclidean hypothesis, mentioned by Euclid but left unsolved: "Let be $\alpha > \alpha^{**}$." This is nothing other than the obtuse-angle hypothesis, in a formulation perfectly symmetrical with the one given by Euclid to the acute-angle hypothesis, and reduced to impossible by the same absurd conclusion: The parallels *a* and *b* meet. Euclid's text fits Aristotle's just as the jawbone of a fossil hominid might match a cranium found elsewhere.

Here we have arrived at a point in time when human thought, having just begun to walk on two legs instead of four, tries immediately to leap into the interplanetary space of mind. The first part of the fragment from Aristotle is something that is missing from Euclid, but the second part is present in Saccheri as Proposition 14 of his *Euclid Freed of Every Stain*. In other words, Greek geometers before Euclid had already drawn up the obtuse-angle hypothesis and had found the fundamental proposition that eliminated it. This proposition, which is also missing from Euclid, appears nowhere else until Saccheri. The formulation given by Saccheri is identical with the one given by Aristotle in the passage cited above, although it is almost certain that Saccheri did not know of the passage. The pre-Euclidean geometers must have been led to draw up this "counter-Euclidean" hypothesis after their attempt to demonstrate the equivalent of Proposition 29, Book I, in a direct way had failed. Evidently they had fallen into a vicious circle, just as Nasr-ed-din and Wallis were to do centuries later, because Aristotle chose just these attempts "to demonstrate the [problem of] parallels" in order to illustrate what *petitio principii* (begging the question) means.

The obtuse-angle hypothesis appears in Aristotle's works in four passages besides the one I have mentioned. The acute-angle hypothesis appears only once, and then only in passing. All five passages in which the obtuse-angle hypothesis appears quote the general counter-Euclidean hypothesis: "The sum of the angles of the triangle is not equal to $2R$." (I explain this by the fact that the acute-angle hypothesis could not be eliminated, and that accordingly the general counter-Euclidean hypothesis remained valid, absorbing the acute-angle hypothesis.)

How far things went in the pursuit of the consequences of the counter-Euclidean hypothesis is shown by the following statement in Aristotle's *De caelo*: "If it is impossible for the triangle to have the sum of its angles equal to $2R$, then the side of the square is commensurable with the diagonal." Here is a powerful and beautiful theorem that is not to be found in either Saccheri or Lambert, and not in the modern creators of non-Euclidean geometry. From the general counter-Euclidean hypothesis it



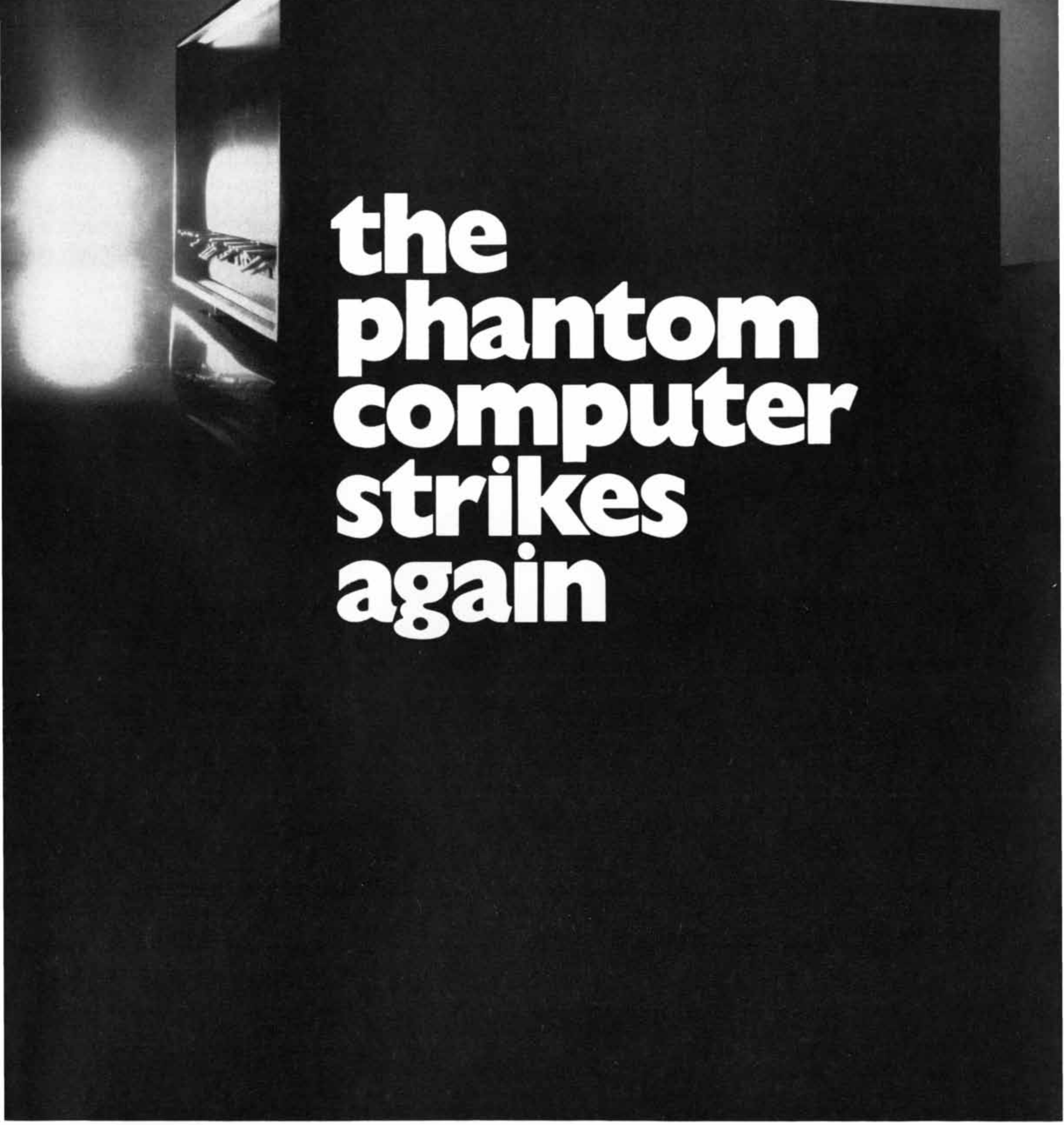
"ABSURD" CONCLUSION must have resulted from the attempts of the pre-Euclidean geometers to demonstrate the equivalent of Euclid's Proposition 29, Book I, in an indirect way by means of the counter-Euclidean acute-angle or obtuse-angle hypothesis. Evidence of their approach is contained in the following passage from Aristotle's *De caelo:* "If it is impossible for the triangle to have the sum of its angles equal to 2*R*, then the side of the square is commensurable with the diagonal." (That is, the ratio of the diagonal to the side can be a rational number, instead of $\sqrt{2}$, as in Euclidean geometry.) In this representation of the problem, proposed by the 14th-century mathematician Levi ben Gerson (Gersonides), all the sides and all the angles of each "Gersonides quadrangle" (*center*) are equal. Assuming the acute-angle hypothesis (*top*), the ratio of the diagonal to the side (d_i/s_i) is equal to or greater than 1 and less than $\sqrt{2}$. In this case the limit of the quadrangle, at which d_i/s_i equals 1, consists of two pairs of parallel straight lines (similar to a pair of hyperbolas). Assuming the obtuse-angle hypothesis (*bottom*), d_i/s_i is greater than $\sqrt{2}$ and less than or equal to 2. In this case the limit of the quadrangle, at which d_i/s_i equals 2, is a paradoxical "straight line closed upon itself" (like a circle). In Euclidean geometry either assumption can be shown to result in the contradiction that the same number is both odd and even.

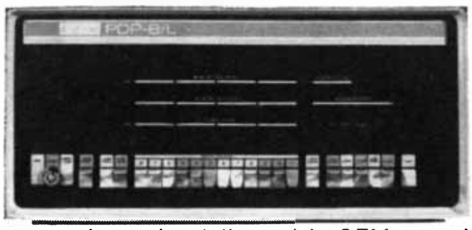
follows that the diagonal of the square could be commensurable to its side [*see illustration at left*]. Failing to find among the consequences of the general counter-Euclidean hypothesis the absurd conclusion "Parallels meet," the Greek geometers probably attempted to obtain from the same hypothesis another absurd conclusion: "The diagonal of the square is commensurable with its side." (That is, the ratio of the diagonal to the side is a rational number.) They knew that from this would follow the contradiction that the same number is both odd and even. Unfortunately, in order to deduce the impossible conclusion "Odd is even," it is necessary to have recourse to the Euclidean proposition "The sum of the angles of the triangle is equal to 2*R*." Thus the absurd conclusion "Odd is even" cannot be reached from the counter-Euclidean hypothesis alone.

What impression was produced by this series of counterpropositions that could not be demolished because no internal contradiction could be found in spite of their obviously pathological character? Some passages in Aristotle's works *Magna moralia* and *Ethica ad Eudemum* throw an unexpected light on the situation.

In these passages Aristotle is putting forward the following conception: Man alone of all beings is free. He is free to choose between good and evil. If the ethical value of the actions changes, this is nothing but the consequence of a change of the principles; if the principle of the actions remains unchanged, the generated actions cannot have opposite ethical values. In the field of geometry we encounter the same situation, namely the possibility of choice between two opposite principles. Each of the systems generated from these principles must be self-consistent; in neither of them can two contradictory propositions coexist, because they would reciprocally destroy themselves.

From these passages it is clear that to Aristotle the *essence* of geometrical propositions lies in their being either Euclidean or non-Euclidean. The Euclidicity and non-Euclidicity of the principles are invariants, and they are transmitted as such to their descendants. As Aristotle writes in *Ethica ad Eudemum:* "Thus if the sum of the angles of the triangle is equal to 2*R*, then the sum of the angles of the quadrangle is necessarily equal to 4*R;* but if the triangle changes [its geometrical essence], then the quadrangle has to change [its essence] too. For example, if the sum of the angles of the triangle is equal to 3*R* or 4*R*, the sum of the angles in the quad-





PDP-8/L is often faceless, nameless, unknown. More people buy PDP-8/L without knowing it than you can imagine. □ Because PDP-8/L comes disguised as part of a milling machine, part of a transistor checkout system, part of a computerized measuring system, part of somebody else's instrumentation package, or numerical control system, or just about part of other people's anything. □ It is the perfect OEM computer. Built to be buried. □ Because it is a full, true, general purpose computer. Because it is fast. Because it is rugged. Because it is an on-line, real-time computer that understands other instruments and machinery and works well with them. □ And because it can be delivered in OEM quantities. Because it is priced for OEM applications. And because it is as reliable as the instruments and machinery that are built around it. □ Write for a brochure on the phantom computer. Before it strikes again.

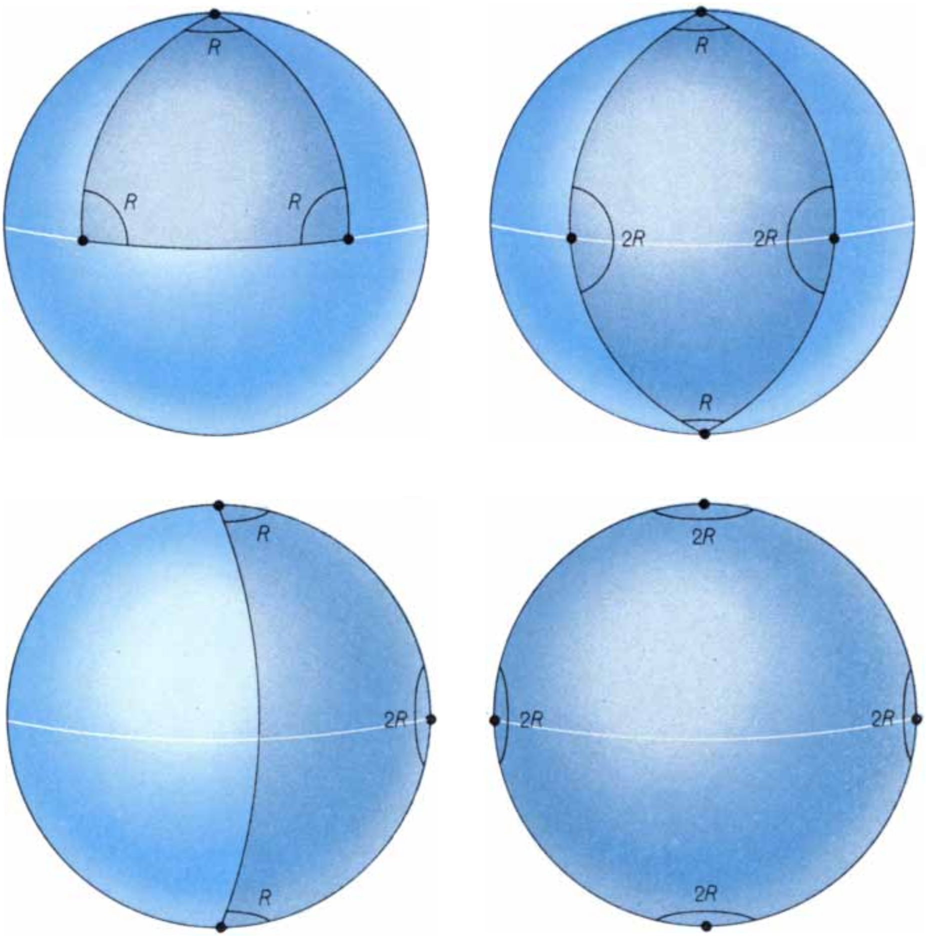
DIGITAL EQUIPMENT CORPORATION, Maynard, Massachusetts, Telephone: (617) 897-5111 / ALABAMA, Huntsville / CALIFORNIA, Anaheim, Los Angeles, and Palo Alto / COLORADO, Denver / CONNECTICUT, New Haven / DISTRICT of COLUMBIA, Washington (College Park, Md.) / FLORIDA, Cocoa / GEORGIA, Atlanta / ILLINOIS, Chicago / MASSACHUSETTS, Cambridge and Maynard / MICHIGAN, Ann Arbor / MINNESOTA, Minneapolis / MISSOURI, St. Louis / NEW JERSEY, Parsippany and Princeton / NEW MEXICO, Albuquerque / NEW YORK, Centereach (L.I.), New York City, and Rochester / NORTH CAROLINA, Chapel Hill / OHIO, Cleveland and Dayton / PENNSYLVANIA, Philadelphia and Pittsburgh / TENNESSEE, Knoxville / TEXAS, Dallas and Houston / UTAH, Salt Lake City / WASHINGTON, Seattle / AUSTRALIA, Brisbane, Melbourne, Perth, and Sydney / CANADA, Edmonton, Alberta; Carleton Place, Ottawa, and Toronto, Ontario; and Montreal, Quebec / ENGLAND, London, Manchester, and Reading / FRANCE, Paris / GERMANY, Cologne and Munich / HOLLAND, The Hague / ITALY, Milan / JAPAN, Tokyo / SWEDEN, Stockholm / SWITZERLAND, Geneva

digital
COMPUTERS · MODULES

rangle will be $6R$ or $8R$." Conversely, as Aristotle writes in *Magna moralia*, "if the sum of the angles in the quadrangle is not equal to $4R$, then neither will the sum of angles of the triangle be equal to $2R$."

In the passage just quoted from the *Ethica ad Eudemum* appears the singular image of a quadrangle the sum of whose angles attains the maximum possible value of eight right angles. The existence of such a figure follows from the hypothesis of the obtuse angle. Today we know that the ordinary sphere provides a model for this geometry [*see illustration below*]. There is no evidence, however, that Greek geometers reached these remarkable consequences of the obtuse-angle hypothesis by studying it on the spherical model. The geometry of the sphere developed much later, and the idea of considering a great circle as a straight line was alien to the spirit of Hellenic geometry. As a matter of fact, the extremely high level of the technique of purely synthetic demonstrations enables us to explain these results without resorting to a spherical model.

In any case, the Greek geometers went far enough in pursuing the consequences of the obtuse-angle hypothesis. Indeed, in considering the proposition on the sum of the angles of the triangle to be a principle—the very essence of the triangle—Aristotle admits the possibility of two opposite principles. In yet another passage he says that the "essence of the triangle is the sum of its angles, which might be equal to, greater than or less than $2R$." This is the only passage where the acute-angle hypothesis is explicitly mentioned, and I should like to point out the remarkable impartiality it shows toward the three Saccherian hypotheses.



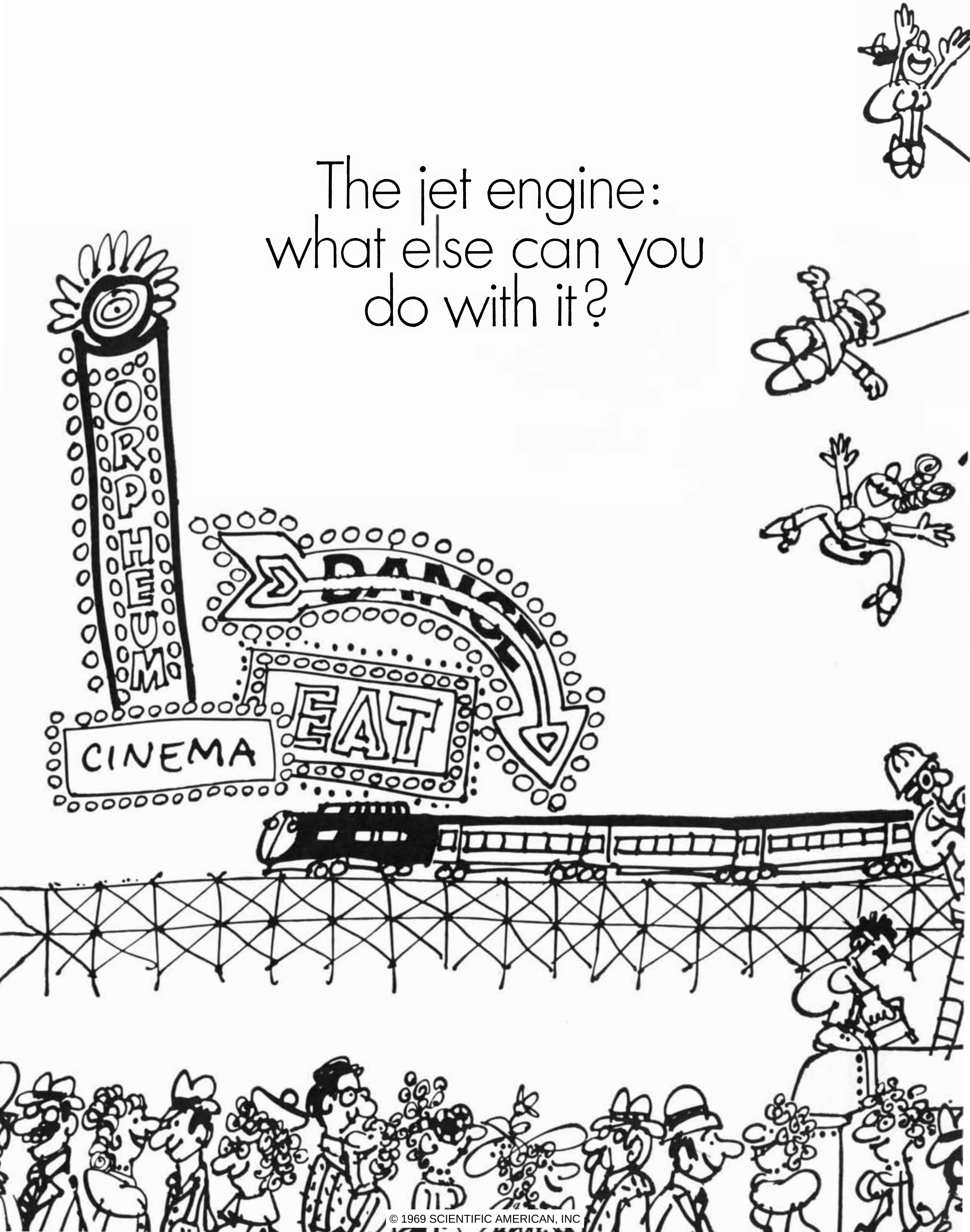
IN RIEMANNIAN GEOMETRY a triangle with the sum of its angles equal to $3R$ can easily be represented as the face of a regular spherical octahedron (*top left*). The sum of the angles of the quadrangle constructed from two such triangles (*top right*) is $6R$. On the same spherical model a triangle the sum of whose angles is $4R$ (*bottom left*) can be doubled to obtain a quadrangle the sum of whose angles is $8R$ (*bottom right*). The latter quadrangle can be regarded as a square degenerated into a single straight line closed on itself, a great circle of the sphere on which the four vertices are each equal to $2R$. All of this is remarkably prefigured in the following passage from Aristotle's *Ethica ad Eudemum*: "Thus if the sum of the angles of the triangle is equal to $2R$, then the sum of the angles of the quadrangle is necessarily equal to $4R$; but if the triangle changes [its geometrical essence], then the quadrangle has to change [its essence] too. For example, if the sum of the angles of the triangle is equal to $3R$ or $4R$, the sum of the angles of the quadrangle will be $6R$ or $8R$."

The Euclidean proposition is considered as being as hypothetical as the other two; all three are presented as being equally possible.

But is free choice possible in mathematics as it is in ethics? In his minutely elaborated parallel between ethics and geometry Aristotle explicitly emphasizes the existence of free choice only in ethics. In geometry he is merely suggesting free choice in the way he takes up the question. It is nonetheless remarkable how calmly he considers the possibility of a metamorphosis of the triangle, the Euclidean theorem of the sum of the angles being so frequently cited as an example of general, immovable and eternal truth. This confused situation is reflected in a remark accompanying Aristotle's parallel between ethics and geometry. "About these things," he writes in *Ethica ad Eudemum*, "one cannot speak at the present moment more accurately, but neither can they be passed over in silence." This is a confession of a kind one seldom encounters in Aristotle. Perhaps his interest in the mathematical achievements of his time was not so casual as is generally believed.

But how could one express such free choice if it actually existed in geometry? Is it possible to accept that the unspecified, abstract concept of "triangle" simultaneously allows both of the predicates "The sum of the angles is equal to $2R$" and "The sum of the angles is not equal to $2R$"? This idea is to some extent discarded in Aristotle's *Metaphysica* when he says (cautiously): "If we were to admit the opinion that the triangle is of immovable essence, we could not admit one time that the sum of the angles is equal to $2R$ and another time that it is unequal to $2R$."

Saccheri speaks of the "inimical acute-angle hypothesis" as if it were a personal enemy; he looks forward to the victory but is aware that he cannot expect it before "a long and desperate fight." Aristotle seems to be attending a contest between two rival hypotheses; he shows the absolute impartiality of an umpire who knows that both competitors have the same right to win. We read in his *Ethica Nicomachea* that this speculative trial—whether the sum of the angles of the triangle is equal or unequal to $2R$—"is not impaired or annihilated by emotional causes, by likes and dislikes, as often happens in ethical trials." Even if both possibilities have equal rights, the decision does not depend on us; this contest is totally different from other competitions and battles. As Aristotle writes in his *Problemata:* "If recollecting, for


The jet engine:
what else can you
do with it?
ORPHEUM
DANCE
EAT
CINEMA

A jet engine does lots more than fly you to far away places. Look around: Our jet-powered generators help provide electrical power for cities. Sleek new TurboTrains use our jet engines. Big pipelines use jet power to pump gas long distances quickly and economically. And United Aircraft jets have gone to sea; giant freighters powered by jets are setting speed records. So, while working towards an ideal transportation system—coordinated, efficient—we're also working on new ways for our jet engines to serve you. Here at United Aircraft. Where new ideas take off.

United Aircraft®

PRATT & WHITNEY AIRCRAFT · HAMILTON STANDARD · SIKORSKY AIRCRAFT · NORDEN · UNITED TECHNOLOGY CENTER · ELECTRONIC COMPONENTS · UNITED AIRCRAFT RESEARCH LABORATORIES · UNITED AIRCRAFT OF CANADA LIMITED ·

instance, the sea battle of Salamis, we feel happy because we have vanquished. We rejoice too if recollecting or expecting to be winners in the Olympic games. But we do not rejoice at all in contemplating the fact of the triangle's having the sum of the angles equal to $2R$ and neither expecting it to be so. Because if we really enjoy speculation, we may be equally pleased whether the sum of the angles is not equal to $2R$ but equal, for instance, to $3R$ or even more than it." We follow a battle or a wrestling match with avowed passion and conscious partiality, but in the contest of the opposed geometrical sentences (even principles) it is the human mind that plays the game in its double aspect of actor and spectator. Here, in the field of thinking, Aristotle continues, "things happen according to their own nature; in this field only the contemplation of the true state of things is pleasurable."

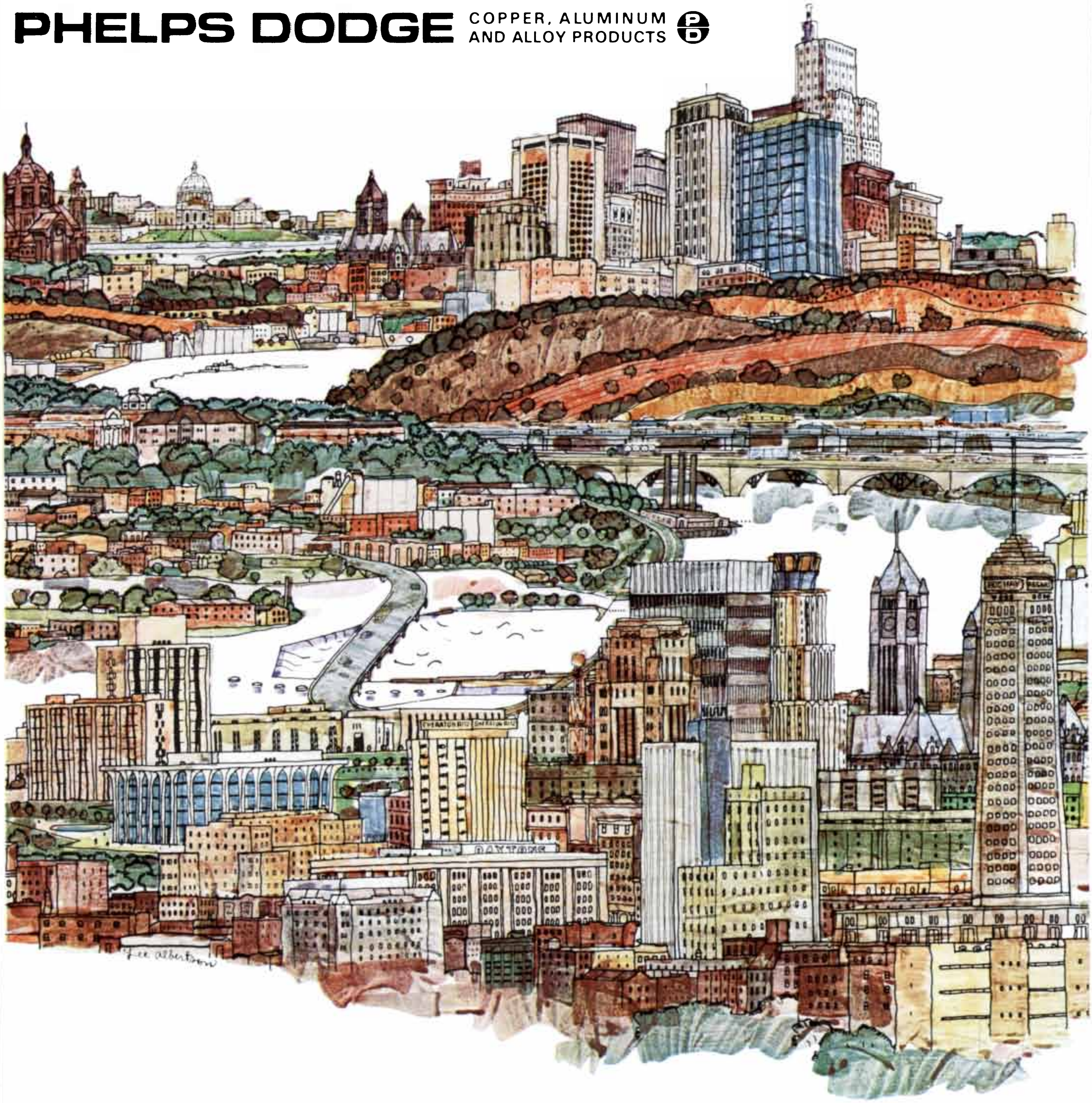
How could the argument be decided? Not by ordinary demonstration, thought Aristotle. The Euclidean proposition "The sum of the angles in the triangle is equal to $2R$" expresses the very essence of the triangle, and it is undemonstrable. The common argument that is regarded as a demonstration is nothing but a quasi-demonstration. The sentence is like a principle, and the truth of such a proposition cannot be decided by a demonstration. As in ethics it is the ethical sense that decides what is right and what is wrong, so here in geometry the decision between principles is made by the *nous,* the intellectual intuition, alone. It is entirely likely that this conception was a theoretical justification for the deadlock in which ended the attempts to demonstrate this fundamental proposition of Euclidean geometry as a theorem of absolute geometry. In any event, the idea that geometry "seeks a principle" (a new postulate) is already pervasive in Aristotle's writings. A generation later saturation is achieved, and the idea is precipitated in Euclid as the postulate of parallels.

The appearance of Euclidean geometry proper was preceded by a stage similar to the one that preceded non-Euclidean geometry at the beginning of the 19th century. The final choice of the Euclidean variant was undoubtedly determined by the fact that the counter-Euclidean hypothesis cannot match geometric figures as they appear in correct drawings. "If this is the straight line," said Aristotle in his *Physica* (probably showing his audience a straight line drawn by a ruler), "then it necessarily ensues that the sum of the angles of the triangle is equal to $2R$; and conversely, if the sum is not equal to $2R$, then neither is the triangle rectilinear." This is a unique passage whose heterodox character struck Sir Thomas Heath. In his *Mathematics in Aristotle* (1949) Heath concluded, however, that the passage appeared as an *ad hoc* example, probably resulting from a pure dialectical play. He could not bring himself to believe that Aristotle could have had such a non-Euclidean idea.

This empiricist philosophy nonetheless fits in well with Aristotle's ethical conception, according to which all that is in accordance with nature is good and all that is opposed to it is evil. A proposition may be called "ungeometrical," Aristotle asserts, "because it does not contain anything geometrical, e.g., a proposition on a musical item [a trivial case], or because it consists in a common fallacy [an irrelevant case, because it is very rare in mathematics] or [the only interesting case] because it contains geometricity in a faulty, distorted way, as, for instance, the proposition 'Parallels meet.' Such a geometry is 'ungeometrical' because it is a wrong and degenerated one." It is remarkable that Aristotle never says that counter-Euclidean propositions are false. The "Parallels meet" proposition, for example, is absurd, but it is nonetheless correct because it does not derive from a logical error. (Indeed, it is a theorem of Riemannian geometry: "Parallels do not exist; all straight lines in the same plane meet.") The free choice between these two opposite geometries appears to Aristotle as an ethical dilemma: one has to choose either the good geometry (in accordance with nature) or the wrong one (against nature).

Digging in these layers of the history of ideas, we have come on a large enough number of fragments to suppose that they are the remains of a unique Athenian vase. I am inclined to believe it is the joint work of an extraordinary team of geometers working in and around the academy of Plato. These men were contemporaries of Aristotle's, and they were particularly under the influence of Eudoxus, whom Aristotle described (in *Ethica Nicomachea*) as being "especially respected and celebrated for his wisdom and fairness." On this antique vase we distinguish the outline of a strange masterpiece: figures having to do with a geometry opposed to the one that was to be constructed by Euclid later on. One is reminded of Picasso's haunting aphorism: "The *against* comes before the *for*."

PHELPS DODGE COPPER, ALUMINUM AND ALLOY PRODUCTS



Minneapolis and St. Paul... a Phelps Dodge double feature.

The Twin Cities . . . a symphony of lakes, river bluffs, park-like suburbs and glittering downtowns . . . the commercial and cultural hub of the upper Midwest. Both are Phelps Dodge cities. You'll find our handiwork everywhere.

Ride on I-94 and I-494, the expressways circling the cities, and Phelps Dodge combined duct lighting cable follows alongside. You'll find our building wire, copper plumbing tubing, aluminum conduits, buried power and telephone cables used in the surrounding suburbs. Manufacturers of electrical motors, generators, relays, radio coils, and other electrical and electronic components are big users of PD magnet wire. Electrical power, the muscles of modern man, travels everywhere on Phelps Dodge high and low voltage power cables . . . on our 15-kv express feeders and 600-volt secondary distribution lines. Our copper, aluminum and alloy rods, bars and shapes are favorites of Twin Cities equipment manufacturers.

We specialize in copper, aluminum and alloy products . . . especially those conducting electricity, liquids, gases, and heat . . . the essential conductors of modern life.

Getting new materials to the assembly line.

Large section boron wing flap skins now possible from "broadgoods" sheets • Spark sintering produces beryllium and titanium parts rapidly and to controlled density • High reliability interconnect matrix modules developed for Polaris and Poseidon programs • Laser alignment system improves accuracy in massive tooling operations.

Manufacturing research, one of the least publicized areas of activity in the aerospace industry, is also one of the most important. New materials and concepts coming out of current R&D efforts require new techniques to turn them to practical use. Manufacturing research provides these needed techniques and processes. The following is a report on some new methods developed by Lockheed for the handling, forming, cutting and processing of new materials and tools.

Large structure boron composites. Boron, combined with epoxy, one of the most promising of the new composite materials, would seem to be an ideal material for such structures as wing slats. But the properties of boron make it difficult to fabricate into large structures. It has a modulus of elasticity of 60 million psi. It is twice as stiff as steel, at less than half the weight. Under an Air Force contract for the design, development, testing and fabrication of boron composite wing slats for ten C-5 airplanes, Lockheed has devised a technique for producing 10x22 foot "broadgoods" sheets of boron composites. A special machine, converted from a conventional lathe, applies 3-inch wide boron filament composite tape continuously on Mylar film to make these sheets. Several pieces of these sheets with specified filament orientations are placed in contoured fixtures for autoclave curing, thus producing the finished wing slat skins.

As a result of this fabrication development, boron composite wing slats are now being made for performance testing and evaluation. Manufacturing research has given aircraft designers freedom to think in terms of large section boron composites for superior stiff-

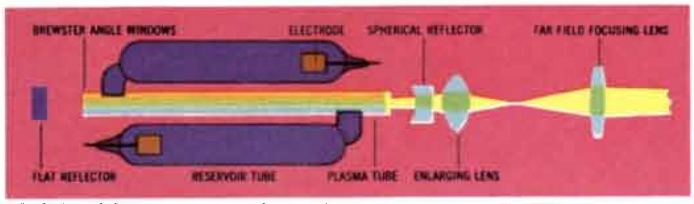
LOCKHEED
LOCKHEED AIRCRAFT CORPORATION

ness and strength with no accompanying weight penalty.

Spark-sintered beryllium. Lockheed is utilizing a breakthrough in the powder metallurgy processing of exotic materials to achieve significant benefits. Simple-geometry parts and specimens are now being produced for a number of customers and require only minor machining to provide finished dimensions. In a new facility, components of beryllium, tungsten-carbide, titanium and other exotic material powders are displaying improved mechanical properties over those found in conventional hot-pressed blocks of the material. The one-shot process has also lowered the fabrication costs of various complex parts, providing controlled and uniform densities with specific strength, thermal, electrical, nuclear and other properties. This is a good example of how manufacturing research turns a new process to practical use, with lowered costs and eventual broadening of applications.

Interconnect matrix modules. New laminated assembly techniques have produced significant advances in the reliability of circuitry for Polaris and Poseidon. These high reliability interconnect matrix modules have virtually no possibility of contact resistance or breakdown and are far superior to conventional printed circuits. In this new process, conductor pattern layers are photo-engraved onto .020 thick sheets of phosphor bronze and then chemically milled. The milled patterns are gold-plated and positioned into a machine-recessed epoxy glass substrate. They are temporarily bonded pending final lamination. Connector tabs placed outside the circuit areas from the layers can be positioned in the same plane as the conductors or diverted in the opposite plane through holes punched in the epoxy substrate. Problems associated with the close proximity of traces at the junction crossovers were eliminated by major redistribution of plane configuration and the addition and deletion of circuitry to avoid complex overlapping. Modules are now being manufactured with 20 to over 1000 interconnections within 42 laminations. Manufacturing research combined a variety of techniques to arrive at this new process, then developed new test equipment to check out these modules, insuring that they would meet stringent requirements for continuity, current carrying capacity, insulation resistance, and voltage breakdown with mechanical strength.

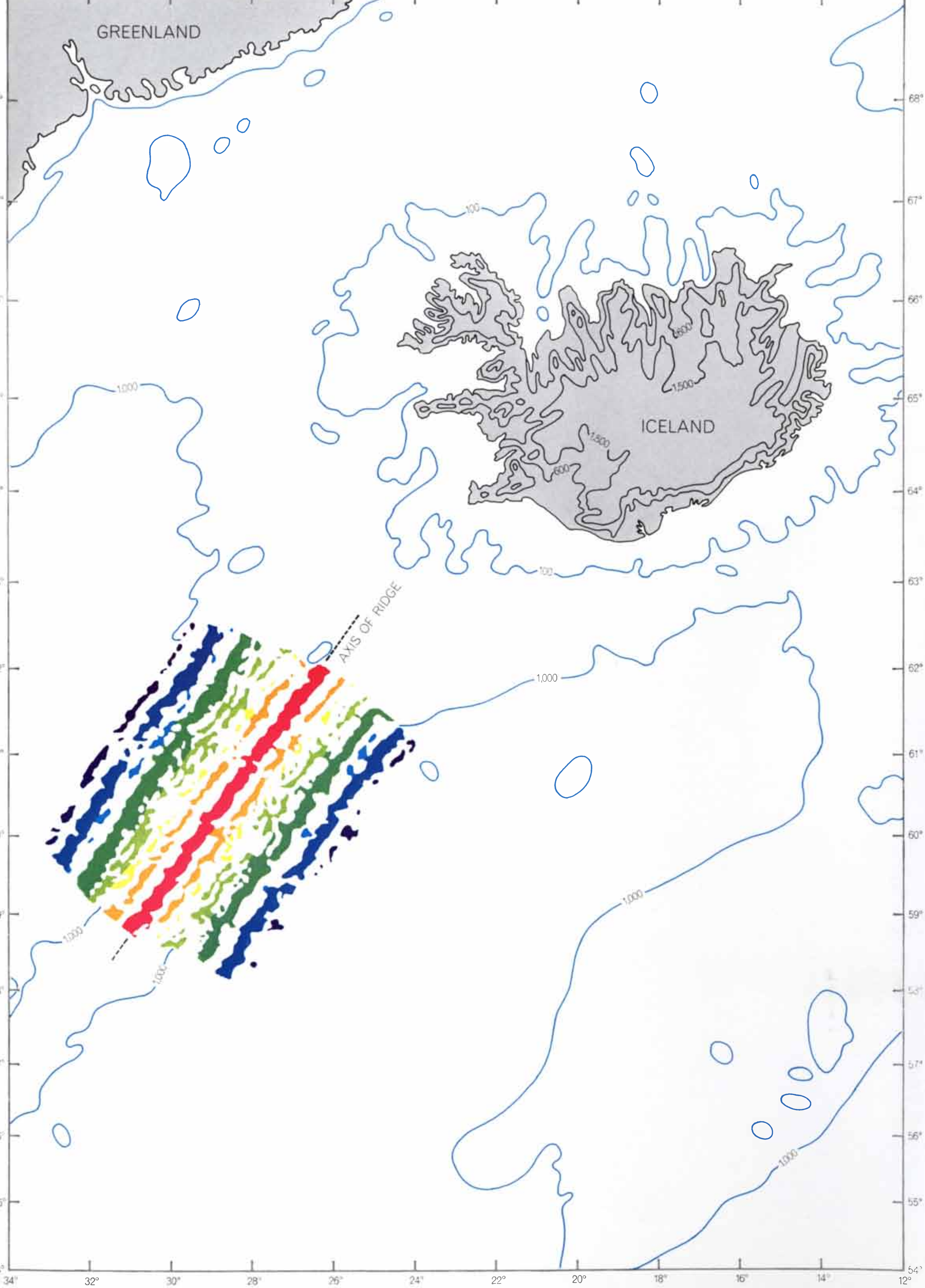
Laser alignment. As larger and larger aircraft are built, tooling jigs become more massive and present difficult alignment and handling problems. A new laser alignment



Modulated laser beam used for alignment.

system at Lockheed has demonstrated exceptional gains in both accuracy and cost savings over conventional optical tooling methods. First tests indicate that the laser system is approximately five times as accurate as the optical system on tools of 60 feet and longer. It is accurate to five-thousandths of an inch at 200 feet, whereas optical systems are accurate to that degree only to 70 feet because of the limitations of the human eye. In comparative tests for cost factors, six major jig test fittings were precisely positioned to the jig test frame. 163 man hours were required to set and inspect all six fittings by conventional methods. In the same test, but using laser alignment and hydraulic positioning controlled by the laser, only 37.3 man hours were required to set and inspect. In subsequent tests, this time dropped to 28 man hours. Here again, manufacturing research was involved in a chain of development. New materials make possible larger aircraft, which require huge tools to build, which require, in turn, new techniques of positioning and handling. This sequence of opportunity and demand led to the laser alignment system.

The activities described here are only a few of the current Lockheed programs in manufacturing research. If you are an engineer or scientist interested in this field of work, Lockheed invites your inquiry. Write: K. R. Kiddoo, Lockheed Aircraft Corporation. Burbank, California 91503. An equal opportunity employer.



GREENLAND
ICELAND
AXIS OF RIDGE
100
1,000
600
1,500
68°
67°
66°
65°
64°
63°
62°
61°
60°
59°
58°
57°
56°
55°
54°
34°
32°
30°
28°
26°
24°
22°
20°
18°
16°
14°
12°

The Origin of the Oceanic Ridges

Current models of sea-floor spreading and continental drift imply convection in the earth's mantle as the driving mechanism. Critical analysis of this assumption leads to an alternative explanation

by Egon Orowan

To a casual observer the surface of the earth is a meaningless jumble of hills, plains and bodies of water. How completely it is dominated, however, by regularities arising from very special circumstances can be seen when it is compared with the surface of the moon. Apart from rather trivial phenomena of soil mechanics, the moon does not seem to show any trace of geology in the literal sense of the word: no mountain belts, no oceans, no oceanic ridges and apparently no volcanoes of the terrestrial kind. Instead it shows meteoric impacts so powerful that they could not have been fended off by the atmosphere of the earth. The fact that only a few small well-preserved impact craters are found on the earth can hardly be the consequence of particular luck in escaping large impacts; it must be due to the ability of the earth's crust to heal wounds quickly, an ability the moon obviously lacks.

The large displacements in the earth's crust revealed by faults, shallow earthquakes, mountain belts, oceanic ridges and the remarkable capacity for obliterating wounds inflicted by large meteoritic impacts indicate that the crucial factor ultimately responsible for geology as we know it is the ability of the crust to slide easily on the mantle, as the skin of the cat can slide on its body. Between the hard "mechanical" crust—the lithosphere—and the hard mantle there must be a relatively soft layer, sometimes referred to by geologists as the asthenosphere. To the physicist this word is misleading, since "astheno-" means weak, not soft. Window glass is weak because it breaks easily, but it is not soft; a nylon filament is strong but not hard. The confusion between fracture and flow has been particularly harmful to geology; it is appropriate, therefore, to replace the time-honored solecism "asthenosphere" by the proper word, "rheosphere," which means an easily flowing layer.

It is usual to assume that the rheosphere is soft because it is hot and because the pressure is not high enough to make it hard. More probably, however, the presence of water, a powerful plasticizer of silica and silicates, plays a decisive role. The preservation of traces of ancient impacts on the moon, therefore, need not mean that the moon has not been hot or that it cooled long ago; it may mean that it has never contained significant amounts of water, or that it lost the water in its infancy.

The thickness of the lithosphere is usually assumed to be between 20 and 50 kilometers and the thickness of the rheosphere between 100 and 400 kilometers. How does the interaction of these two layers with each other and with the mantle give rise to the phenomena of geology? What kinds of forces are acting on them? These are the fundamental questions of geology. The simplest and perhaps most effective clues should be provided by the geography and geology of the mountain belts, continents and oceans, including above all the oceanic ridges.

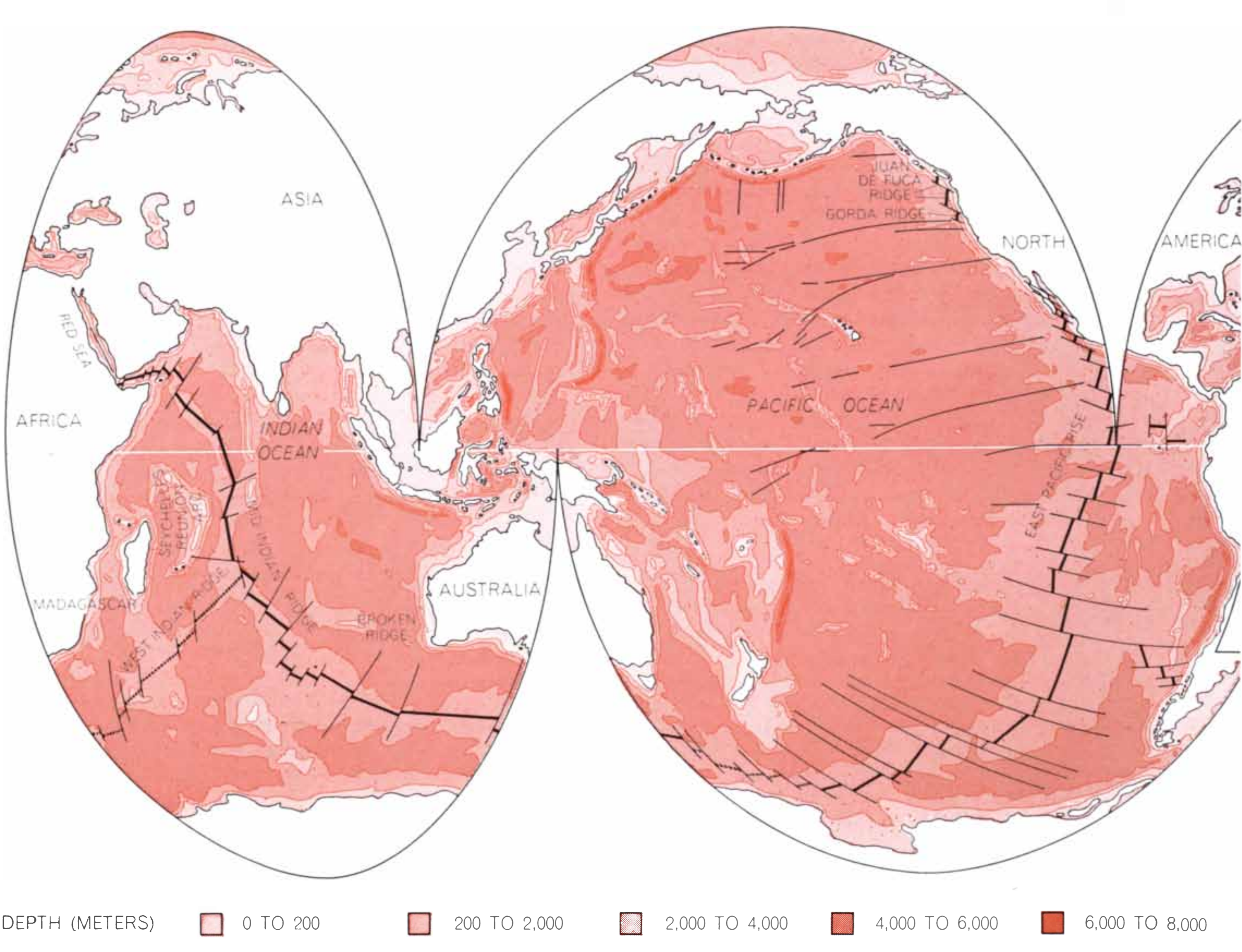
From the 1830's to the 1920's geology was dominated by the theory that the mountain belts were wrinkles caused by the thermal contraction of the earth's interior; then in 1928 the modern form of the convection theory was put forward by Arthur Holmes. Under the name "sea-floor spreading" Holmes's theory currently enjoys great popularity. It is, in fact, far more efficient in providing answers to some of the important questions than its predecessors were. Just when its popularity began to increase rapidly, however, it became apparent that many of its shortcomings could not be expected to disappear in the course of further development: they suggested a basic flaw in the foundations of the theory. Some of the main difficulties cannot be recognized without considering points from the physics of solids that have not been treated in the literature of geology because they were not needed in the past. The purpose of this article is to scrutinize the convection theory, particularly

STRIKING EVIDENCE in support of the hypotheses of sea-floor spreading and continental drift is presented on the opposite page in the form of a colorful "magnetic-anomaly pattern" associated with the Reykjanes Ridge, a portion of the Mid-Atlantic Ridge southwest of Iceland. The color scheme in the area of the pattern is based on the theoretical interpretation by F. J. Vine and D. H. Matthews of Princeton University of a magnetic survey carried out in 1962 by a joint research team from the Lamont Geological Observatory and the U.S. Naval Oceanographic Office. The colored strips correspond to ocean-floor rock formations with a positive magnetization, that is, rock that bears the earth's present north-south polarity. The white spaces between the colored strips correspond to rock formations with a negative magnetization, that is, rock formed during a period when the polarity of the earth's magnetic field was the reverse of what it is today. Presumably basaltic lava emerges from a rift at the crest of the oceanic ridge, and it "paints" its flanks with positive or negative magnetization as the ridge flows apart. Youngest rock is red; oldest is purple. In the theory of sea-floor spreading it is usually assumed that the ocean floor spreads as a rigid plate and that its velocity far from the ridge is the same as that deduced from magnetic-anomaly patterns for the movements of the ridge flanks. In the author's view, however, there is no adequate evidence for this conclusion. On the contrary, the mid-oceanic position of many ridges seems to contradict the possibility of rigid movement of the sea floor.

from the point of view of physics and mechanics, and to consider how it could be repaired or replaced by a picture more in accord with the facts of geology and the laws of physics.

If continental drift is a reality, it represents the most monumental demonstration of the ability of the lithosphere to slide over the mantle. The remarkable similarity of the Atlantic coasts of Africa and South America had excited the interest of such diverse early observers as Francis Bacon and Alexander von Humboldt. It was interpreted as the result of a continental disruption first by Antonio Snider in 1858 and later by Frank B. Taylor. The presence of characteristic fossils such as the ferns Glossopteris in the Gondwana formations of the southern continents has been known for almost 100 years. Around the turn of the century Alfred L. Wegener collected abundant evidence for the common origin of these continents, with particular emphasis on the implications of Permocarboniferous glaciation in those regions of South Africa and South America that face one another across the Atlantic. Finally, Alexander L. du Toit demonstrated in great detail in 1937 the similarity of the Paleozoic layers in Africa (the Karroo formation), South America (the Santa Catharina and other formations) and India. In recent years the paleomagnetic work of P. M. S. Blackett, E. Irving and S. K. Runcorn has supplied almost conclusive proof of movements of the continents, particularly in the course of the past 200 million years.

If two pieces of layer cake fit together along otherwise irregular contours and their layers match remarkably accurately, they are likely to be parts of the same cake. The possibility that they have come from two similar cakes can be eliminated in the case of the earth: the fossil animals and plants of the Gondwana layers could not have developed independently in Africa and in America. The hypothesis of former land bridges linking fixed continents would not explain why the Gondwana flora and fauna are missing on the northern continents, which are now in contact with several Gondwana regions, nor could it account for the fit of the southern Atlantic coastlines. The only explanation available at present, therefore, is that the Gondwana areas are fragments of an ancient continent that drifted apart after the breakup. The absence of Gondwana life on the northern continents (North America and Eurasia without India) suggests that these continents were once widely separated from the Gondwana continent; to attribute this absence to climatic changes would imply vast continental displacements. The similarity of their paleozoic flora and fauna indicates that they were parts of another ancient continent: Laurasia. This is the picture developed by Du Toit. Wegener assumed only one ancient continent, leaving open the question of how the Gondwana and



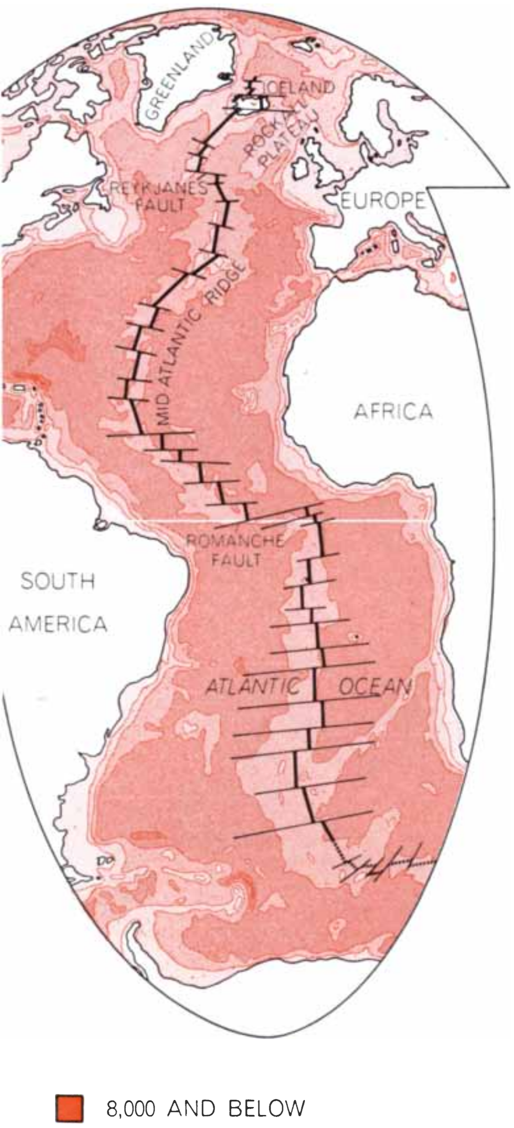
OCEANIC-RIDGE SYSTEM is depicted in two ways on this world map. The colored depth contours give a rough picture of the topography of the ocean bottom. The thick black lines indicate the crest of the ridge, as established by magnetic studies. The thin black

Laurasian floras and faunas could remain so clearly separated.

Continental disruption has also been attributed to an expansion of the interior of the earth. While Gondwana and Laurasia were being disrupted, however, some of their fragments drifted into contact with one another, a process that seems to eliminate global expansion as the cause of the disruption.

Since no serious alternatives and no serious physical objections are known to the assumption of continental disruption and drift, the question arises: How could it be the subject of a violent controversy for more than half a century?

The first and probably main reason for the disagreement was psychological. Man is used to the idea of "terra firma" under his feet. The heliocentric theory of Aristarchos of Samos had to wait for 18 centuries before it was vindicated by Copernicus. Almost equally unpalatable was the suggestion that continents or subcontinents moved with velocities of the order of centimeters per year, although experience provided almost direct confirmation for this. The San Francisco earthquake, and the evidence of displacements by earlier earthquakes along the same fault, showed four years before Wegener's first publication that North America was moving in relation to the Pacific, at least along the San Andreas fault, with a mean velocity close to that deduced by Wegener for the transatlantic drift. It was difficult to realize, however, that a jerky motion, if repeated long enough, would represent a drift. Sir Lawrence Bragg told me recently that in the early 1920's, fascinated by the idea of continental drift, he asked Wegener to write an article on his theory for a British scientific periodical. Bragg gave the translation of Wegener's manuscript to a famous geologist, who provided the only occasion in Bragg's life on which he saw a man literally foaming at the mouth.



lines denote wrench faults, which cut sharply across the ridge, dividing it into segments.

The other reason for the widespread feeling against the idea was that the possibility of continental displacements could not be understood without some knowledge of the mechanical behavior of solids beyond the elastic range. This chapter of physics, however, was in its infancy before the appearance of A. A. Griffith's epoch-making paper in 1920, 10 years after the publication of Wegener's first paper, and at least 50 years before more than elementary problems of the mechanics of solids were to be treated in books on geology. It took another five years after Griffith before R. Becker made the first steps in the physical theory of plasticity, and nine more years to the realization that dislocations in crystal structure play a role in plasticity. Understandably but unfortunately Wegener was not content with supplying weighty arguments for the movement of the continents; he tried to explain its cause by manifestly unsound physical hypotheses. It is equally unfortunate that the unsoundness of a hypothesis is often interpreted as a demonstration of the unsoundness of the facts to be explained.

Typical of the misunderstandings in and around Wegener's hypotheses was the treatment of the question of how the continental plates could move against the resistance of the hard oceanic crust. Wegener gave little attention to this problem; many of his critics, on the other hand, regarded it as a grave and even fatal difficulty. Yet it was pointed out early by Holmes, and concluded from seismic data by Beno Gutenberg, Charles F. Richter and Hugo Benioff, that the continental lithosphere can ride up on the oceanic lithosphere just as an icebreaker rides up on the ice floes. Holmes suggested that the basalt of the depressed oceanic crust would sink into the mantle as it was transformed into the heavier eclogite by increasing pressure.

Altogether the polemics on continental drift are a sobering antidote to human self-confidence. Practically all arguments against it, and many for it, some of them coming from prominent scientists, now appear to have been fallacious.

Most of the recent attempts to explain continental drift and the driving forces of geological events can be traced back to a paper published in 1906 by the Austrian geologist Otto Ampferer. Fortunately Ampferer did not know that seismic shear waves had been identified in the mantle seven years earlier and that they therefore suggested the mantle must be solid. In the belief that it was liquid, Ampferer suggested that mountains were folded and elevated by the viscous drag of thermal convection currents in the mantle. Convection in liquids attracted much attention after the appearance of H. Bénard's famous paper in 1901. That Bénard's beautifully regular "convection cells" resulted from capillarity and not from convection was not known until 1956.

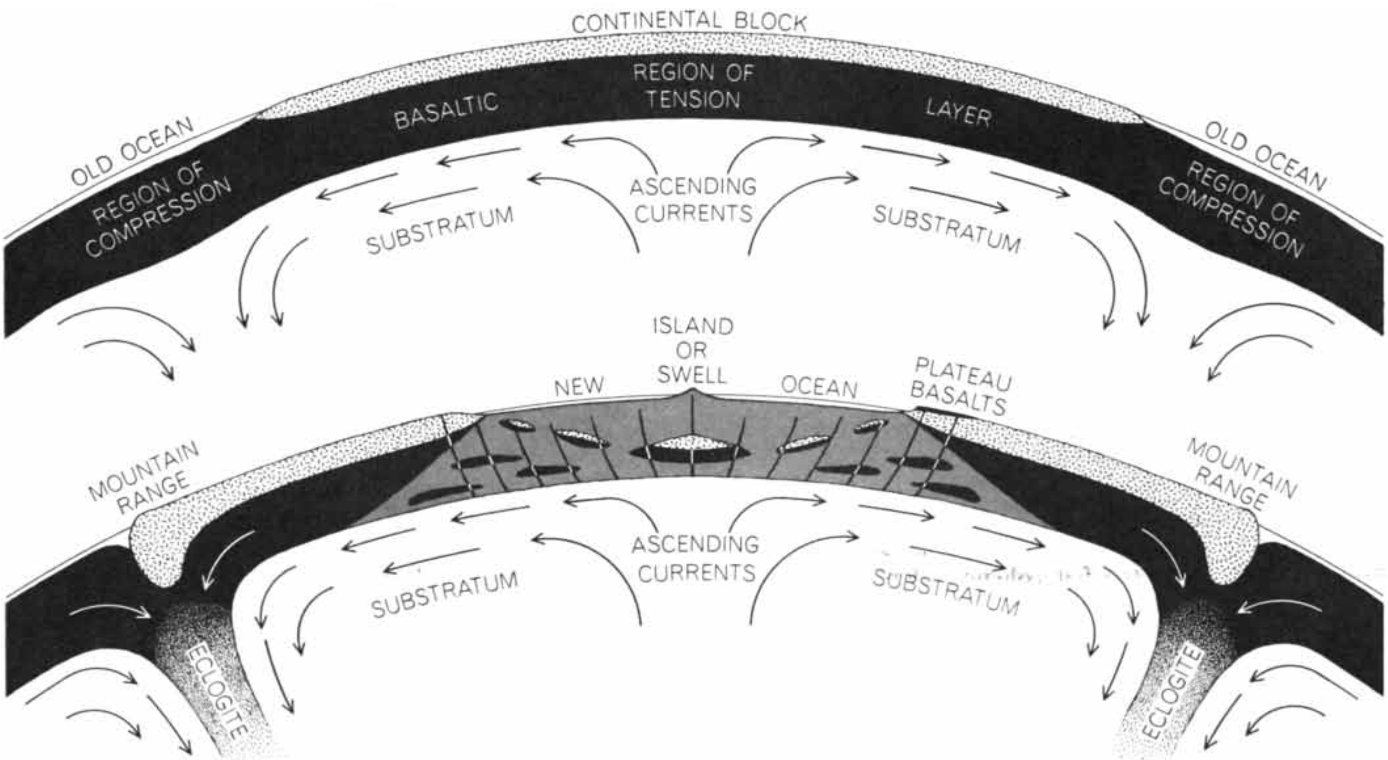
Ampferer's suggestion was shelved for a time because it was thought that convection could not occur in a solid mantle. In 1916, however, Lord Rayleigh investigated the mathematical condition of convection in a horizontal layer heated at the bottom, and 10 years later Harold Jeffreys extended Rayleigh's condition to the case of a spherical shell such as the mantle of the earth. Surprisingly, Jeffreys found that convection due to thermal instability in the mantle was possible even at extremely high values of the coefficient of viscosity, provided that the dimensions of the "convection cell" exceeded some 2,000 kilometers.

Two years after the work of Jeffreys, Holmes put forward the modern form of convection theory. In his original scheme [*see illustration on next page*] Holmes suggested that the oceanic ridge was elevated by the rising hot column of the mantle convection. He assumed, in addition, that the cold column of the convection descended along the oceanic deeps, or trenches; for instance, those in front of the Pacific island arcs and the Andes. "To sum up," he wrote, "during large-scale convective circulation the basaltic layer becomes a kind of endless travelling belt on the top of which a continent can be carried along, until it comes to rest (relative to the belt) when its advancing front reaches the place where the belt turns downwards and disappears into the earth." At these places of descent the basalt would first be transformed to eclogite and sink, and then melt and rise to the surface. "Most of the basaltic magma, however, would naturally rise with the ascending currents of the main convectional systems until it reached the torn and outstretched crust of the disruptive basins left behind the advancing continents or in the heart of the Pacific. There it would escape through innumerable fissures, spreading out as sheet-like intrusions within the crust, and as submarine lava flows over its surface. Thus, in a general way, it is possible to understand how the gaps rent in the crust come to be healed again; and healed, moreover, with exactly the right sort of material to restore the basaltic layer."

According to Holmes's theory (which remained almost unknown to many geologists for 33 years), the oceanic crust is formed from the rising convective column at the oceanic ridge, and then it travels toward the adjacent continents, either pushing them ahead (creating what Holmes called in his original figure a "new ocean") or diving down into the mantle at the continental margin. This theory, which was given the name "sea-floor spreading" by Robert S. Dietz in 1961, received strong support about 1956 when Maurice Ewing, Bruce C. Heezen and their collaborators at the Lamont Geological Observatory recognized that not only the Atlantic but also other oceans have "veins" in the form of ridges that join to form a global network [*see illustration on preceding two pages*]. The theory was independently rediscovered by H. H. Hess in 1960. The only difference between Holmes's picture and later versions of the theory is that in his original figure Holmes drew fragments of the disrupted sialic (silicon-aluminum-rich) continental crust at the point of stagnation where the rising column of convection diverges under the crest of the oceanic ridge. He was perhaps influenced by Charles Darwin's "most extraordinary fact" that in 1845 the German naturalist Christian Gottfried Ehrenberg had found "siliceous-shielded, fresh-water infusoria" in one of Darwin's volcanic rock specimens from Ascension (*The Voyage of H. M. S. Beagle,* Chapter XXI). Although there are apparently no sialic remnants under the crests of the ridges, Holmes was on the right track in suspecting their presence. If there is no stagnation wedge, the physical basis of the theory of sea-floor spreading by convective rise is bound to appear problematical.

In the theory of sea-floor spreading the surface of the earth is subdivided into areas bounded by ridges and trenches. Oceanic crust emerges along the ridges and sinks down along the trenches; each crustal area thus moves in the direction from its trailing edge (the ridge) toward its leading edge (the trench). Like ships frozen into drifting plates of ice, the continents are carried by the plates of oceanic lithosphere in which they are embedded.

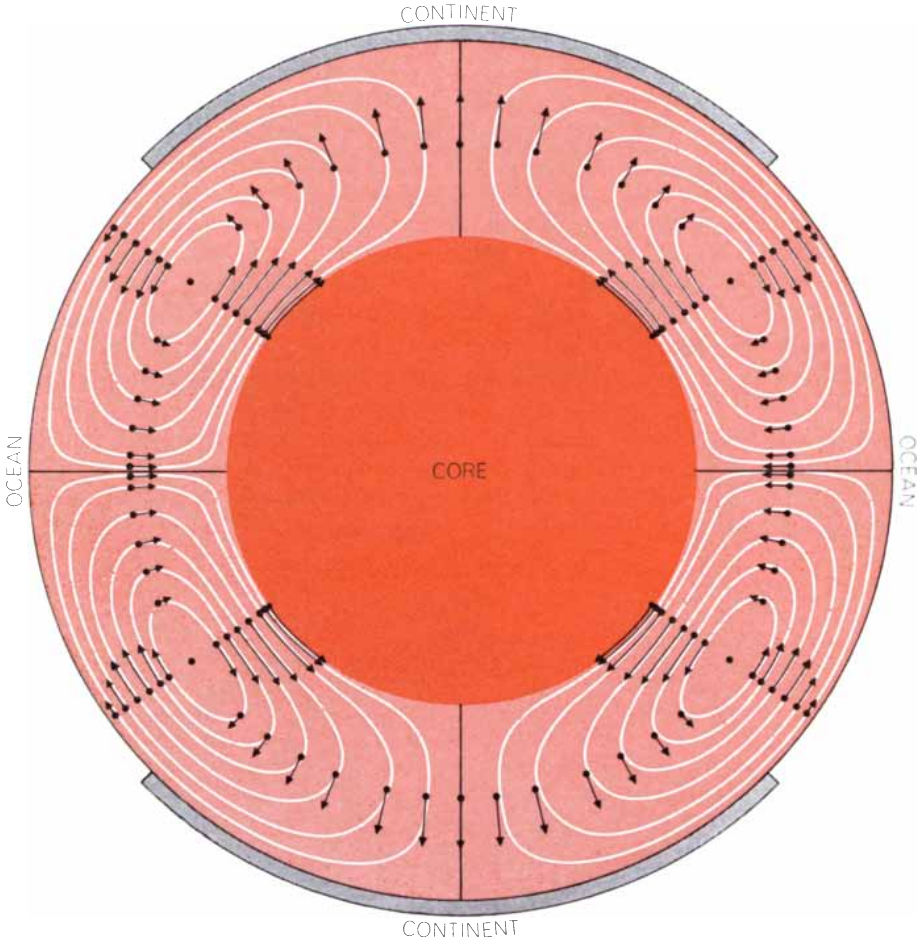
As I mentioned earlier, Holmes's theory and other convection theories lost much of their attractiveness in recent years under the pressure of facts that they could not explain and with which they seemed incompatible. Some of the difficulties were mentioned by Sir Edward Bullard in a recent article in this magazine [see "The Origin of the Oceans," by Sir Edward Bullard; SCIENTIFIC AMERICAN, September]. In what follows a more detailed critique will be given. Nonetheless, the theory did not disappear altogether; like the Cheshire Cat, which faded away while its smile remained, the convection picture was



EARLY CONVECTION MODEL was proposed around 1930 by Arthur Holmes. In this version of Holmes's scheme, adapted from an illustration in his textbook *Principles of Physical Geology* (published in 1944), the formation of the oceanic ridge (referred to as "island or swell" in his original illustration) is attributed to a rising hot convection column in the earth's mantle. The top part of the picture shows an early stage in the convection cycle; the bottom part shows a later stage, during which the two halves of the original continental block are being pulled apart, resulting in the formation of mountains and oceanic deeps in the vicinity of the leading edges, where the convective currents are descending, and the creation of a new ocean floor in the vicinity of the gap, where the convective currents are ascending. Holmes wisely abstained from any assumption about the lower part of the convection "cell."

survived by one of its corollaries. This was the picture of the crust as consisting of rigid plates created at the ridges and sliding toward the trenches. The theoretical model serving as a basis for this conclusion was replaced by an empirical appeal to a recently discovered fact: the existence of strips of magnetic anomalies often symmetrically arranged to the right and left of the ridge crest [*see illustration on page 102*]. An excellent review of the resulting picture, without speculations about any convective mechanism, was given by H. W. Menard in another recent article [see "The Deep-Ocean Floor," by H. W. Menard; SCIENTIFIC AMERICAN, September].

In the most recent literature on the subject convection has been invoked in a very interesting way, which is actually the reverse of the traditional convection theories. In general four types of convection have been considered since the beginning of this century. In convection of the "second-order spherical harmonic" type [*see illustration at right*] the flow is rotationally symmetrical around the vertical axis. Such patterns of convection are at best remotely similar to features of the earth, but they allow exact mathematical treatment, at least if somewhat unworldly assumptions are made about the material of the earth. (The matter is a little reminiscent of the story told about the eminent German mathematician who was said to have solved the problem of the movement of Saturn's rings with the simplifying assumptions that the rings were square and at an infinite distance from Saturn.)



HIGHLY SCHEMATIC MODEL of convection in the earth's mantle was computed in 1935 by Chaim L. Pekeris. The pattern shown, which resembles the second-order spherical harmonic type, was based on the assumption of two opposite warm poles and a cool equatorial belt. The mantle was assumed to have a constant viscosity. Although such models permit exact mathematical treatment, they cannot be directly correlated with geography and geology.

Holmes sketched only the upper part of the convective circuit; how the circuit was to be completed by a return flow deeper in the interior of the earth remained an unsolved problem. According to the deep-convection hypothesis [*see upper illustration on next page*], the convection "cells" would extend over most or all of the depth of the mantle down to the core of the earth. In shallow convection [*see lower illustration on next page*] the cells are confined to the rheosphere. In both cases the hot column rises under the oceanic ridge and the cold column sinks at a trench.

Finally, there is a scheme of convection suggested recently to which the name "new global tectonics" has been given [*see top illustration on page 112*]. It is fundamentally different from the earlier theories in which the subcrustal convective flow was supposed to move the crust by exerting a viscous drag on it. In the new picture, which will be referred to here as crustal-plate tectonics, the crustal plate is moved by some force acting on its trailing edge along the ridge, and the rheosphere flows in the *opposite* direction, so that its viscous drag resists the motion of the plate instead of causing it. The flow circuit is closed by the rigid-body movement of the crustal plate; this movement may be a translation, or a rotation around an axis going through the center of the earth. Such a combination of flow (in the rheosphere) and rigid-body movement (of the crust) may be called a translational convection or, briefly, a transvection.

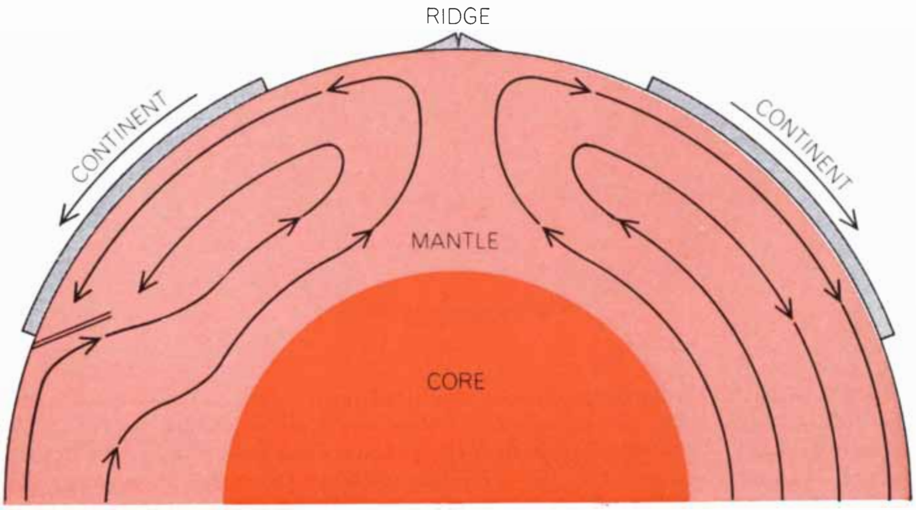
Each of these four types of convection raises difficulties that have not been eliminated so far. Mathematical schematizations of the convective flow in simple spherical harmonic patterns could not be correlated convincingly with geography and geology (except for the first-order pattern, which, in the form of transvection, will be discussed below). Deep convection is made unlikely by the fact that the elastic modulus of the mantle near the core boundary is about 115 million pounds per square inch, approximately equal to the modulus of diamond at room temperature and almost four times that of steel. Now, it is a general empirical rule—easily understood on an atomic basis—that the resistance of materials to plastic and viscous deformation goes remarkably parallel to their elastic moduli: the highly rigid materials are also the very hard ones. In the deep mantle the temperature is high, but there is no reason to assume that the high pressure would raise the resistance to flow and creep much less than it raises the elastic moduli. It is often said that the process of "diffusion creep" offers a way out of the difficulty, but this belief has its root in a misunderstanding. In diffusion creep the viscosity is inversely proportional to the square of the diameter of the individual crystal grains; we know nothing about how large the grains are in the mantle, but one would hardly expect a diameter less than, say, about one kilometer in a material that has been "annealed" at a high temperature for millions of years while its rate of deformation has been extremely low. The possibility of sufficiently rapid diffusion creep was presented in 1965 on the assumption of a grain diameter of about one millimeter; as a consequence many geologists came to believe in the possibility of sufficiently fast diffusion creep in the mantle. Since the most plausible guess of the actual grain diameter is 10^6 times greater than one millimeter, however, the rate of diffusion creep may well be some 10^{12} times lower than the estimate based on the one-millimeter assumption.

The shallow-convection models assume that the flow returns within the rheosphere, so that the centers of the two opposite flows are vertically less than 200 kilometers apart, whereas the horizontal length of the convection cell may be 25 to 100 times greater. Such very thin lamina flowing over one another in opposite directions without breaking up into smaller convection cells seem unlikely.

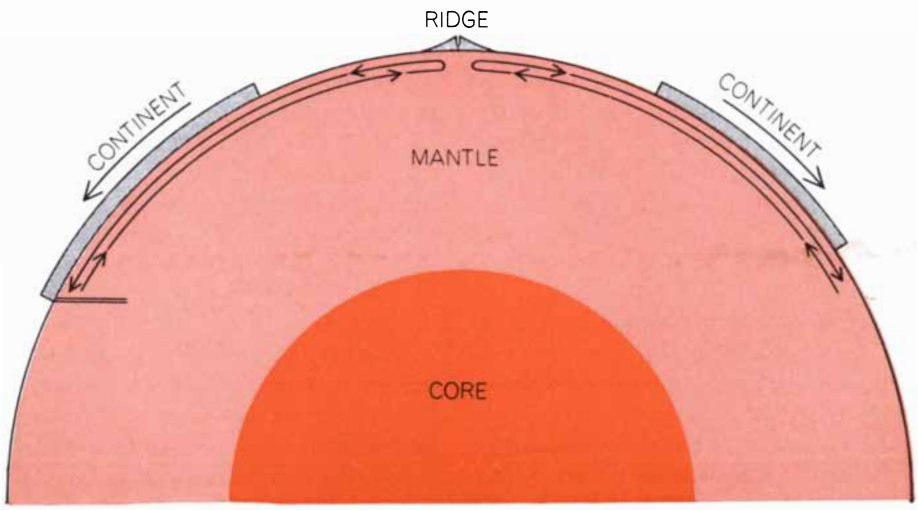
The most recent hypothesis, that of crustal-plate tectonics, is perhaps the most difficult to maintain. The Indian Ocean Ridge has the shape of a vast inverted Y on a map. If the three crustal plates separated by the ridge are driven away from the crests of the ridge, the sea floor moving westward from the ridge would tend to press Madagascar to Africa. Between Madagascar and the ridge, moreover, is the great arc of the Seychelles, Amirante, Mauritius and Réunion islands connected by the Saya de Malha and Cargados Carajos banks. Neither these nor Madagascar is of recent volcanic origin; on the contrary, they contain Precambrian rocks almost certainly of the same origin as Africa itself. The Seychelles-Réunion arc must also have been torn away from the parent continent of Africa by a tensile stress; yet sea-floor spreading from the Indian Ocean Ridge would have had the opposite effect of pressing the Seychelles-Réunion arc and Madagascar to Africa. East of the Indian Ocean Ridge there is the submarine Broken Ridge, which seems to have been torn off Australia; sea-floor spreading from the Indian Ocean Ridge, however, would tend to press it to Australia. The geography of the Indian Ocean could be understood on the basis of the Holmes-Hess convection theory: if Africa and Australia were moved by a viscous drag acting on their undersides, they could be driven apart by an upwelling along the Indian Ocean Ridge, and at the same time their trailing edges could be plucked off if the velocity of convective flow and with it the undertow increased with the distance from the ridge. Sea-floor spreading by pressure acting on the margins of crustal plates, however, would produce the opposite effect. Either the geography of the Indian Ocean antedates the Indian Ocean Ridge, and then the ridge is little more than a recent decoration without much significance, or the hypothesis of crustal-plate tectonics has to be abandoned.

Similar difficulties are encountered in many other parts of the earth. Greenland has been pulled away from North America without evidence of more than an embryonal ridge-rift between them, the Rockall plateau has been pulled away from Europe, and the British Isles and Scandinavia have been pulled away from the rest of the continent, although sea-floor spreading from the North Atlantic Ridge would have pasted them to the mainland.

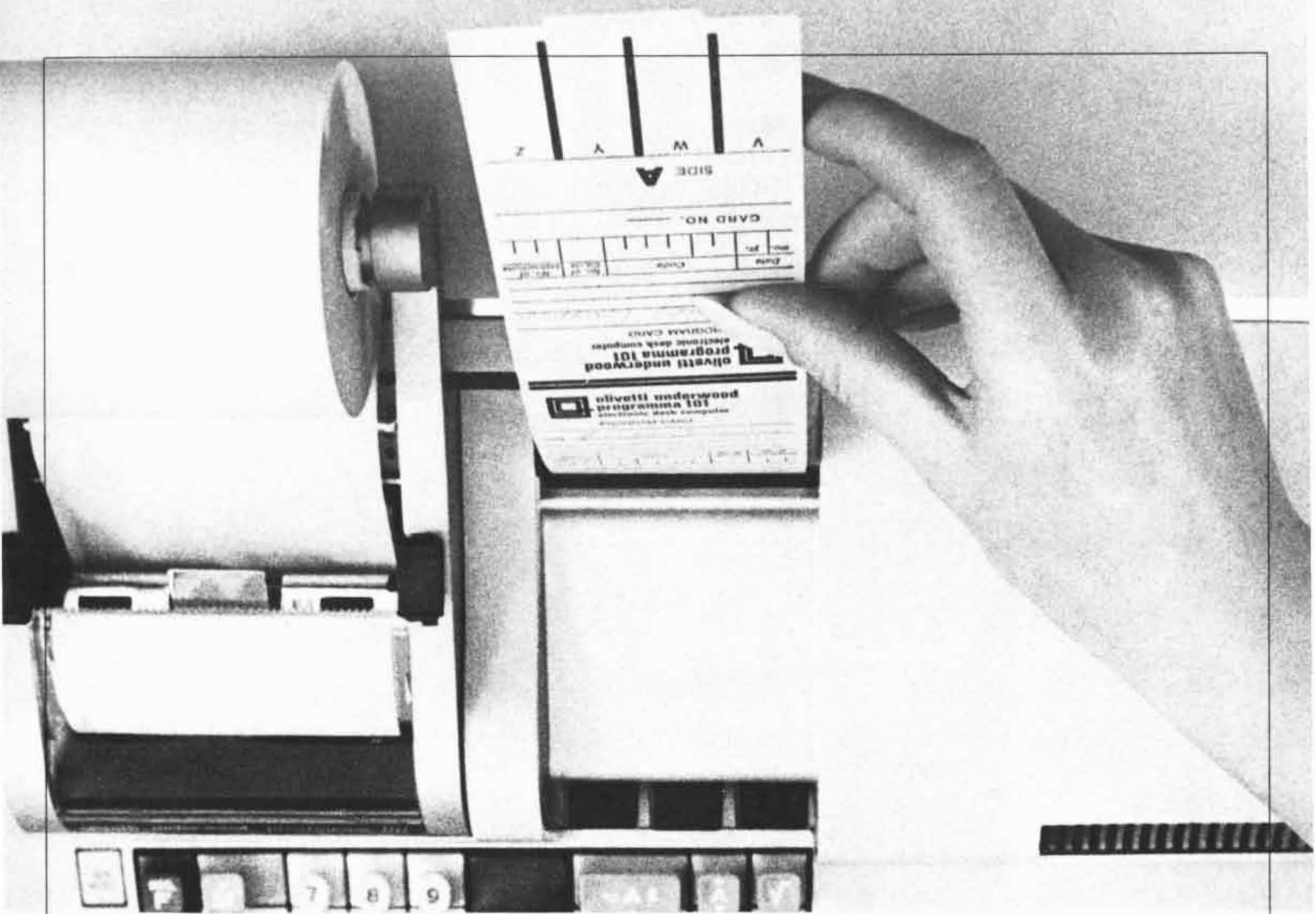
The remark that the British Isles, Scandinavia and the Rockall plateau have been torn away from Europe, Greenland from North America and Madagascar from Africa may seem surprising. After all, they are separated from the mainland only by shallow channels under which the continental crust is thinner but not absent. Apparently the continental crust does not fail under tension with a brittle fracture as a plate of glass does, however; the failure is preceded by the formation of a tensile "neck," a strip of local thinning in which



DEEP-CONVECTION MODEL is simply one way of completing the Holmes theory; the convection cells are assumed to extend over the entire mantle down to the core. As in the Holmes model the hot convective column rises under the oceanic ridge and flows apart, carrying the continents on its back. The two thin lines at left show a seismic shear layer dipping under the continent at a marginal trench. Deep convection is made unlikely by seismic data, which indicate that the deep mantle is too rigid and therefore too hard to convect.



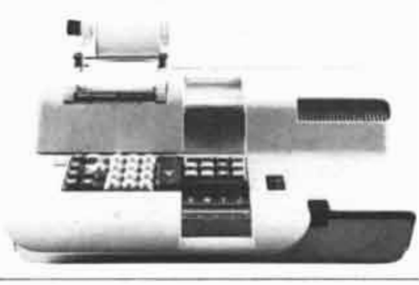
SHALLOW-CONVECTION MODEL is another extension of the Holmes theory, in which the convection cells are confined to the rheosphere. The centers of the two opposite flows are assumed to be vertically less than 200 miles apart, whereas the horizontal length of the cell may be 25 to 100 times greater. It seems unlikely that such very thin lamina could flow over one another in opposite directions without breaking up into smaller convection cells.



The computer that doesn't make you wait.

The Olivetti Programma 101. The do-it-yourself, desk-top computer.

Easy to program. Create your own. Or use ours. Programs are stored on a unique magnetic card about the size of a dollar bill. Applications in statistical and numerical analysis. Mathematics and engineering. Curve fitting. And there are many more. Easy to operate. You enter variables. Programma 101 does the rest. From a simple calculation to a complex equation. Cost? Less than a month's rental for many large computers. Why wait?



olivetti

One Park Avenue
New York, N.Y. 10016
Attention: Marketing Services

Please send information on the P-101.

Name/Title__________

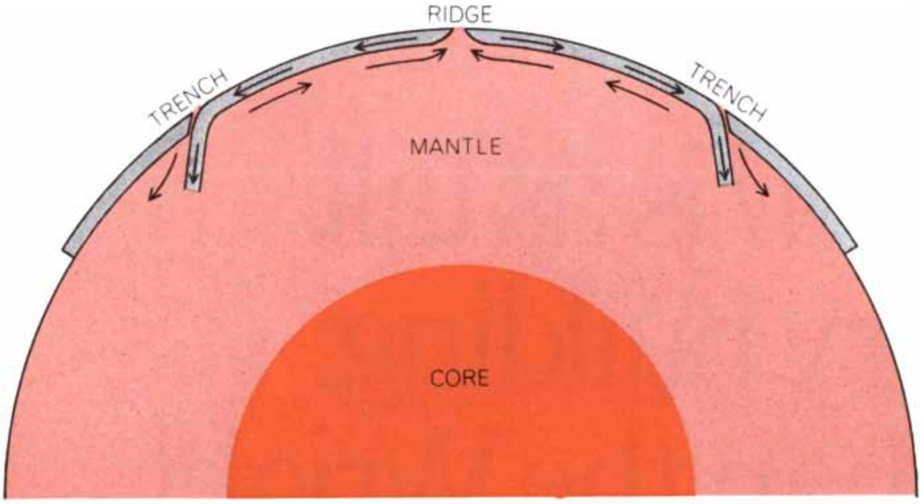
Firm__________

Address__________

City__________ State__________ Zip__________

SA 11-69


747
BOEING
747



MOST RECENT CONVECTION MODEL is fundamentally different from the earlier theories in which the subcrustal convective flow was supposed to move the crust by exerting a viscous drag on it. According to the recent model (to which the name "new global tectonics" has been given), the crustal plate is moved by some force acting on its trailing edge along the ridge. The rheosphere flows in the opposite direction, so that its viscous drag resists the motion of the plate instead of causing it. The flow circuit is closed by the rigid-body movement of the crustal plate. Such a combination of flow (in the rheosphere) and rigid-body movement (of the crust) may be called a translational convection, or, briefly, a transvection. The author cites several objections to this particular form of the transvection idea.

ultimately fracture may occur, as in specimens of ductile metals. Wherever a large island or a peninsula is separated from the mainland by a channel, no matter how shallow it is, there is always a strong possibility that the channel is the initial stage of a tensile failure, the result of a "necking" of the continental crust. No other plausible process of thinning the continental crust along channels has been suggested so far.

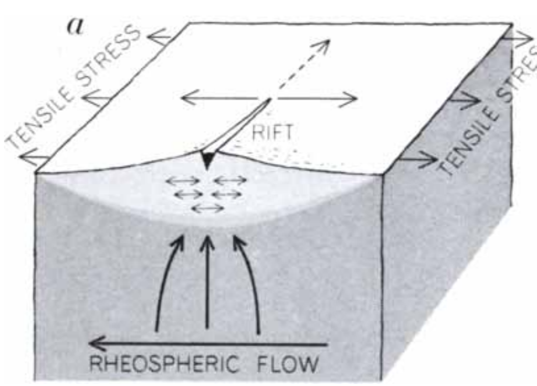
The Holmes theory, both in the deep-convection and the shallow-convection forms, faces yet another difficulty so simple that it is rarely deemed worth mentioning in the literature. The hot column of a convection has probably never been seen to rise in the form of a very long continuous curtain, corresponding to the linear network of the ridges, unless the flow is confined between guiding walls. The upward flow finds much less resistance if it rises in the form of a column of small surface area in comparison with its cross section, as in the puffs of hot air that rise from the ground on a hot morning to form cumulus clouds. Holmes and his successors assumed the existence of a convective rise in the form of long curtains because Bénard's "convection cells" formed a network of ridges on the surface of the liquid. In 1956, however, Myron Block found that the Bénard cells did not arise from convection: they disappeared when a surface-active layer of molecular thickness was spread on the liquid. Evidently they were caused by the temperature-dependence of the surface tension.

Equally damaging to the convection theory is the spectacular fact that the oceanic ridges are not continuous lines: they are cut sharply into segments by what appear to be wrench faults (sometimes called strike-slip faults), which are nearly perpendicular to the adjoining ridge segments. This phenomenon is particularly striking along the North Atlantic Ridge, the segments of which are oriented approximately north-south, whereas the dissecting faults have an east-west direction in the equatorial belt. Since the oceanic crust has no directional properties as a fabric or a veneer of wood does, the segmentation of the ridge cannot have a structural cause. But then how can a rising column of mantle convection disappear without a trace over the 800 kilometers of the Romanche gap in the South Atlantic, or over the 600 kilometers of the Reykjanes fault in the North Atlantic? More drastic still is the case of the very short Juan de Fuca Ridge and the even shorter Gorda Ridge, both off the Pacific coast of North America; they are separated by a fault longer than the Gorda Ridge itself. What kind of convection could produce sharply cut-off short lines of upwelling separated by large distances?

One of the most impressive difficulties is the mid-oceanic position of many ridges (rather more than half of the length of the ridge network is mid-oceanic). The accuracy of centering of the Mid-Atlantic Ridge is astonishing: each segment is individually centered between the continental slopes at each end of the strip, which is framed by the ridge faults at the ends of the segment.

A glance at the Holmes scheme of convection [*page 106*] shows the extent of the conflict between this fact and the theory. The convective flow assumed in the theory "knows" only its sources and sinks, the ridges and trenches; it is "unaware" of the Atlantic coastlines between which the ridge is centered. The ridge, in turn, seems to "know" only the Atlantic coastlines—which play no role in the theory! Equally sharp is the contradiction between the recent crustal-plate tectonics and the mid-oceanic positions of the ridges. If the crustal plates between the ridges and the trenches are rigid units, how can the position of the ridge be determined by the continental margins, which in this hypothesis are little more than lines drawn on the rigid boards of the crustal plates?

For the first hint of a possible answer one must go back to 1955, when Hess suggested that the oceanic ridges might be swellings caused by absorption of water. The mantle probably has the chemical composition of peridotite. Its main component is then olivine and various high-pressure modifications of olivine. Below about 500 degrees Celsius olivine can combine with water to form serpentine; it can take up water up to about a fourth of its own volume, so that the resulting serpentine has a volume about a fourth greater than that of the olivine. Serpentine is a characteristic component of mountain belts of the Alpine type. Hess suggested therefore that the oceanic ridges might be largely products of serpentinization of olivine



FAULT-SEGMENTED STRUCTURE of the oceanic ridges is accounted for in the author's hypothesis by the following mechanism: A horizontal tensile stress first produces a short crack, or rift, at the surface of the crust (*a*). The stress concentration (*small double-headed arrows*) associated

by water rising from the mantle—for instance at convective upwellings. The suggestion was not a popular success, partly because a surface layer of the oceanic crust, about a kilometer or two thick, is mainly a "glaze" of basalt rising molten from the central rift of the ridge and containing too little serpentine. Hess himself did not follow it up and adopted the Holmes theory instead, probably without knowing the historical antecedents.

But suppose now that the convection theory is abandoned and the cue for its replacement is taken from the most spectacular feature of the oceanographic map: the parallel direction of the ridge segments within large areas. This suggests that the segments might have arisen as tensile fractures—cracks, rifts or "normal faults"—produced by a large-scale tensile stress; how such a stress might arise will be discussed below. There are concentrations of tensile stress around the cracks; water and other fast-migrating substances may penetrate both into the crack and into the expanded material around it. They may produce swelling, as in Hess's 1955 picture. A ridge may arise as a "welt" of a lighter material (for example serpentine) floating on the heavier mantle. And now a point of crucial importance emerges. Such a ridge of lighter material tends to flow apart by gravity, just like a heap of mud or a glacier tongue where it leaves the confining valley. The gravity spreading force is easily calculated; with the usual assumptions about the thickness and the viscosity of the soft layer, I found in 1964 that the spreading force of the Mid-Atlantic Ridge would drive South America westward with a velocity equal in the order of magnitude to that of continental drift as deduced from geological, paleontological and geophysical data.

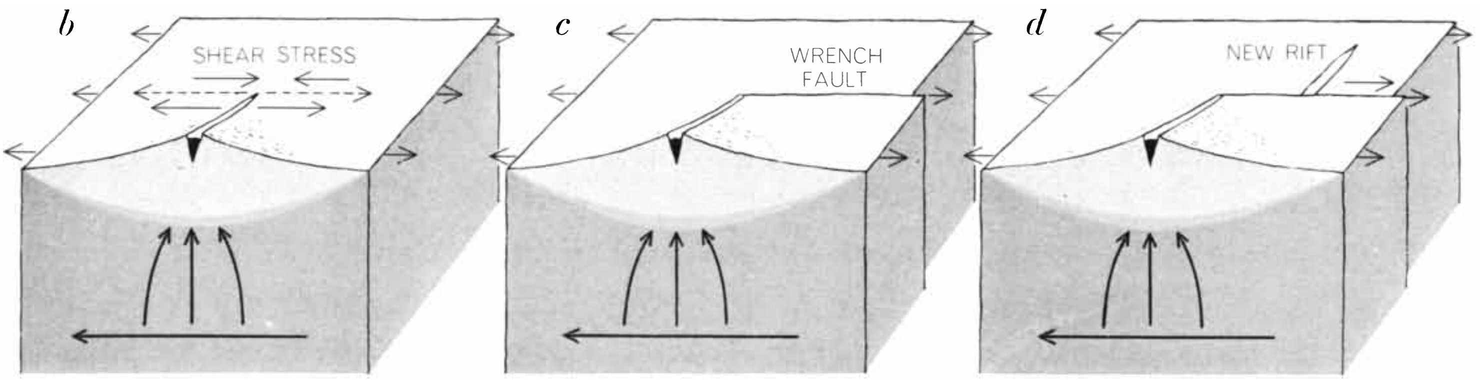
This could change the picture radically. No convection would be needed to drive continental drift; once the ridge is there, it could supply much or most of the driving force. Nor would convection be necessary for building the ridge: it would be raised by the swelling around a crack, an "infection" of the skin of the earth caused by water and other materials migrating toward a wound, the tensile rift [*see illustration below*]. Light or volatile substances such as water move toward the rift and cause a swelling, and the resulting ridge flows apart by gravity. When the lighter components of the material of the rheosphere are extracted by the tension-expanded material of the ridge, the heavy residue sinks back. The flow can be regarded as a small-scale local convective circulation under the ridge; instead of the ridge's being produced by convection, however, it produces its own little convection and it can take the convection wherever it migrates.

If the process is to continue, fresh rheospheric material has to be supplied to the ridge. Such a flow can be maintained by a process (transvection) that does not seem to involve the difficulties of the convection theories.

This mechanism could first of all easily account for the fault-segmented structure of the ridges.

The swelling of the walls of the initial rift is equivalent to a rigid plate's being rammed into it, its thickness being equal to the amount of swelling. The simplest case would be a plate of constant thickness, because the stresses caused by thrusting such a plate into the earth are very well known in the physics of solids. They are the stresses arising from the presence of "edge dislocations" at the vertical edges of the plate. If the edges are well rounded, the material at them is in a state of hydrostatic tension: the tensile stress is the same in all horizontal directions. In addition, shear stresses are present in the plane that contains the edge and is perpendicular to the plate, and their magnitude at the edge happens to be equal to that of the tensile stress. In addition to expanding the rock and promoting the penetration of water, the tensile stress can extend the rift in its own plane. Alternatively, however, the shear stress in the perpendicular plane may produce a kind of tearing by shear in this plane, by the formation of a wrench fault at right angles to the rift. This process is seen when a snowplow is at work. The "rift" arising when its blade is sunk into the snow does not extend in its own direction. Instead two "wrench faults" are produced at the ends of the blade, more or less at right angles to it, by shear fracture between the strip of snow pushed forward and compressed by the blade, and the layers of snow to the right and left of it. Since, in the simple case of swelling just considered, the rift-extending tensile stress and the cross-faulting shear stress are equal, under the usual conditions of soil mechanics accidental circumstances will decide which of the two processes takes place.

If the rift is cut short by cross-faulting, the friction between the fault walls produces a tensile stress in the wall away from the rift; combined with the regional tensile stress, this can sooner or later initiate a new rift on the other side of the cross fault, parallel to the first rift but displaced to the right or left. This rift may again be cut short by another fault, and so on. The tensile-stress mech-



with the rift causes the migration of water and other light substances toward the rift, swelling the rock. The swelling raises a ridge, which flows apart slowly by gravity; the migration of light material toward the rift, however, may raise the ridge as much as, or more than, it is lowered by spreading (*b*). If the rift is deep enough, its two ends act like two "edge dislocations." As hydrostatic tension expands the crust and promotes further swelling, the rift may either be extended horizontally by the tensile component of the stress distribution, or it may be cut off at right angles through the formation of a wrench fault produced by a shear stress in the perpendicular plane (*c*). If cross-faulting does take place, the shear drag between the fault walls produces a tensile stress parallel to the regional tensile stress; the new combined stress field may in time start a new rift from the opposite side of the cross-fault wall (*d*).

anism of ridge formation therefore accounts in a simple way for the rift-fault structure, which was a notorious stumbling block of the convection theories.

Water not only swells some rocks; it dissolves in basalt and reduces its melting point by hundreds of degrees Celsius. The late Jesse Talbot Littleton of the Corning Glass Works, the inventor of Pyrex glass, was so impressed by the low temperatures reported of the liquid basaltic lava of the Hawaiian volcanoes that he traveled to Hawaii with one of his colleagues to find out whether or not the stories were true. In fact, the temperatures were fabulously low, and the prospects for the glass industry seemed revolutionary. They took samples, and when they remelted them in the laboratory at Corning, the miracle had disappeared and the melting point was higher than in Hawaii by hundreds of degrees. The water had evaporated from the basalt.

The tensile-rift mechanism also offers an explanation for the mid-oceanic position of many ridges.

If the lithosphere consisted of rigid plates bounded by ridges and trenches, a ridge would not move into mid-oceanic position because it could not "feel" where the coasts and the middle of the ocean were. It must be concluded, therefore, that the oceanic crust is horizontally compressible even if the compression does not produce easily observable or seismically active faults. As the ridge flows apart it will then compress the ocean floor on each side. Where the continent is far away, there are long strips of sea floor between it and the ridge, and they will be shortened by compression more than the strips on the opposite side of the ridge, where the continent is nearer. The ridge, therefore, will not spread symmetrically: it spreads more toward the distant continent, and its center line moves toward the mid-oceanic position. If a ridge segment is between two faults, the wrench fault on the side of the distant continent moves more, and the chances are that the next new rift beyond the fault will arise closer to the distant continent. It is an interesting question whether the numerous parallel ridges on the flanks of a ridge were perhaps old central rifts inactivated by the birth of a new parallel rift.

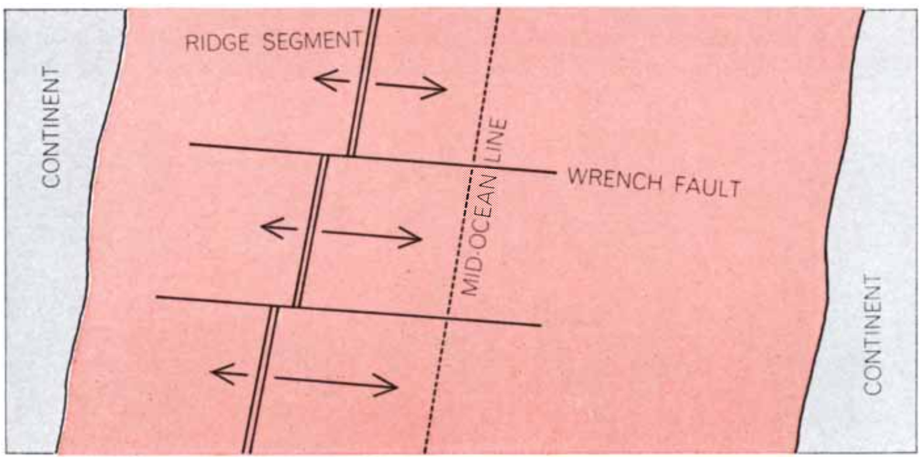
Closely related to the problem of the centering of the ridges is a phenomenon that is at the center of attention in geology at present: the strips of magnetic anomalies that run parallel to the ridges and are frequently remarkably symmetrical on each side of the crest. The details of this phenomenon have already been treated in the articles by Bullard and Menard, so that only a few words have to be said about it here. The adjacent strips of positive and negative rock magnetization are attributed to changing signs of the earth's magnetic field. Basaltic lava emerges from the rift at the crest of the ridge and "paints" the flanks with layers of alternating magnetization as the ridge flows apart, usually with velocities between one and 10 centimeters per year on each flank. This discovery, made by F. J. Vine and D. H. Matthews in 1963, is often regarded as a striking verification of the Holmes theory of convective upwelling under the ridges. In reality, however, no way has been found so far to reconcile it with the theory. The problem can be recognized clearly from Menard's article. He pointed out that the magnetic strip pattern may be symmetrical even if the velocity of spreading is different along the two flanks of the ridge. All that is needed is that the rift migrate with a certain velocity. If it is filled with lava and this freezes, it should crack and form the new rift in the center, even if, for example, one wall of the old rift has not moved at all, so that the strip of lava is off-center with respect to the position of the rift before it was poured out. The "spreading center" therefore has to move with a velocity that is the algebraic mean of the velocities of the two rift walls; in this case the magnetic lava strip pattern will be symmetrical.

Menard's condition shows the predicament of the convection theory very lucidly. If the two crustal plates adjoining along the ridge are rigid, they will be driven apart by whatever force acts on them at velocities determined by their size, shape and other factors. If the spreading center between them had an independent existence, for instance as a convective upwelling, there is no reason why its velocity of migration should be just one-half of the algebraic sum of the velocities of the two plates. If the strip pattern is symmetrical because Menard's condition is satisfied, this seems the most direct disproof conceivable of the convective genesis of the ridge. It shows that the ridge is not made by convection; on the contrary, whatever convection there is around the ridge is generated by the ridge itself.

Menard's principle also shows that velocities of "sea-floor spreading" cannot be deduced from magnetic strips unless the velocity of migration of the ridge is known. If flank velocities of 10 centimeters per year are deduced from the strip pattern, one flank may move at 20 centimeters per year and the other stand still, or the other way round, or anything in between.

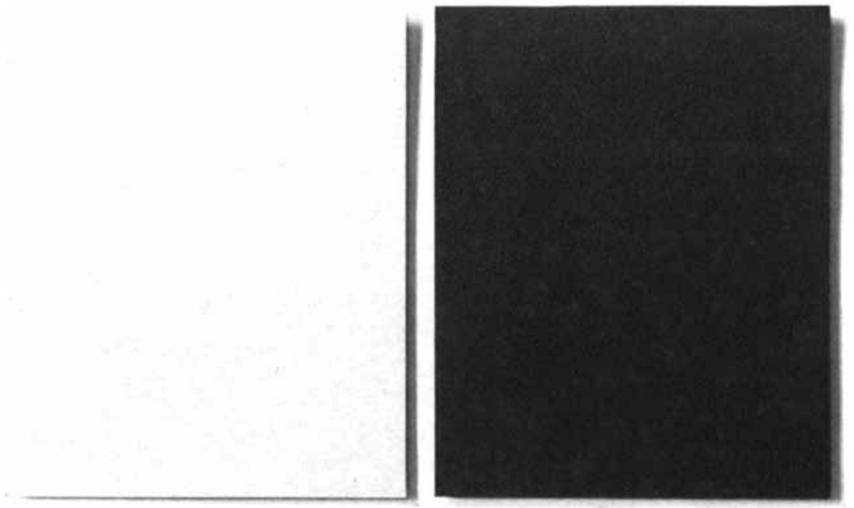
If the oceanic ridge is initiated by a tensile fracture, what may be the origin of the necessary tensile stress? The ridges of the Atlantic Ocean and the Indian Ocean are much too large to be attributed to local effects. But if the deep mantle is too hard to convect and the rheosphere too thin to accommodate convection cells of vast horizontal extent, how can stresses of global dimensions arise?

One way in which a worldwide flow



SELF-CENTERING MECHANISM responsible for keeping more than half of the oceanic-ridge system (including almost all the Mid-Atlantic Ridge) accurately centered between the continental slopes can be explained by assuming that the oceanic crust is horizontally somewhat compressible. As the ridge flows apart it presses with the same force toward both flanks. The amount of compressive shortening, however, will be greater where the strip of the ocean floor between the ridge and the continent is longer. Consequently the center of the rift will tend to move toward the more remote continent. The assumption of crustal compressibility can also be used to show that each successive ridge segment will probably arise in the side of the previous segment that is closer to the mid-oceanic position.

The difference between ordinary paper and the sophisticated materials Riegel develops from it is the difference between white and black.



We started with paper and came up with a thin, opaque, contamination-free material. What can we start for you?

Nothing is quite as unremarkable as what you're looking at now, right? Good old paper. Common as air.

Well, Riegel's business is first to make paper—and then to make it remarkable.

For instance, if you wanted it absolutely free of radiation, we'd drill 450-foot wells to get chemically pure water.

Moreover, if you wanted this same paper completely opaque, so that even the most microscopic pinholes were eliminated—well, we'd work in our laboratories until we'd made it completely opaque.

And if, in addition, you wanted it to be machinable to extremely tight tolerances, to be compatible with many other materials, to possess exceptional strength and a thin uniform caliper, we'd take care of those details too.

Matter of fact, we've already done it for the photographic industry.

From tree to technology

Light-tight materials are some of many we pioneered for the photographic industry. Others include moisture-proof black wrap; dri-mount tissues and heat-seal coatings; x-ray inter-leaving; slide mounts; pouch wraps; and colored and black duplex for film envelopes.

In each case Riegel's paper machines and converting equipment produced the substrate, or basic paper stock. But at this point, where most paper mills leave off, we like to start in.

Stumped? Tell Riegel.

You might be surprised what can be done with paper. Particularly in the hands of a research-minded company that knows paper backward and forward. Particularly when it's Riegel. When you are really stumped for a solution, and you're practically certain paper can't do it, come to Riegel. If any one can, Riegel can. Riegel Paper, Technical Papers Group, 260 Madison Avenue, New York, N.Y. 10016.

Riegel

Much more than paper.

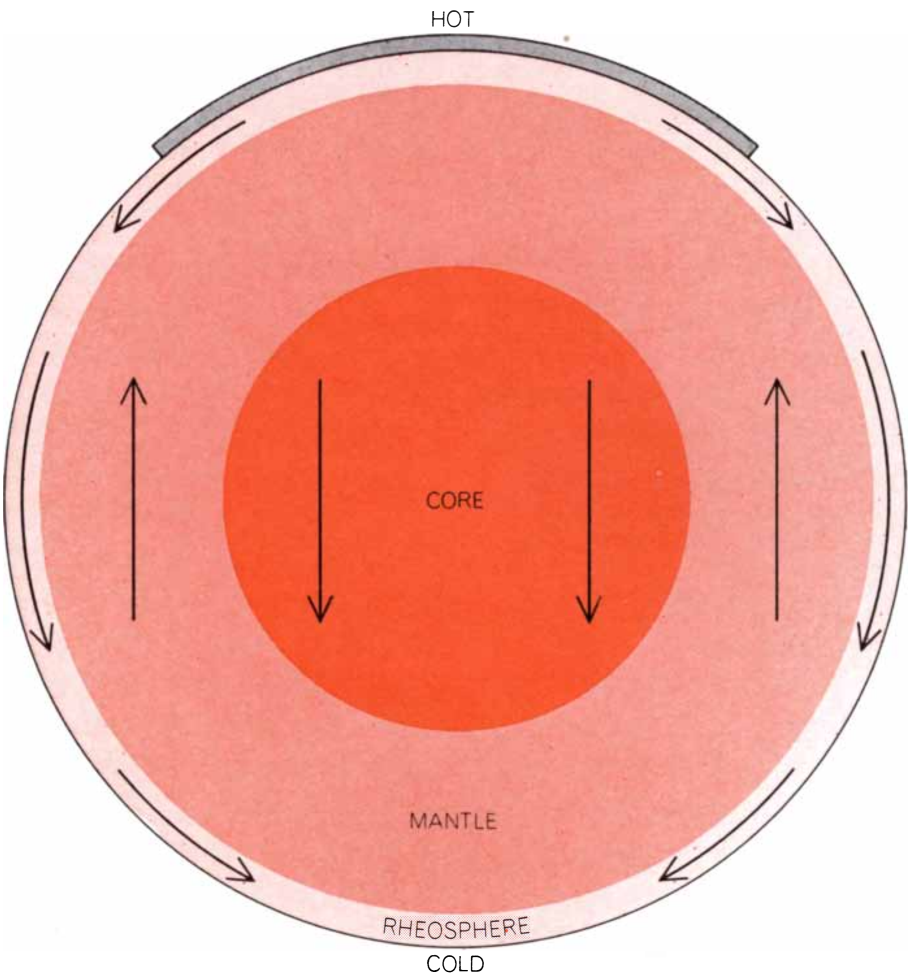
can take place in the rheosphere without any return flow either in the deep mantle or in the rheosphere itself is illustrated below. If the firm lower part of the mantle moves in the direction opposite to the flow in the rheosphere, it closes the circuit by means of a rigid-body movement. This is another case of what was referred to above as a translational convection, or transvection, although in this case the movements are quite different.

A transvective circulation of this type may be caused by local heating in one region, and it may continue for a long time if there is little difference between the chemical composition of the mantle and the rheosphere, so that the rheosphere is soft merely because it is under a lower pressure. In this case, although the firm mantle moves from a cooler region toward the heated region, the boundary between the mantle and the rheosphere stays in place because the mantle softens to rheosphere as it rises at the cooler region, and the rheosphere solidifies to mantle as it sinks to depths of higher pressure at a cooler region. Of course, the mantle carries with it the iron core, but blocks or grains or dissolved ions of the mantle may be transported fast enough by the liquid core in the direction from the cooler to the hotter regions in order to reduce the gravitational energy of the core and thus enable the core to stay close enough to its central position.

For many years it has been customary to regard radioactivity as the main source of heat to drive convection. Nonetheless, a more adventurous possibility cannot be excluded. The round maria of the moon seem to have been created by the impact of very large (asteroid-sized) meteors. Since most of the maria are on the visible side of the moon, the meteors may have belonged to one swarm, which struck within a very short time (not more than a few days), since otherwise the hits would be more widely spread on the moon. According to an analysis of data gathered in the course of recent lunar missions, the round maria are sites of mass concentrations ("mascons") indicated by gravity anomalies. The greatest of these may have been produced by asteroids some 100 kilometers in diameter. With a velocity of, say, 30 kilometers per second and a specific gravity of six, the kinetic energy of such a large ball is about 3.5×10^{35} calories. If the mean specific heat of the rocks in the outer layers of the earth is taken as .2, this energy would raise the temperature of a circular layer of the crust and the rheosphere 100 kilometers thick and 3,000 kilometers in diameter more than 1,800 degrees C. Evidently, then, a large impact such as the moon has suffered could start a very intense prolonged rheospheric flow and continental drift in the earth. In view of the drastic happenings in the Indian Ocean that have led to the flight of India to Asia and of Australia away from Gondwana, the possible role of such rare or unique processes in geology in starting a continental disruption and perhaps even the transition between the Precambrian and the Cambrian periods cannot be discounted. In spite of the several hundred million years between the Precambrian and the intensification of continental disruption, the two events might have been started by the same asteroidal impact.

The possibility of geological effects resulting from meteoritic or asteroidal impacts on the earth has already been considered by Heezen and his colleagues at Lamont; they suggested that geomagnetic reversals might have been triggered by such events [see "Tektites and Geomagnetic Reversals," by Billy P. Glass and Bruce C. Heezen; Scientific American, July, 1967]. There is some possibility that the asteroid swarm that created the maria on the moon might also have initiated the Cambrian period on the earth.

It is easy to verify that the mean velocity of the rheosphere as it flows from a hotter to a cooler region increases in proportion to the sine of the "latitude" counted from zero at the hot pole to 180 degrees at the cold pole. Consequently

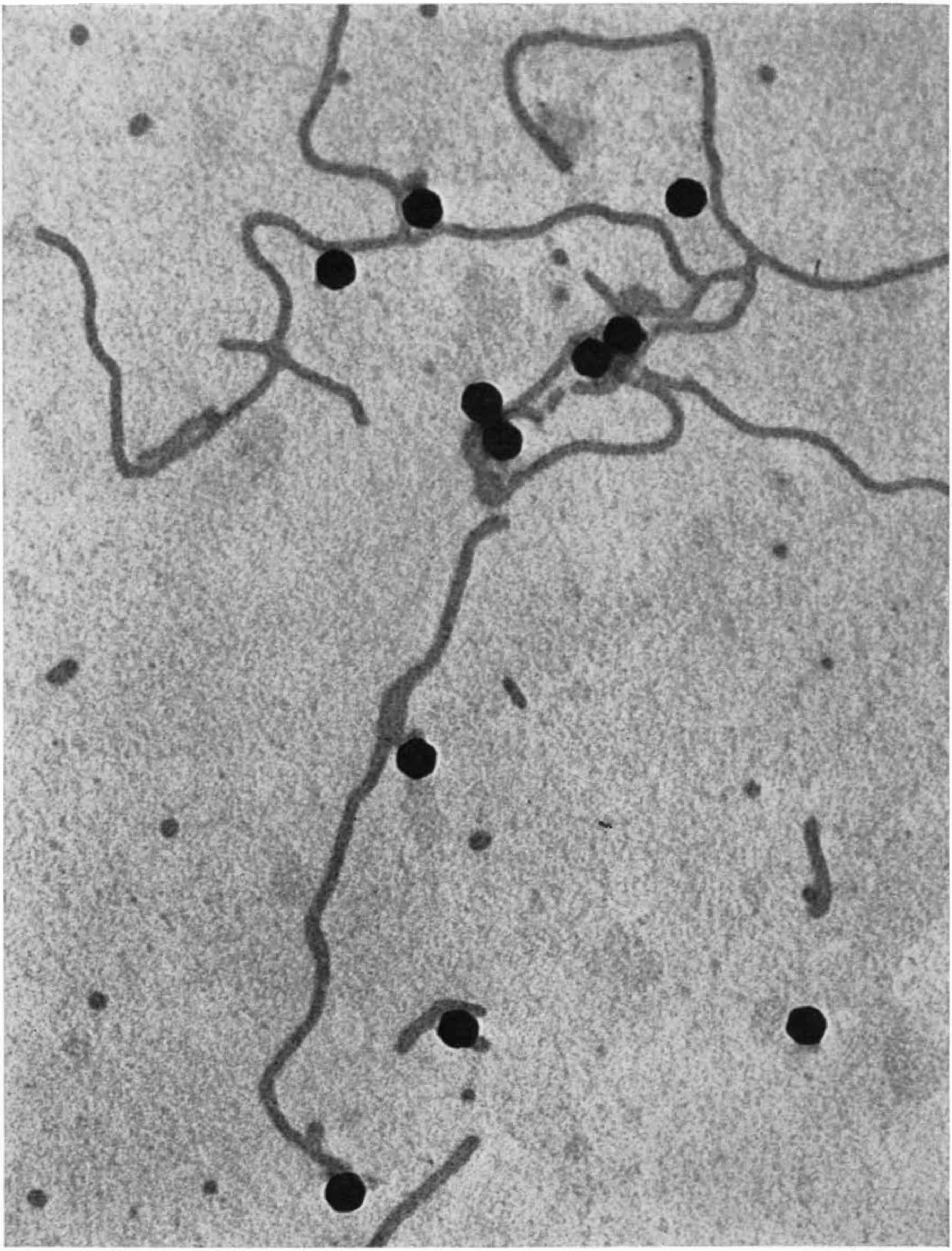


GLOBAL-TRANSVECTION MODEL is suggested by the author as one way in which a worldwide stress-producing flow can take place in the rheosphere without any return flow either in the deep mantle or in the rheosphere itself. Local heating in one region of the earth (caused either by radioactivity or by a large meteoritic impact) could cause an expansion of the rheosphere and hence a prolonged rheospheric flow in the direction of an antipodal cold region. If the firm lower part of the mantle moves in the direction opposite to the flow in the rheosphere, it closes the circuit by means of a translational, or rigid-body, movement. At the same time the iron core moves back toward the center of the earth, as material is detached from the mantle and carried by fast convection in the liquid core in the direction from the colder region toward the hotter surface region. According to the author's hypothesis, the entire process of continental drift could have been set off perhaps as early as the end of the Precambrian period, with the center of disruption (corresponding to the hot pole in this scheme) located somewhere in the northwestern part of the Indian Ocean.

the viscous drag exerted by the rheosphere on the crust increases from zero at the hot pole to a maximum at the transvective equator between the poles, and then diminishes to zero at the cold pole.

If the ancient continent of Wegener, or the two continents of Du Toit, occupied approximately the central part of what is the continental hemisphere today, and the center of disruption corresponding to the hot pole in this scheme was somewhere in the northwestern part of the Indian Ocean, the shear drag on the underside of the continents increased with the distance away from the hot pole. The opening of the Atlantic is then understandable, and also the disruption of Africa from Madagascar and from the Seychelles-Réunion arc, and of the Broken Ridge from Australia. The difficulties of the hypothesis of sea-floor spreading are avoided. What remains unexplained is the youngest part of the ridge network, the East Pacific Rise. This, of course, could hardly be a consequence of the original Wegener-Du Toit continental disruption, with the transvective hot pole somewhere in the present Indian Ocean; it must be a late secondary effect of the disruption.

This crude model offers a possibility of understanding a striking feature of the earth: the equatorial position of the Pacific Ocean, which is so accurately centered between the North and South poles. It was stated above that gravity could keep the core in the center of the earth in spite of transvection going on. Of course, the core would always remain a little off center, in the direction toward the hot transvective pole. This means that the centers of gravity of mantle and core are pulled apart a little along the line connecting the hot and cold transvective poles; this line is then the axis of smallest moment of inertia of the earth, and the directions at right angles to it are the axes of maximum moment of inertia. Now, the rotation of a solid is most stable around an axis of maximum moment of inertia. This is the reason space capsules are built as very flat cones to prevent their tumbling around a transverse axis. In the case of the earth the axis of rotation will be perpendicular to the diameter between the transvective poles; if the cold pole is near the center of the Pacific, according to the assumptions made about the mode of disruption, the Pacific will be more or less centered between the North and South poles, with its center on the Equator. In fact, the center of the Pacific seems to lie within five or 10 degrees of latitude on the Equator.



BACTERIAL VIRUSES, or phages, are adsorbed to the surface of cells they infect at specific receptor sites that have a particular chemical structure. In the case of phage ε^{15} the site is in the lipopolysaccharide (fats and sugars) component of the surface of the bacterium *Salmonella anatum*. In this electron micrograph, made by Cecil E. Hall of the Massachusetts Institute of Technology, ε^{15} particles, enlarged about 185,000 diameters, have become bound to ribbons of lipopolysaccharide isolated from *S. anatum* cells. Such particles do not become bound to lipopolysaccharide isolated from other bacterial strains that are not subject to infection by this virus.

The Receptor Site for a Bacterial Virus

A virus infects a bacterium by attaching itself to a specific site on the bacterium's surface. The chemistry of such a site is studied by seeing how it changes after dormant infection by the same virus

by Richard Losick and Phillips W. Robbins

The critical first step in the infection of a bacterial cell by a phage, or bacterial virus, is the attachment of the virus to the surface of the cell. The precise sequence of events in that process has been most clearly outlined in the case of the phage designated T4 and its host, the bacterium *Escherichia coli.* Phage T4 is a complex virus consisting of a head that contains the genetic material, a tail through which the material is injected into the cell and six long tail fibers. Electron micrographic studies, such as those conducted by Lee Simon and Thomas F. Anderson of the Institute for Cancer Research in Philadelphia, show that the phage's tail fibers first make contact with the cell and then bring a pronged plate at the base of the tail into close proximity with the surface. After the base plate is fixed to the cell surface the tail contracts and the phage genetic material is injected into the cell, leaving the empty virus particle behind [*see top illustration on next page*].

Bacterial viruses vary in structure, and many of them are quite different from T4; no doubt their modes of attachment differ too. One important generalization can be made, however, about phage adsorption to the surface of a cell: It is a highly specific process. A particular phage generally attaches itself only to a limited range of bacterial strains. Moreover, a bacterium that adsorbs many kinds of phages may mutate, giving rise to a new strain that no longer adsorbs one of these phages but still serves readily as a host for others. The surface of a bacterium is apparently a mosaic of different receptor sites, each of which is able to form a highly specific bond with a particular phage.

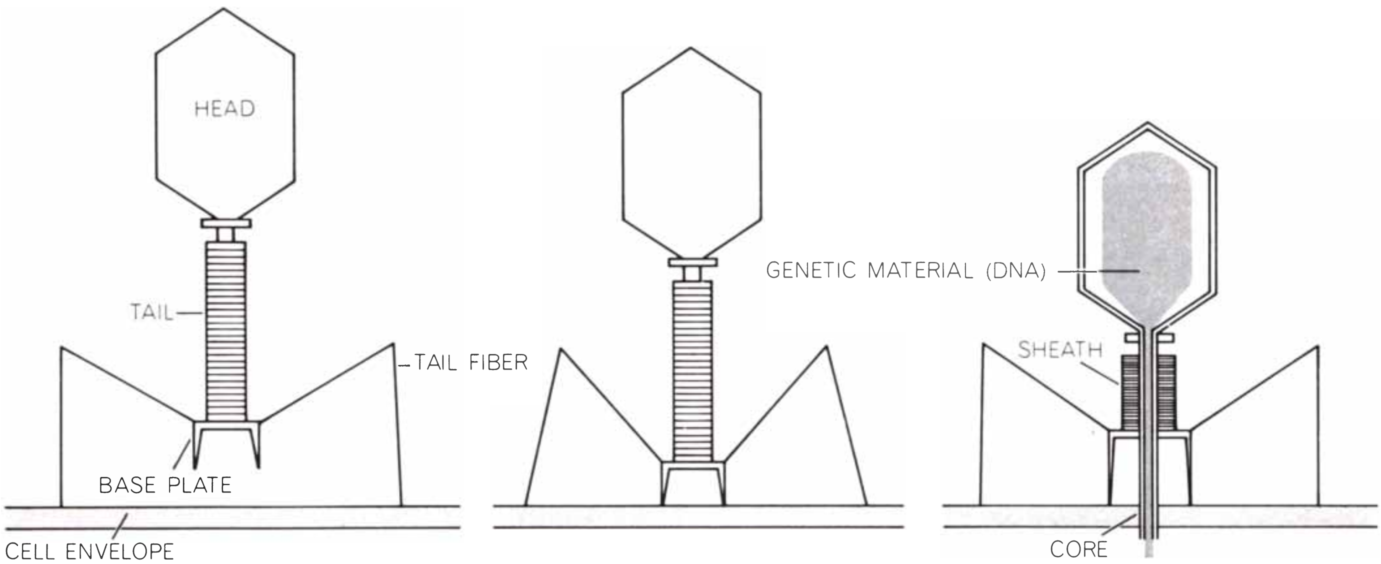
It is clear that receptor sites are located on and characteristic of the surface of bacteria. T4, for example, is adsorbed by empty cell walls of *E. coli* from which the cellular contents have leaked. Some strains of phage are adsorbed by particular sites on appendages of the cell surface. Phage f1 attaches itself to the tips of hairlike structures known as F-pili on the surface of male bacteria; phage R17 attaches itself to the sides of these pili and phage χ attaches itself to the whiplike flagella of certain other bacteria.

The nature of these various receptor sites has remained obscure in spite of considerable investigation ever since the early days of phage genetics. Much was learned about the details of phage infection, and the specificity and diversity of the sites was amply demonstrated, but it was not possible to describe the architecture of even one such site at the molecular level. Then in 1952 Shoei Iseki and Tatsuo Saki of Gunma University in Japan discovered a new bacterial virus, ε^{15}, that is adsorbed by only a few bacteria. One of them is *Salmonella anatum.* In the past 10 years a great deal has been learned about the structure of a particular component of *S. anatum*'s surface. In our laboratory at the Massachusetts Institute of Technology we have been able, by studying the relation between ε^{15} and this component, to determine some structural details of the ε^{15} receptor site. At the same time, and from another point of view, we have made use of the process of phage infection to analyze the chemistry of the bacterial surface.

The cell envelope, the outer portion of the bacterial cell, is a complex structure consisting of an inner plasma membrane and a rigid mucopeptide layer, the cell wall proper, that confers strength and shape [see "The Bacterial Cell Wall," by Nathan Sharon; Scientific American, May]. The gram-negative bacteria, such as *E. coli* and the salmonellae, have in addition an outer membrane, composed in part of lipopolysaccharide, a combination of fats and sugars. It was found that ε^{15} particles are bound to lipopolysaccharide extracted from *S. anatum* but not to the same kind of material extracted from other bacteria that do not adsorb ε^{15} [*see illustration on opposite page*]. This demonstrated that the lipopolysaccharide fraction of the outer membrane contains the receptor site.

One component of this fraction is a long polymer, or molecular chain, composed of the sugars galactose, rhamnose and mannose repeated in sequence. This polysaccharide chain, known as O-antigen, was shown to be a necessary part of the receptor site for ε^{15} because mutants of *S. anatum* that do not synthesize O-antigen also fail to adsorb ε^{15}. O-antigen has two structural features that turn out to be important in receptor-site chemistry. Sugars are joined to one another by either of two linkages, known as alpha and beta; the galactose units in O-antigen are joined to the adjacent mannose units by an alpha linkage, in which the linkage to mannose is below the plane of the galactose unit. And the galactose units in O-antigen have a substituent, an acetyl group, attached to them [*see bottom illustration on next page*].

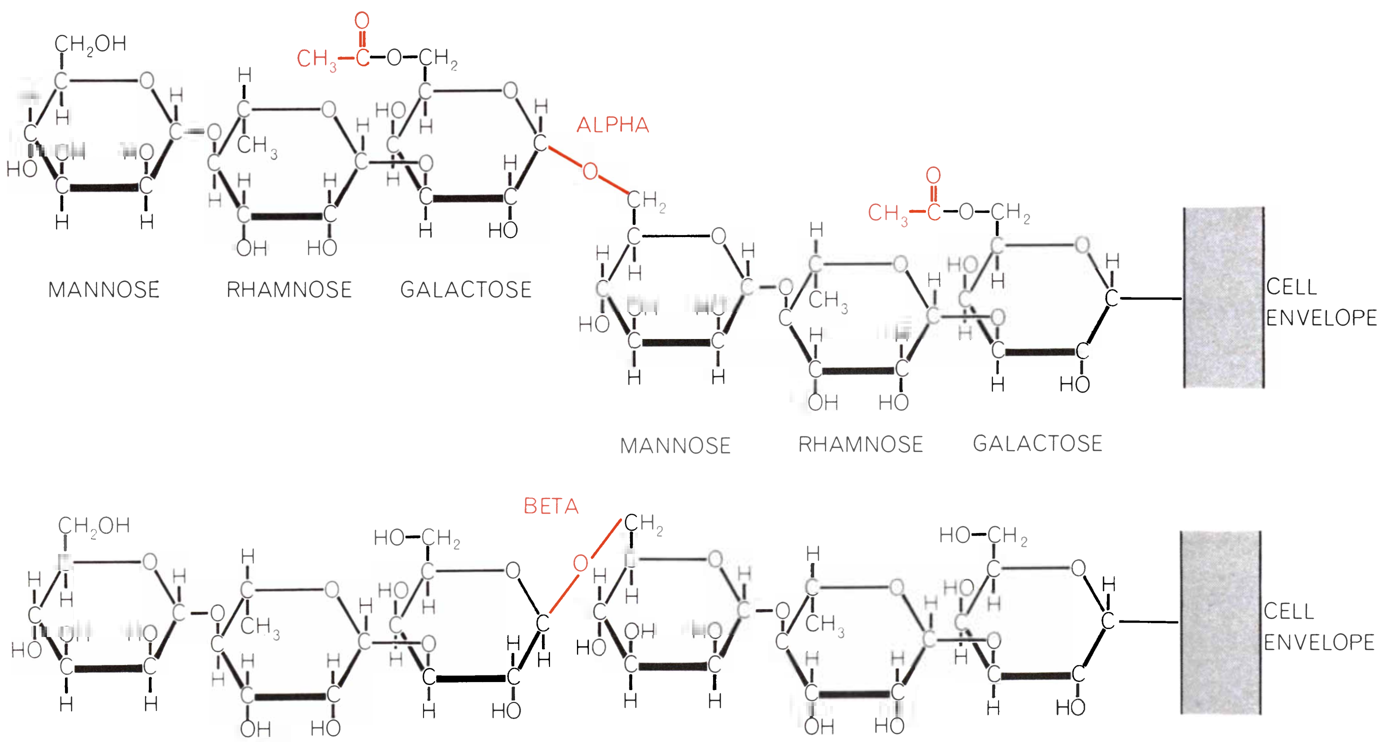
Working with one of us (Robbins), Andrew Wright and Marcello Dankert unraveled the biochemical steps involved in the synthesis in cells of the O-antigen chain. They linked radioactively labeled sugars by means of enzymes taken from *S. anatum* and found that in a cell-free laboratory system the galactose, rhamnose and mannose are not immediately joined to one another in repeating chains. There is an intermediate step. A carrier molecule—a lipid, or fat—serves as a site on which the trisaccharide galactose-rhamnose-mannose is assembled. Then the trisaccharide



PHAGE T4 attaches itself to the surface of a bacterial cell in a series of steps. First the long tail fibers become affixed to the surface of the cell (*left*). Then the base plate is brought into contact with the surface (*middle*). Finally the tail sheath contracts, driving the phage's tubular core through the cell envelope, and the genetic material in the head enters the cell through the core (*right*).

units are polymerized into long polysaccharide chains by the enzyme alpha-polymerase, the galactose of one trisaccharide being attached to the mannose of another unit in an alpha linkage. (In a mutant that lacks alpha-polymerase the trisaccharide units, linked to their lipid carriers, simply accumulate in the cell envelope and no O-antigen is formed.) After polymerization another enzyme, transacetylase, attaches acetyl groups to the galactose units. Finally the O-antigen polymers are released from their lipid carriers and become attached to a new site on the outer surface of the cell [*see top illustration on opposite page*].

There is a special form of virus infection of a bacterium in which the phage genetic material, instead of subverting the machinery of the cell to produce hundreds of virus particles and kill the cell, enters into a dormant state and is integrated into the bacterial chromosome. This dormant form of infection is called lysogeny [see "Viruses and Genes," by François Jacob and Elie L. Wollman; SCIENTIFIC AMERICAN, June, 1961]. In 1956 Hisao Uetake of Kyoto University made a remarkable observation. He noticed that cells that were lysogenic for ε^{15} (cells dormantly infected by ε^{15} and therefore containing its genetic material) would not adsorb new ε^{15} particles; lysogenic infection by the phage apparently modified the bacterium's sur-



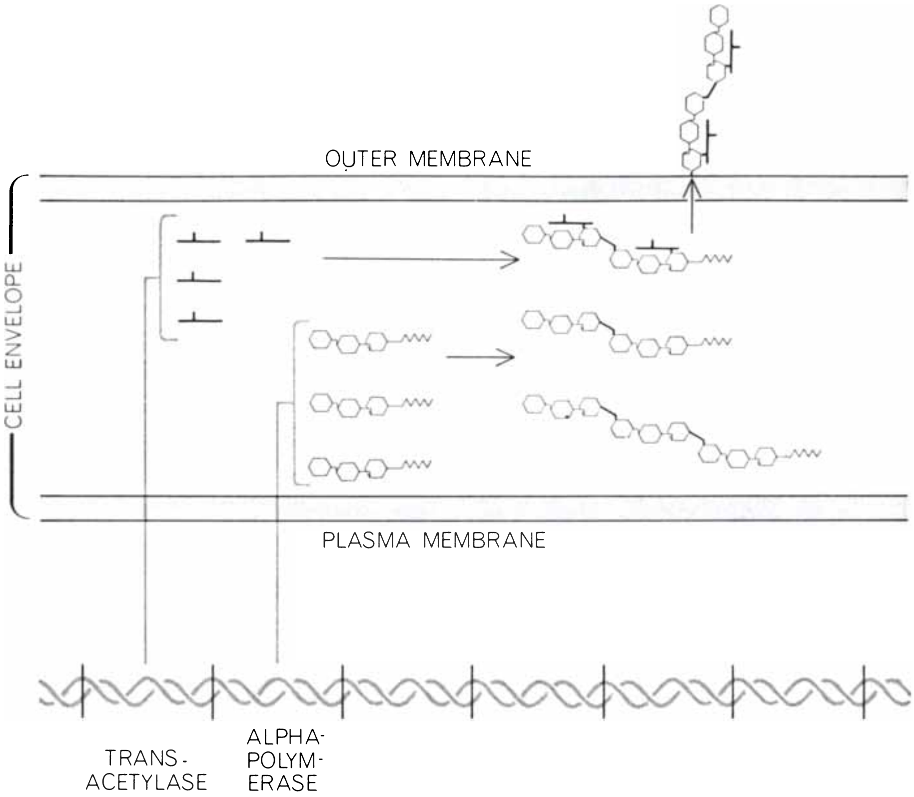
O-ANTIGEN, the component of *S. anatum* lipopolysaccharide to which ε^{15} attaches itself, is a repeating chain of three sugars: mannose, rhamnose and galactose. In O-antigen from uninfected cells the galactose units carry an acetyl group (CH_3CO) and trisaccharide (three-sugar) units are joined by alpha linkages, in which the linkage to the mannose is below the plane of galactose (*top*). In O-antigen from cells that are lysogenic for ε^{15} the acetyl is missing and the linkage is beta; the linkage is above the galactose (*bottom*).

face and made subsequent adsorption impossible. Six years later one of us (Robbins) and Takahiro Uchida were able to learn the chemical basis for this observation: O-antigen that is synthesized after infection by ε^{15} has a different structure from normal O-antigen. The polysaccharide chains lack the acetyl substituent. And the galactose-mannose linkages are beta linkages rather than alpha; they can be cleaved experimentally by the enzyme beta-galactosidase and not by alpha-galactosidase. In other words, lysogeny by ε^{15} changes the chemical structure of its own receptor site.

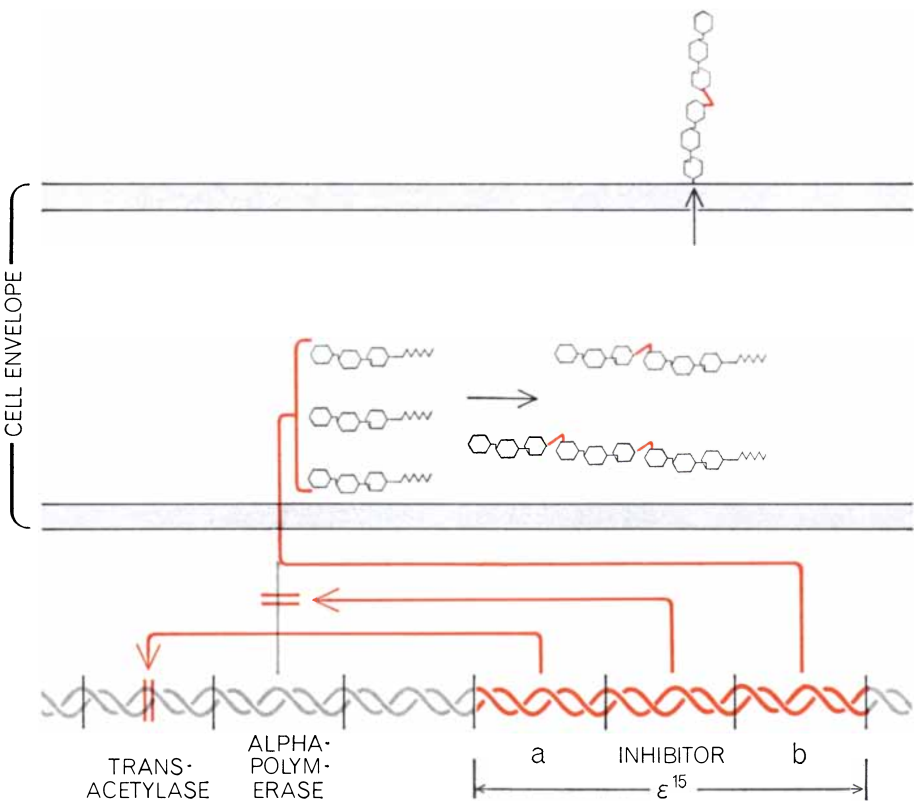
With John Keller and Dennis Bray, we undertook genetic and biochemical experiments in an effort to determine the events that accomplish this change in structure. We had reason to believe these events were mediated by specific phage genes, since Uetake and Uchida had in 1959 isolated ε^{15} mutants that do not alter receptor-site structure in the usual way. We were able to track each change in the receptor to a specific phage gene incorporated in the bacterial chromosome. One gene, called gene *a*, blocks the synthesis of *S. anatum* transacetylase and thus eliminates the acetyl groups on O-antigen polymer that is synthesized after infection. Another gene, *b*, makes a new enzyme, beta-polymerase, that replaces the *S. anatum* alpha-polymerase; like the normal enzyme, the new enzyme links up trisaccharide units, but it does so by generating beta linkages, in which the linkage to mannose is above the plane of the galactose instead of below it. Finally, another ε^{15} gene inactivates the host's alpha-polymerase by making a protein that inhibits it [*see bottom illustration at right*]. The protein has been purified in our laboratory, and it inhibits alpha-polymerase in a cell-free system.

It is interesting to note that two different mechanisms of inactivation are employed here. The transacetylase is repressed, that is, its synthesis is blocked; this is shown by the fact that after infection there is no increase in transacetylase activity but enzyme that was previously synthesized continues to show activity. The alpha-polymerase, on the other hand, continues to be synthesized by the cell but is inactivated by an inhibitor; this is shown by the fact that after infection its activity decreases or ceases altogether.

Since ε^{15} alters the structure of O-antigen both by preventing the attachment of acetyl groups and by changing the linkage, the question arises: Which of these two structural features is essential



BIOSYNTHESIS of O-antigen takes place in the cell envelope. Genes in the bacterial chromosome (*gray helix*) encode instructions for two enzymes involved in the synthesis. One, alpha-polymerase, links individual trisaccharides on lipid carriers into long polysaccharides, forming the alpha linkage. The other, transacetylase, attaches acetyl groups (*black bars*) to galactose units. Then the O-antigen chains are transferred to the cell surface.



MODIFICATION of O-antigen in cells lysogenic for ε^{15} is directed by three phage genes (*colored helix*) integrated into the bacterial chromosome. Host-cell alpha-polymerase is inactivated by a protein made by the inhibitor gene. Gene *a* is responsible for repressing the synthesis of host transacetylase. Gene *b* makes beta-polymerase, which generates an O-antigen with beta linkages that replaces the normal alpha-linked O-antigen on the cell surface.

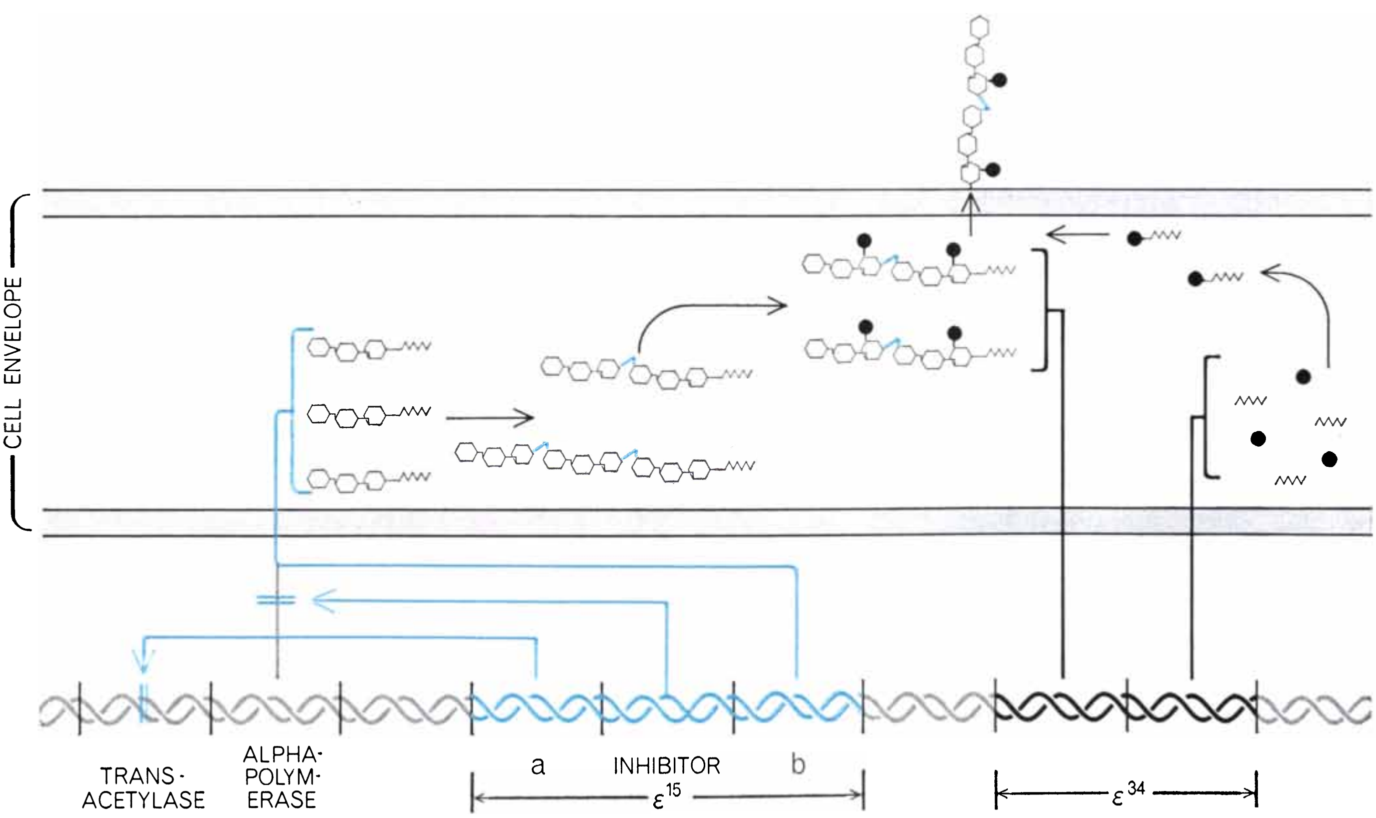
for adsorption—or are both of them essential? In order to answer this question Bray isolated mutants of S. *anatum* that synthesize an alpha-linked polymer without acetyl groups. He found that ε^{15} particles are adsorbed by these transacetylase mutants, and so the acetyl groups must not be prerequisites for attachment. The next step was to test an O-antigen that retains acetyl groups but has the beta linkage. S. *anatum* was made to synthesize such a polymer by infecting it with a mutant phage, $\varepsilon^{15}a$, that changes the linkage without removing the acetyl groups. Bacteria so lysogenized fail to adsorb ε^{15}, showing that the alpha linkage is the critical feature of the receptor site.

We soon found that alteration of the receptor site is not the only consequence of lysogeny by ε^{15}. The very changes that prevent the attachment of more ε^{15} particles make possible the attachment of a different virus, ε^{34}. In fact, ε^{34} will infect only cells that are lysogenic for ε^{15}; it requires a receptor site without acetyl groups and with beta linkages. Thus one phage makes it possible for another phage to infect the same bacterial cell.

Phage ε^{34}, in turn, brings about a further modification of the cell surface. It adds glucose to the galactose units of the O-antigen. Wright determined the steps that occur in this modification [*see illustration below*]. By working with various mutants of ε^{34} he found that one phage gene is responsible for linking glucose to a carrier molecule, a lipid similar to the one involved in the polymerization of O-antigen. Another gene is responsible for transferring glucose from the carrier to the polysaccharide chain. The presence of the glucose makes it impossible, in turn, for new ε^{34} particles to be adsorbed by cells that have been lysogenized by ε^{34}.

There may be significant similarities between the way ε^{15} or ε^{34} alters the surface of bacterial cells and the action of tumor viruses on animal cells. Tumor viruses such as simian virus 40 and the polyoma virus transform normal cells into cancer cells. Renato Dulbecco and his colleagues at the Salk Institute for Biological Studies have shown that the transformation, the process of "cancerization," is induced by viral genetic material [see "The Induction of Cancer by Viruses," by Renato Dulbecco; SCIENTIFIC AMERICAN, April, 1967]. Dulbecco's group has recently reported evidence that the chromosome of tumor viruses sometimes persists in the cancerized cell in a dormant state, integrated into the host cell's chromosome. This dormant state may be analogous to the lysogenic state of such phages as ε^{15} and ε^{34}. Moreover, cancerized cells appear to have altered surface characteristics. A new, virus-specific antigen that can be detected by immunological techniques usually appears on the cancer cell's surface. Contact inhibition, the process whereby the persistent undulating motion of normal cells is stopped when one cell touches another, is ineffective in a cancer-cell culture; the cells keep on moving even after touching, suggesting again that their surface has somehow been altered. It may be that a specific gene of the tumor virus is responsible for each change in the cell surface just as specific ε^{15} genes bring about changes in the surface of S. *anatum*.

The analogy is nice, but it should be treated with caution. Normal cells can be transformed into cancer cells by chemical agents or radiation without any apparent viral influence at all, implying that cancerization does not always require specific virus genes. Moreover, even when a virus is the agent, it may be that cancerization is incidental to the entire process of infection and integration into the host chromosome and does not involve specific genes. As work on phage infection and on cancerization proceeds, it will be interesting and perhaps helpful to watch for further similarities and differences between the mechanism by which a phage alters the surface of its bacterial host and the more complex mechanisms by which a tumor virus affects animal cells.



FURTHER MODIFICATION of O-antigen occurs when a second phage, ε^{34}, infects cells that are already lysogenic for ε^{15}. One gene of the new phage's genetic material (*black helix*) is responsible for the attachment of glucose (*black disks*) to a lipid carrier. A second phage gene transfers the glucose to the galactose units of the beta-linked O-antigen chains, which are then transferred to the surface.

Dropped.
Fried.
Dried.
Frozen.
And fought over.

How do you build an air conditioner that can take a brutal beating, and still come up working?

That was the Carrier Special Products Group's assignment from Litton Industries when work started on the Marine Tactical Data System.

New huts for the Marines

MTDS is a highly sophisticated control system created to detect enemy aircraft and direct Marine firepower. Both with computerized speed and accuracy.

Communication gear. Display devices. Computers. All are housed in a series of huts that are rugged. Light. Compact. Made to ride into battle under a helicopter. In a plane. Or on a truck. Ready to perform in all climates. Without fail.

And each hut had to have air conditioning.

Cool it in the heat of battle

MTDS was unprecedented in scope and concept. And the standards set for its air conditioning were equally unique.

First, the air conditioner had to control the environment for the electronic equipment (as well as personnel efficiency). And in outside temperatures ranging from 65°F below zero to 125°F above. An integrated system to ventilate, heat, cool and filter air.

Next, it had to meet the restricting dimensional and physically torturous requirements MTDS demanded.

So with a bundle of military specs, our Special Products Group got down to work and came up with a one-of-a-kind air conditioner. A 2x4-foot traveling unit weighing 206 lbs. Secured inside the MTDS hut for travel, and bolted to the outside for operation.

Big volume in a little space

Probably one of the Group's toughest jobs was getting the large capacity called for in such a small space unit. But it was done, and even exceeded the specified capacity of 17,000 Btuh.

Completed, the Carrier air conditioners were put into their huts and drop-tested by the Marines in helicopters. Jounced on trucks over 4x8's at speeds up to 25 mph. Frozen in a freezer-like climate. Broiled in an oven-like atmosphere. And then (using specified preventive maintenance) run for 2500 hours with no major maintenance, 14,000 hours without overhaul.

Special delivery

The first major MTDS sub-system was delivered to the Marines in 1966. Today, there are more than 500 in action.

In the military. In aerospace. Oceanography. Transportation. Food storage. In commerce and industry. In every one, our Special Products Group is at your service.

Just call or write: Special Products Group, Carrier Air Conditioning Company, Syracuse, New York 13201.

Carrier®

HOW BIRDS SING

The mechanism of bird song has traditionally been compared to either a wind musical instrument or the human vocal apparatus. The analysis of the bird songs themselves points toward an entirely different system

by Crawford H. Greenewalt

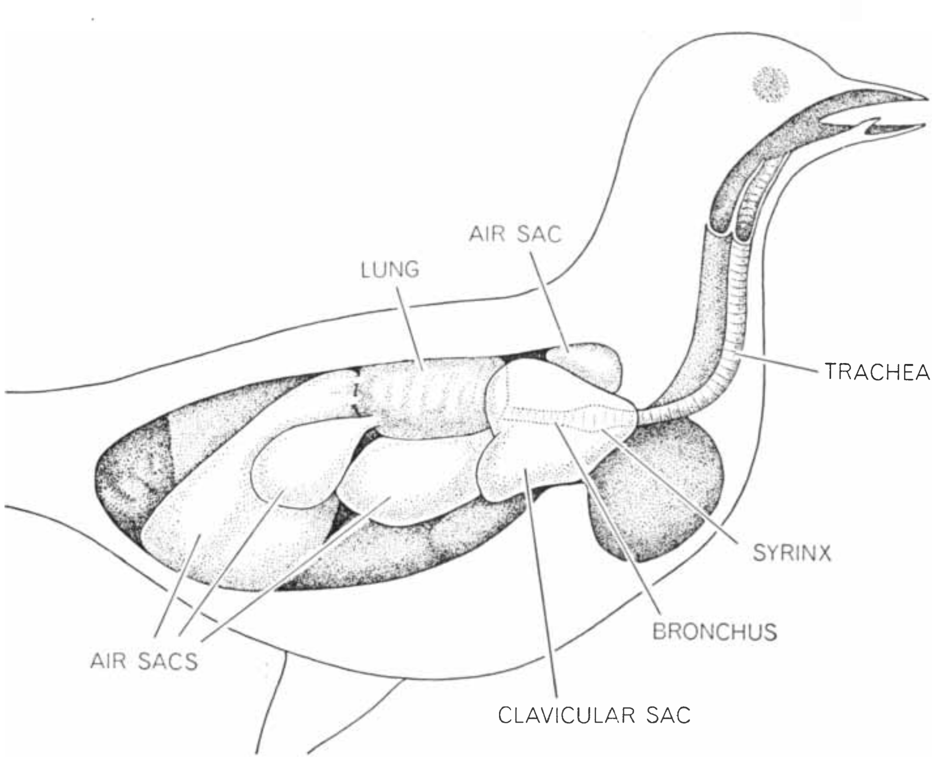
Some years ago I read a book in which, among other things, the author summarized the theories that had been advanced to explain the physiological processes employed by a singing bird. These theories, it seemed to me, were totally unacceptable, and I determined, perhaps with more rashness than wisdom, to see if I could find an explanation that did no violence to either the anatomical findings or the laws of physics.

The theories to which I took exception compared the singing bird either to a musical instrument—the clarinet, the oboe or the trumpet—or alternatively to the human voice (that is, a bird was believed to employ the same devices we do when we speak or sing). In a clarinet, for example, pitch and timbre are controlled by the effective length of the barrel of the instrument; the vibrating reed is in effect driven by and so forced to conform to the harmonic spectrum of the resonator. If a bird sang in this fashion, pitch could be varied only by extension and retraction of the neck. For a song sparrow, with its two-octave range, the neck would have to be extended by a factor of four—clearly a physical impossibility. Furthermore, birds sing many phrases in which the wave form is sinusoidal (a train of pure sine waves), without an associated harmonic spectrum. Resonators such as the barrel of a clarinet must by definition produce harmonics and cannot generate a sinusoidal wave form.

As for the human-voice theory, when we speak we produce at the glottis a series of puffs of air mathematically equivalent to a harmonic spectrum with an infinity of components, and with a fundamental frequency corresponding to the interval between puffs. The resulting acoustical disturbance is then modulated in our oral passages to produce the spoken word. This mechanism does not (and cannot) produce a purely sinusoidal wave form; every speech sound is associated with a harmonic spectrum of considerable complexity.

Negatives such as these are neither satisfying nor sporting, and to remedy both potential criticisms I undertook to develop a physiological and acoustical model that would describe a singing bird. The approach was to study in detail the bird songs themselves, and to develop from their constituent parts some notion of the associated acoustical processes. It seemed useless to employ an anatomical approach because the many excellent anatomical studies in the scientific literature have produced no useful conclusions, and because the acoustical analysis of human speech has been so strikingly successful in elucidating our own vocal performance.

I might pause here to note that I have written a book—*Bird Song: Acoustics and Physiology* (Smithsonian Institution Press, 1968). In it much detail is given on the state of the literature, the instrumentation my associates and I have developed for the analysis of bird song, and



VOCAL APPARATUS OF A SONGBIRD is shown in schematic form. The vocal organ itself is the syrinx, which is located at the point where two bronchi join to form the trachea.

the proofs, mathematical and otherwise, for the findings I am about to describe. This article is in effect a summary of that book, and if the reader should wish more detail he will find it there.

The vocal organ of birds is called the syrinx. It is located deep in the thoracic cavity, at the point where the two bronchi join to form the trachea [*see illustrations below*]. It is difficult to decide from the detailed anatomy of the syrinx exactly what parts are important in vocalization, and precisely how sound is produced. At one time or another almost every anatomical element in the syrinx has been assigned a role in vocalization; unhappily there is an overabundance of possibilities.

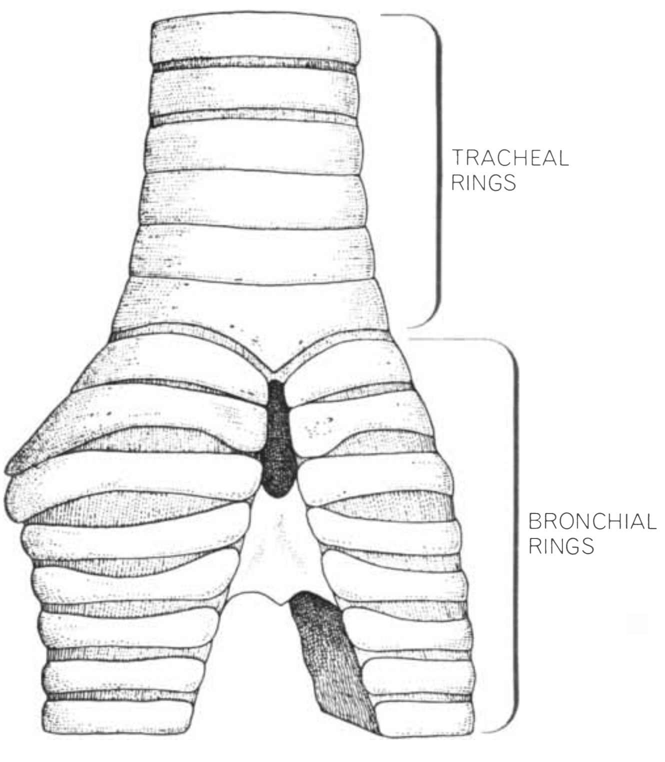
I have proposed the relatively simple functional structure depicted in the top illustration on page 132. The elements are the tympanic membranes and their associated musculature, the external labia and the system of internal air sacs (not shown in the illustration) that force air through the bronchial passage and bulge the membrane into the bronchial lumen, or passageway. The illustration depicts the highly evolved syrinx of a songbird. At successively earlier stages of avian evolution the syrinx loses the external labia, then the syringeal muscles and ends up as a simple tube with a membrane on its periphery and contained within an air sac.

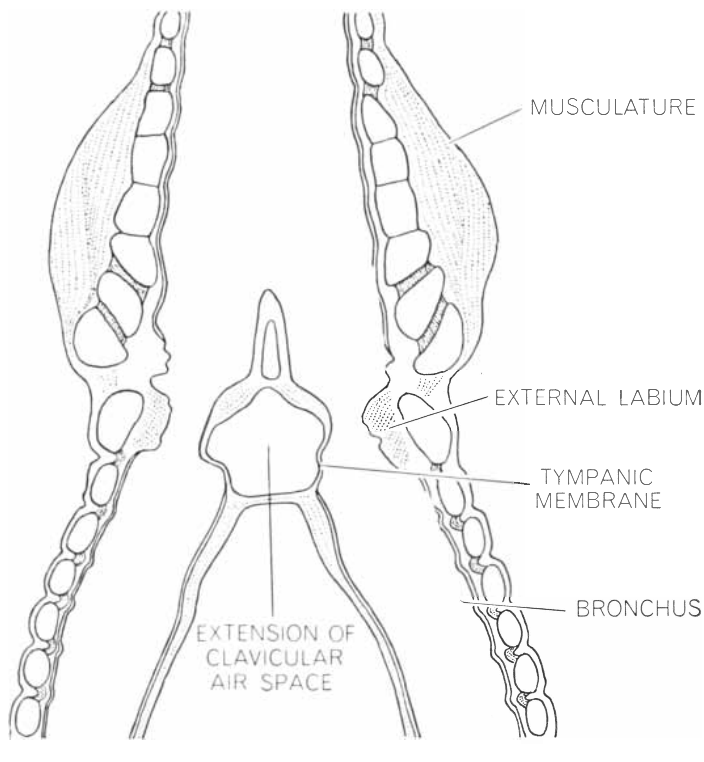
It should be kept in mind that the functional elements of the syrinx are doubled, that is, there is a set for each bronchus. Since each bronchial passage has its own membrane, musculature and nervous system, it is evident that birds can control each passage independently of the other; they can sing what might be called an internal duet.

The system operates as follows: When a bird undertakes to sing, it in effect closes a valve between the lung and the syrinx. Then it compresses (with its chest muscles) the air in a system of sacs. Pressure in the clavicular sac, which surrounds the syrinx, forces the exceedingly thin tympanic membrane into the bronchial passage, closing it momentarily. Tension is then applied to the syringeal muscles, which, acting in opposition to the sac pressure, withdraw the bulged membrane from the opposite bronchial wall, thus creating a passage through the bronchial tube. Air streaming through the passage past the tensed membrane stimulates it to vibrate, and song is produced. If only one of the two voices is to be used, no tension is applied to the other membrane, and the corresponding bronchial passage remains closed. When a duet is sung, both membranes are under tension; there are two airstreams and hence two vibrating membranes and two simultaneous sounds.

The illustrations on the next two pages show three complex songs analyzed in terms of amplitude and frequency. The songs embrace almost every vocal gymnastic of which birds are capable. Note, for example, the extremely rapid amplitude modulations, the wide range in the shape and extent of the amplitude envelopes and the relatively large frequency intervals: from just over two to just under seven kilocycles per second. The amplitude displays are precise and require no qualification. In the frequency displays there are a time delay and an integration of frequencies over small time intervals that introduce ambiguities for precise analysis.

The question of whether or not a bird can use its two acoustical sources independently is vital to the elucidation of the mechanics of bird song. If, for instance, the bird's trachea behaved like the barrel of a clarinet, both sources, that is, both vibrating membranes, would be forced to conform to the resonances of the trachea. The two sources would then merely reinforce each other, increasing the amplitude, or loudness, of the sound. If, on the other hand, the bird's trachea behaved like human oral cavities, and the vibrating membranes were analo-



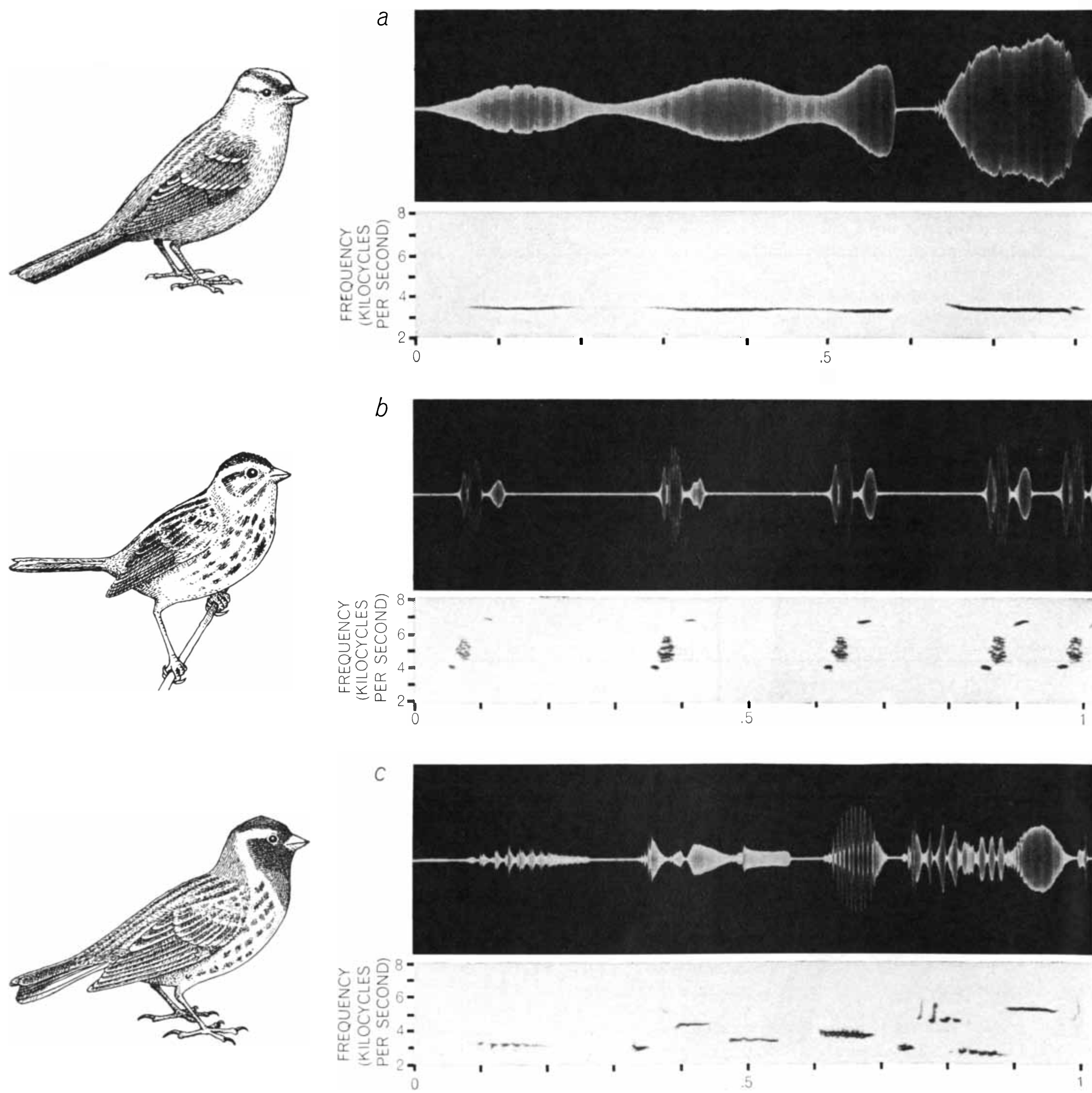


ANATOMY OF THE SYRINX sheds little light on its function. At left is the skeleton of the syrinx in the magpie; at right, a section through the syrinx of the European blackbird. Both drawings are based on studies published by the anatomist V. Häcker in 1900.

gous to the human glottis, the resulting wave form should be equivalent to a harmonic spectrum of great complexity involving three fundamental frequencies, one for each source and one for the first resonance of the trachea.

The fact that the two sources can produce two harmonically unrelated sinusoidal wave forms was first shown by Ralph K. Potter, George A. Kopp and H. C. Green of the Bell Telephone Laboratories, for a fragment of the song of a brown thrasher. Since the phenomenon is an important one, I have searched for and found many similar examples in species representative of most families of birds. One example is provided by a relatively long phrase from the song of a mockingbird [*see illustration on page 130*]. In the displays the two voices were separated with sharply discriminating acoustical filters; frequency was determined by measuring time intervals for 10 successive sine waves. The two voices are harmonically unrelated and overlap on the time axis. We must conclude that neither the musical-instrument theory nor the human-voice theory is operative.

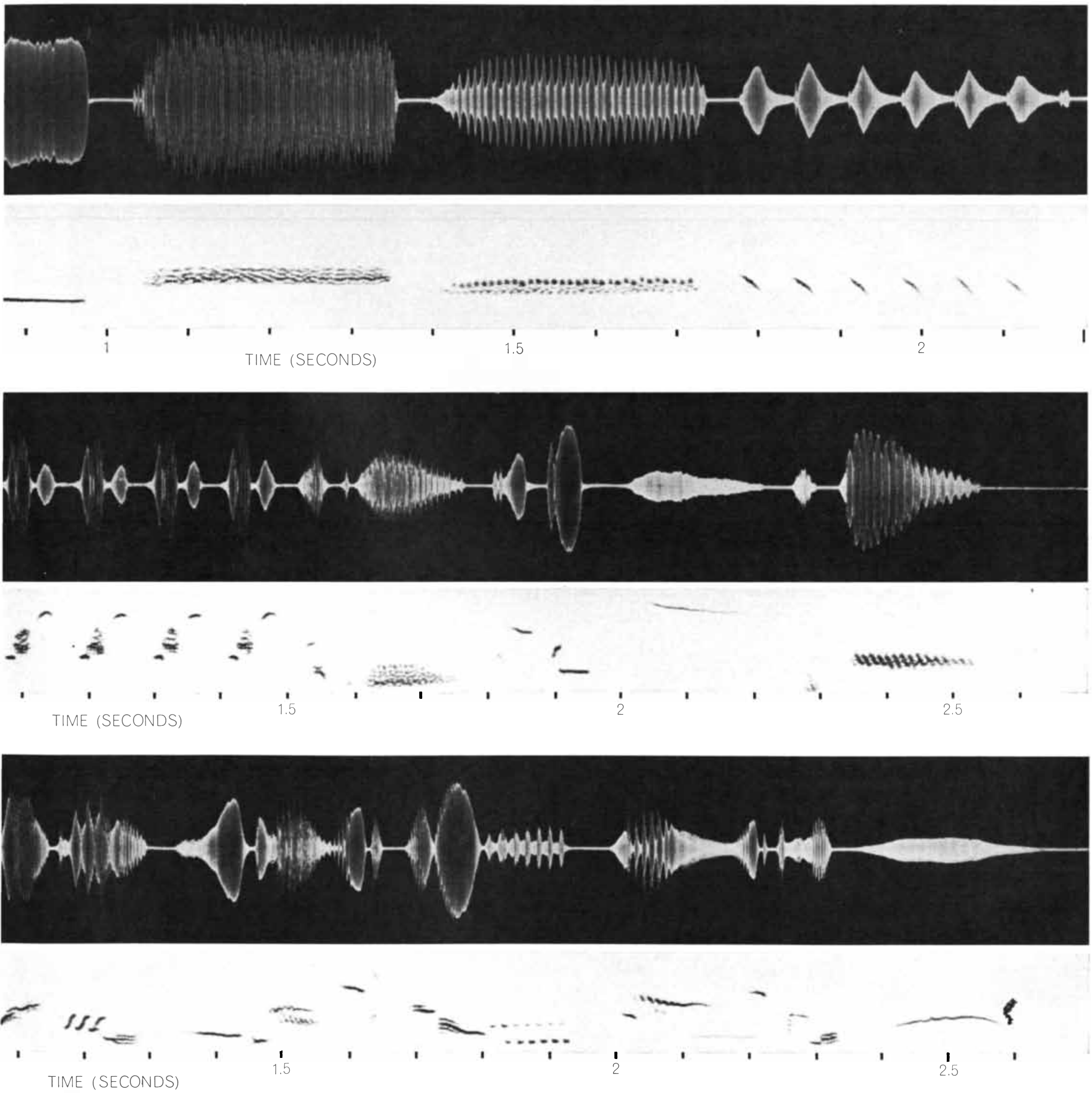
Other evidence may be found in phrases that cover a large frequency range. If such a "glissando" traverses a tracheal resonance or antiresonance, there should be a marked increase or decrease in amplitude at the appropriate frequencies. An example is a glissando



SONGS OF THREE SONGBIRDS are presented in traces that show amplitude and frequency as a function of time. The top record for each bird is an oscillogram that displays amplitude; the bottom record is a sound spectrogram that displays frequency. At *a* is the

sung by a yellow warbler that embraces a full octave [*see illustration on page 131*]. There are changes in amplitude in the glissando, but in an acoustical sense they are very small (less than three decibels), and it is most unlikely that they are associated with tracheal resonances. The two minima at 90 and 130 milliseconds appear at frequencies of 4.8 and 6.5 kilocycles per second. One of these could correspond to a resonance; both of them could not.

Although the evidence seems quite conclusive that the primary vibrations produced in the syrinx traverse the trachea without attenuation, amplification or change of any sort, we must offer a valid explanation for this rather surprising acoustical inertness of the trachea. The trachea is a tube. A tube, if its acoustical losses are small, exhibits resonances, and one would expect it to be an effective modulator, even with its relatively soft walls. A possible explanation—indeed, the only likely explanation—is that the impedance of the source (the vibrating membrane in its constricted passage) closely matches the impedance of the trachea. Calculations based on reasonable assumptions indicate that this condition will obtain where the mean cross-sectional area of the trachea is about 10 times that of the syrinx. The curve relating this area ratio to attenuation is fairly flat and shows that the ratio can be substantially higher or lower than



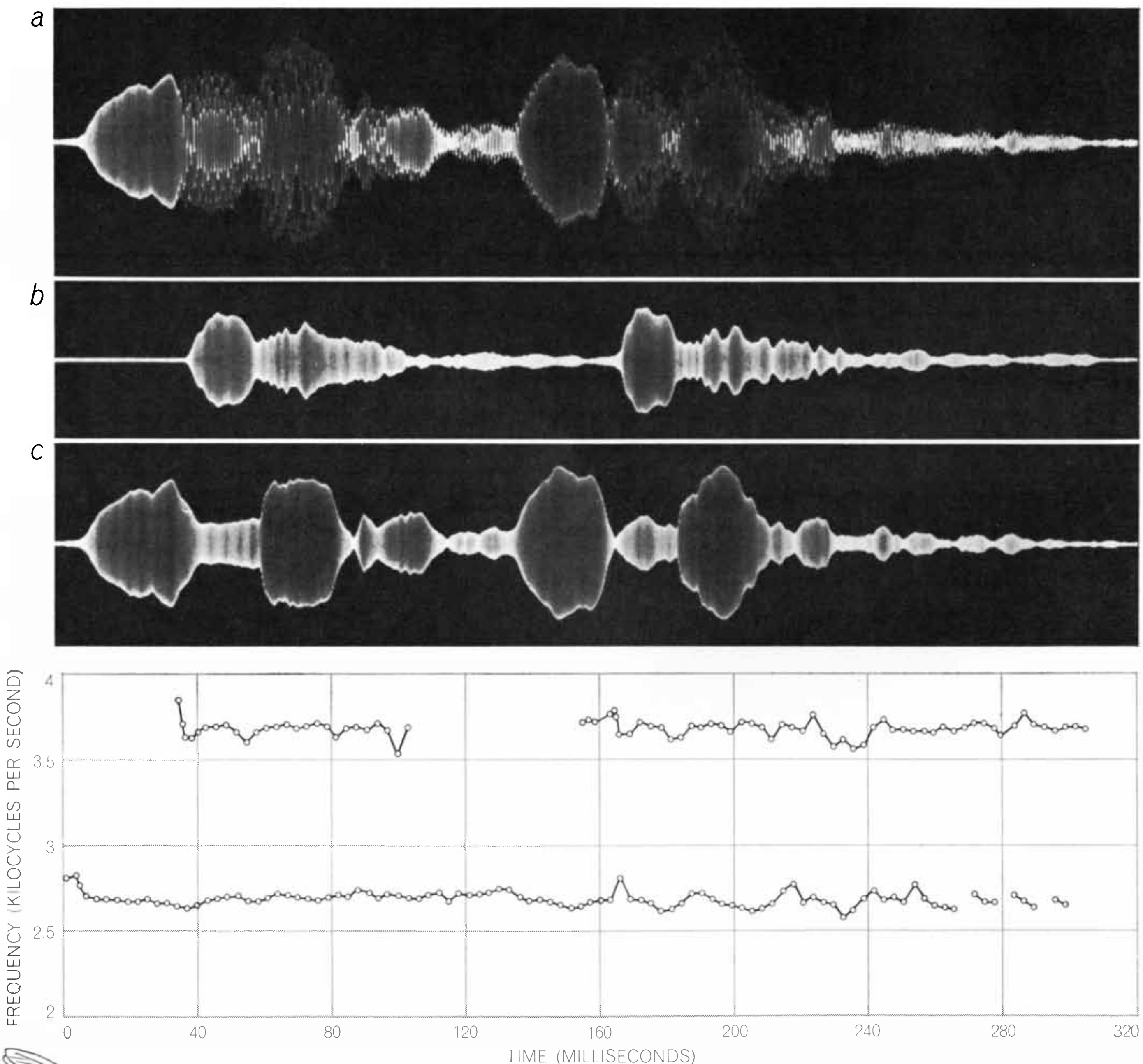
song of a white-crowned sparrow; at *b*, the song of a song sparrow; at *c*, the song of a Lapland longspur. These song recordings illustrate the extraordinary virtuosity of these singing birds. No human performance can match the complexity of the individual phrases.

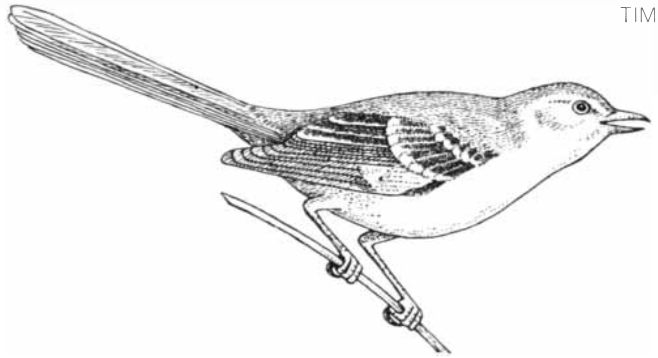
10 before tracheal attenuation becomes important.

There is another way of expressing the effect of variation in the relative cross sections of source and resonator that may be easier to comprehend. If we take a tube closed at one end and open at the other, the resonances will occur at multiples 1, 3, 5, 7 and so on of the fundamental frequency. If the tube is open at both ends, the resonances will occur at multiples 0, 2, 4, 6 and so on of the fundamental frequency. If now we begin with a tube open at both ends and gradually close off one end, we will in due course arrive at a twilight zone within which the tube shows no resonant effects. This presumably is the zone in which birds sing.

Perhaps this is another example of Nature's ingenuity in dealing with the needs of her diverse creatures. The necks of birds (which contain and limit the trachea) have many functions, for instance feeding, preening and nest-building. It would be odd if such functions operated to restrict freedom in an activity as important as song.

So far I have undertaken to show that bird song has its origin in the syrinx,





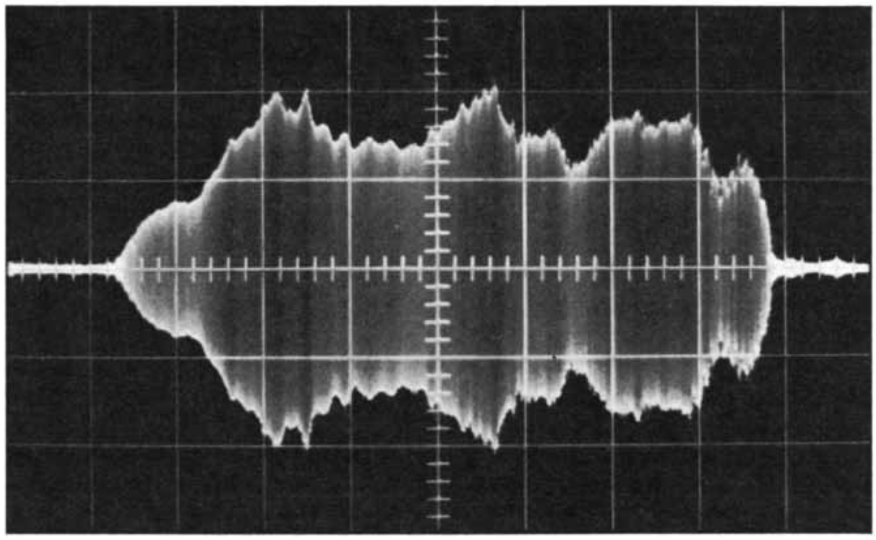
"INTERNAL DUET" is revealed when a phrase of the song of a mockingbird (*a*) is dissected into two phrases (*b, c*) with appropriate filters. The plots at bottom show the frequencies in *b* and *c*.

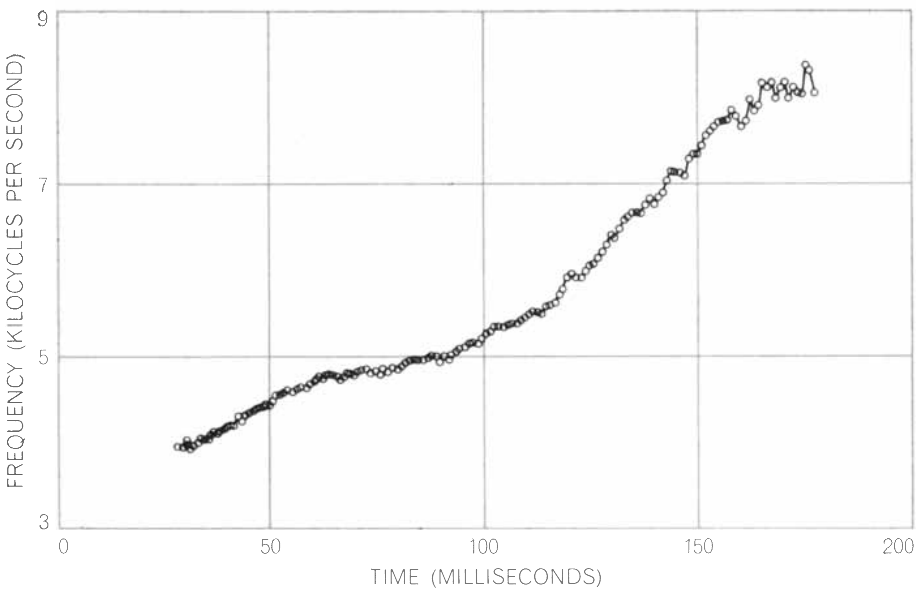
that the syrinx contains two acoustical sources which can be independently controlled to produce an internal duet, and that the sounds which originate in the syrinx pass through the trachea to the ear of the listener without further modulation. I must now show precisely how the sources in the syrinx operate and how the extraordinarily rapid and complex modulations so common in bird song are produced.

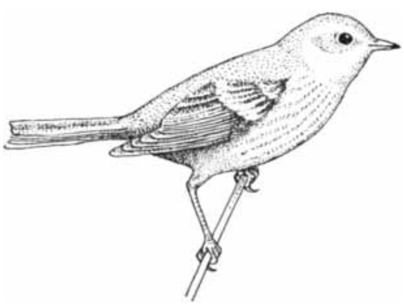
I have noted that pressure in the clavicular sac forces the internal tympanic membranes into the bronchial lumen against tension applied to the membrane by the muscles of the syrinx. Air streaming past the resulting constriction in the bronchial passage stimulates the membrane to vibrate. If this postulate is correct, it follows that in any rapid modulation amplitude and frequency must be coupled; an increase in frequency can be produced only by increasing the tension in the syringeal muscles. At a given pressure in the clavicular sac this tension will increase the cross section of the bronchial passage, which, as we shall see, must produce either an increase or a decrease in the acoustical amplitude.

The direct coupling of frequency and amplitude, that is, amplitude increasing with increasing frequency, appears in curves for Townsend's solitaire in which frequency and amplitude are simultaneously plotted against time [*see upper illustration on page 133*]. The reverse process, with amplitude falling as frequency rises, appears in similar curves for the red-winged blackbird [*see lower illustration on page 133*]. In curves for the song sparrow one can see both types of coupling within the modulating period: amplitude rises with increasing frequency up to 6.8 kilocycles and falls above that frequency.

These phenomena can best be explained by referring again to the simplified model of the syrinx shown on the next page. If we start with just sufficient tension (T) on the tympanic membranes to balance the pressure P, the distance across the passage (D) will be zero and no air will flow. A small increase in tension will open the passage and allow the membrane to vibrate at a frequency corresponding to the tension, but the amplitude of vibration will be restricted to low values by the small distance across the passage. As the tension increases, frequency and amplitude will increase together, as a larger distance across the passage permits greater vibrational amplitude in the membranes. With a further increase in membrane tension, and a correspondingly larger distance across the passage, we arrive at a point where the airflow through the syringeal constriction can no longer stimulate the membrane to vibrate through the allowable distance D, and the amplitude will decrease. At the limit—a tension sufficiently large to open the passage fully—the amplitude will be zero, and the airflow from the bird's lungs will escape through the trachea without sound production. It should be understood that airflow is used here in an acoustical sense; it is a time-varying flow that is controlled by the vibration of the tympanic membrane. Air could well leak past the syringeal constriction, but this airflow would be continuous and would produce no sound.

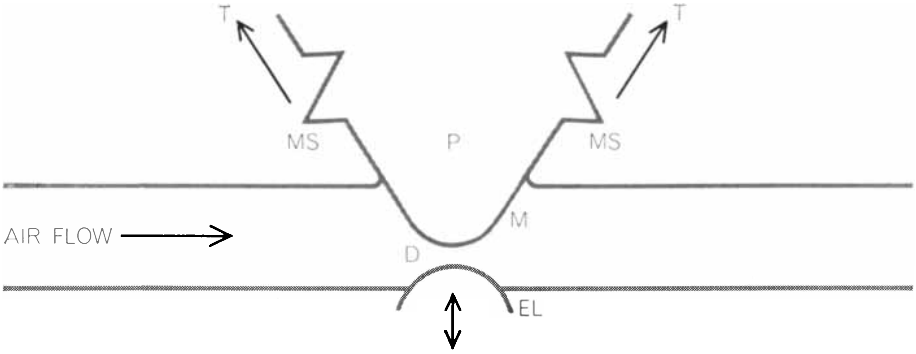




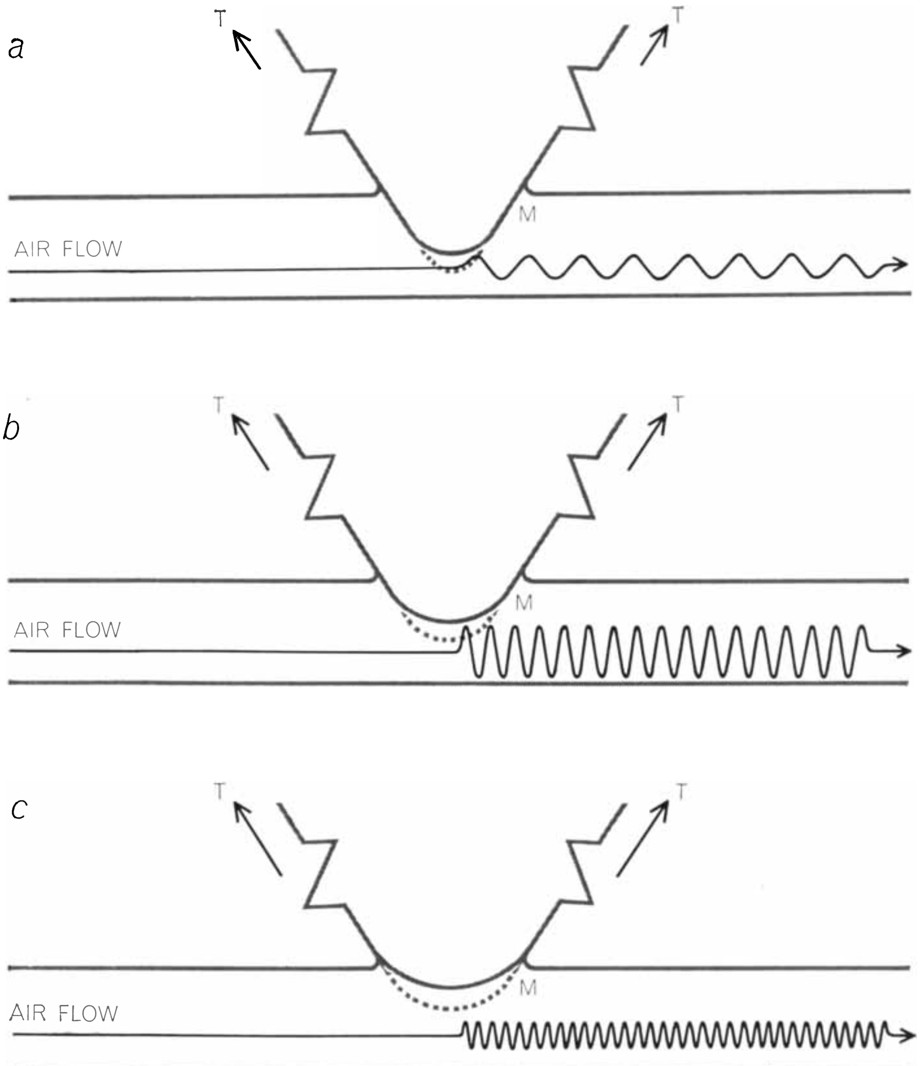


IN THE GLISSANDO of the yellow warbler frequency rises smoothly (*plot at bottom*) but amplitude (*top*) rises only moderately.

Amplitude-frequency coupling always accompanies rapid modulation, but it is not necessarily present when the period of the modulation is long. Consider a song sparrow phrase about 50 milliseconds long [*see top illustration on page 134*]. The modulating frequency remains constant at about 300 cycles per second, as does the frequency excursion. There is, however, a gradual rise and fall in amplitude from the beginning of the phrase to the end. Such comparatively long-term changes in amplitude without



SIMPLIFIED MODEL OF THE SYRINX indicates how it functions. Air flows from the lungs at left into the trachea at right. The pressure of air in the clavicular sac (*P*) forces the tympanic membrane (*M*) into the bronchial lumen, or passage. Tension (*T*) is produced in the membrane by the syringeal muscles (*MS*). The resultant of the two forces *P* and *T* determines the distance across the lumen (*D*) and also the vibrating frequency. *EL* is external labium, which can regulate distance across bronchial lumen without affecting frequency.



AMPLITUDE AND FREQUENCY ARE COUPLED when changes in either are rapid. In *a* a tone of a given amplitude is produced when the airflow causes the tensed tympanic membrane (*M*) to vibrate. In *b* the tension (*T*) is increased, causing the membrane to vibrate faster. Increased tension also retracts membrane, further opening the passage. The membrane can therefore vibrate through a larger distance, increasing amplitude of song. This is termed positive coupling. In *c* tension on the membrane has increased beyond a threshold; the tension is so great that the membrane can no longer vibrate across the entire passage. Beyond this threshold amplitude decreases with frequency, that is, the coupling is negative.

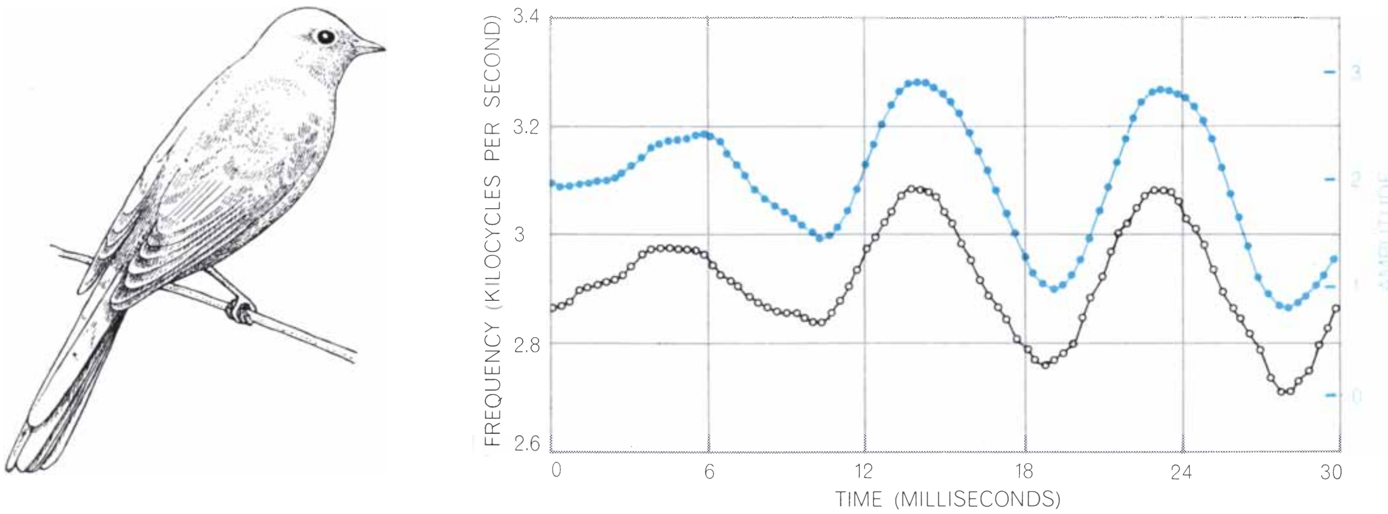
an accompanying change in frequency are relatively common. We find the necessary control mechanism in the external labium [*"EL" in the bottom illustration on page 134*], whose movement into and out of the bronchial lumen provides a simple device for changing amplitude without affecting muscle tension and hence without changing the frequencies produced in the tympanic membrane.

I must confess that this role for the external labium rests on a somewhat shaky anatomical base. Many anatomists report the presence of this member, particularly in the songbird syrinx; it is a pillowlike structure erected on the bronchial half-ring directly opposite the tympanic membrane. Only one anatomist, however, describes muscular attachments that could move the labium into and out of the bronchial lumen. I trust that some interested anatomist will note this opportunity for a definitive publication.

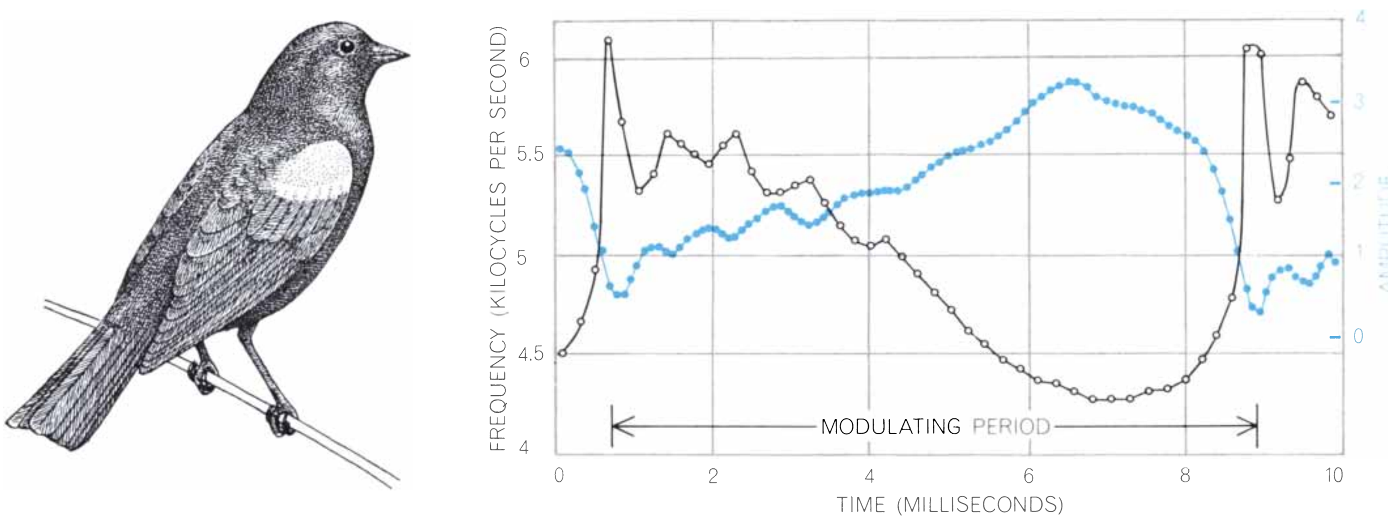
The bird could, of course, change the air pressure in its sac system, but this would have two effects whose resultant is at best ambiguous. Increasing the sac pressure would drive the membrane farther into the bronchial lumen, thereby decreasing its cross section; at the same time it would increase muscle tension. The increased pressure drop across the constricted bronchial passage would increase flow and hence would compensate for the reduction in cross section. I am not rash enough to predict the resultant of these phenomena, but it seems clear that they could not produce the overall rise and fall in amplitude exhibited in the song sparrow phrase on page 134. For that we need the external labium or some other anatomical feature that affects amplitude alone.

In any event, these rapid and highly variable modulations are so common as to be a prime characteristic in the songs of most birds. They can be repetitive, that is, there is a modulating frequency that continues for 50 to 500 milliseconds, or they can be nonrepetitive, that is, they may vary rapidly and randomly in amplitude, in frequency or in both.

Many phrases sung by the song sparrow comprise modulations that are repetitive. The modulating frequency averages 300 cycles per second, and amplitude ranges much more widely within a modulating period than frequency. In the songs of other species the modulating frequency ranges from about 100 cycles per second to something over 400. We are dealing here with two oscillating systems superimposed on each other. The first is the tympanic membrane vibrating



POSITIVE COUPLING of amplitude and frequency is shown in this analysis of a phrase sung by Townsend's solitaire. The colored curve traces the amplitude of the phrase; the black curve traces the frequency. The curves rise and fall together over the time period.



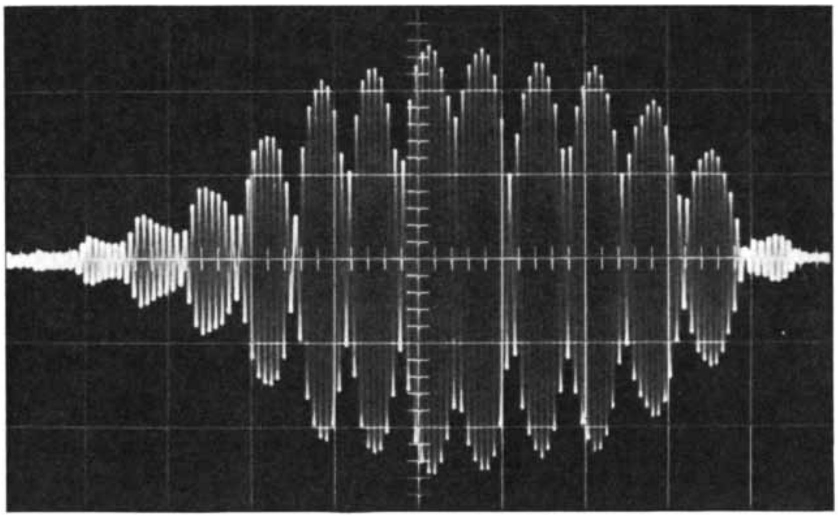
NEGATIVE COUPLING is represented in this analysis of a phrase sung by a red-winged blackbird. Here when the amplitude (*colored curve*) decreases, the frequency (*black*) increases. Curves also show the characteristic modulating period of this phrase in the song.

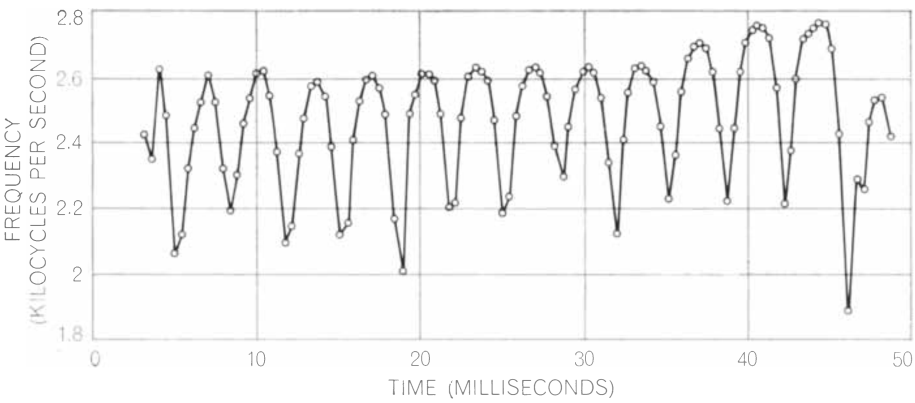
at its natural, or "carrier," frequency. The second includes the syringeal muscles, whose vibration would produce a periodic change in tension and hence a change in both frequency and amplitude. The mass and elasticity of the muscle system presumably limits its natural frequency of vibration to relatively low values—up to, say, 500 cycles per second. The much lower mass of the tympanic membrane allows vibration up to 10 kilocycles per second.

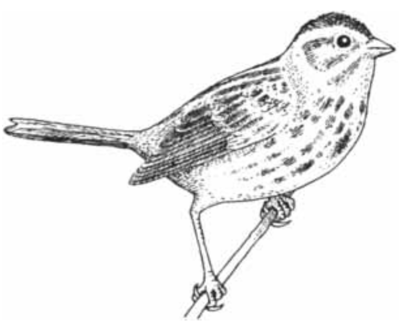
Nonrepetitive modulations can be seen in the courtship song of the brown-headed cowbird [*see illustration on page 136*]. This bird of unprepossessing appearance and habits is the undisputed winner in the decathlon of avian vocalization. Roger Tory Peterson characterizes the first phrase in the illustration as "glug" and the second phrase as "gleeee." Consider the following features: The frequency range in the two phrases is, by a large margin, wider than it is in any other bird song. It extends from .75 to 10.7 kilocycles per second—nearly four octaves! The maximum frequency at 10.7 kilocycles per second is higher than what we have found for any other bird, just nosing out the 10.5 kilocycles per second at the top of the blackpoll-warbler song. Both voices are used in the second subphrase of the "glug," and the frequency spread between the two voices (two full octaves) is exceeded only by that of the American bittern.

The first note in the "gleeee" is the shortest I have encountered. It lasts a bit less than two milliseconds and comprises a packet of 12 sine waves at 6.4 kilocycles per second. The glissando at 50 milliseconds in the "gleeee" is one of the most rapid, covering the range from five to eight kilocycles per second in four milliseconds and 23 sine waves, an average of 130 cycles per wave. The modulating frequency of the high voice in the second subphrase of the "glug" is about 700 cycles per second, higher by a large margin than any other. These performances are truly remarkable. What purpose is served by a "glug" comprising five widely different subphrases, together with a "gleeee" including a note of negligible duration, two rapid glissandi and a peak frequency of 10.7 kilocycles per second, only Madame Cowbird will know.

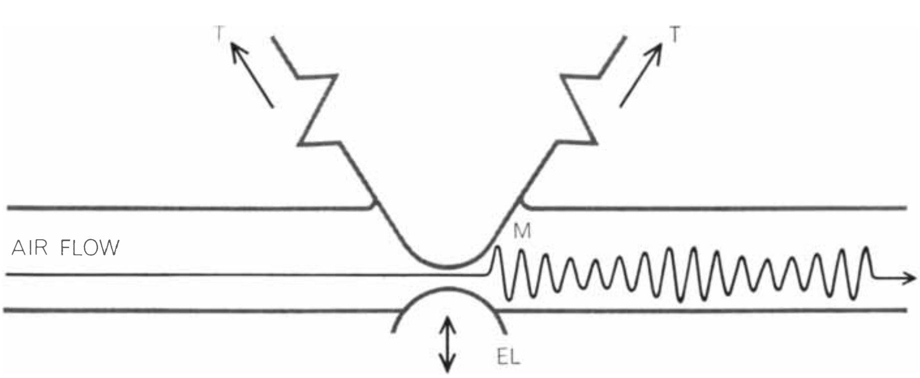
There is not much point in comparing the impressions bird songs make on human and on avian ears. The difference is moot; birds do not sing to us but to their own kind—to seduce a willing female, for example, or to warn off a potential male competitor. We are inclined to put everything we see or hear into our own standard of reference, and we wax lyrical over the song of a nightingale or a







REPETITIVE MODULATION is shown in a phrase sung by a song sparrow. The frequency of the modulation (*curve at bottom*) is fairly constant at some 300 cycles per second. The amplitude of the modulation, however (*oscillogram at top*), rises and falls from the beginning of the phrase to the end.



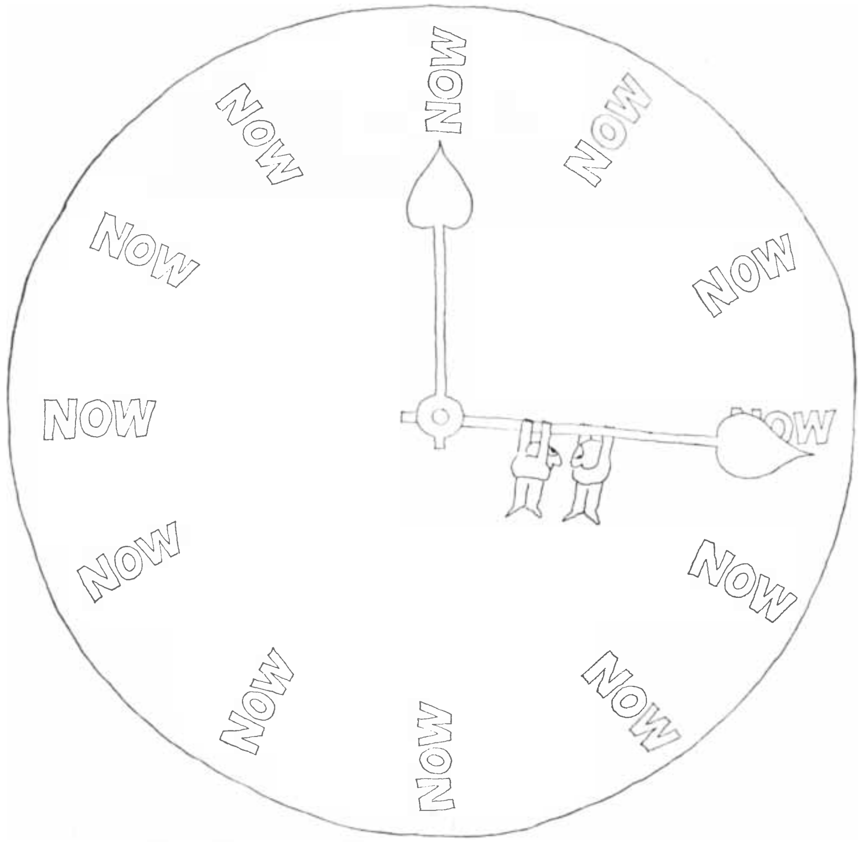
EXTERNAL LABIUM (*EL*), which is located in the bronchial wall opposite the tympanic membrane (*M*), may be the structure that enables the song sparrow and other birds to modulate amplitude without increasing or decreasing frequency. Normally a change in one alters the other, when both are controlled by tension on the tympanic membrane. If, however, the bird can diminish the distance across the bronchial lumen by extending the labium into it, it would be able to lower amplitude without affecting frequency, since this means of changing distance across the passage would not affect tension on the tympanic membrane.

thrush simply because they happen to sing within the ambit of our own musical experience. The fact is that the rapid modulations we have been discussing cannot be perceived as such by human ears. They are smeared, as it were, to produce what to us seems like a note of a different quality, or timbre, and one that on the whole we find unpleasant. It is easy for us to resolve a trill or a tremolo if its frequency is 30 cycles per second or less, but when the frequency rises to 100 cycles per second or more we hear a note of a rather unpleasant buzzy quality. Hence the beautiful complexity of the Lapland longspur song is completely lost in our ears, whereas it seems more than likely that Madame Longspur finds it delightful and enticing.

I have examined experimentally frequency perception and time perception for the avian ear. I conclude that its frequency discrimination, as expressed in the relation $\Delta f/f \times 10^{-3}$ (f is of course frequency), lies between 2 and 5. Time discrimination appears to be no greater than .5 millisecond. For human ears frequency discrimination is about the same, but time discrimination is perhaps 50 to 100 times worse for humans than it is for birds. There is then the strong presumption that birds hear *as such* the rapid modulations so characteristic of their songs, and that the information content even in relatively simple songs must be enormous. One readily understands how birds of the same species can recognize individuals from subtleties in their songs that are imperceptible to a human listener.

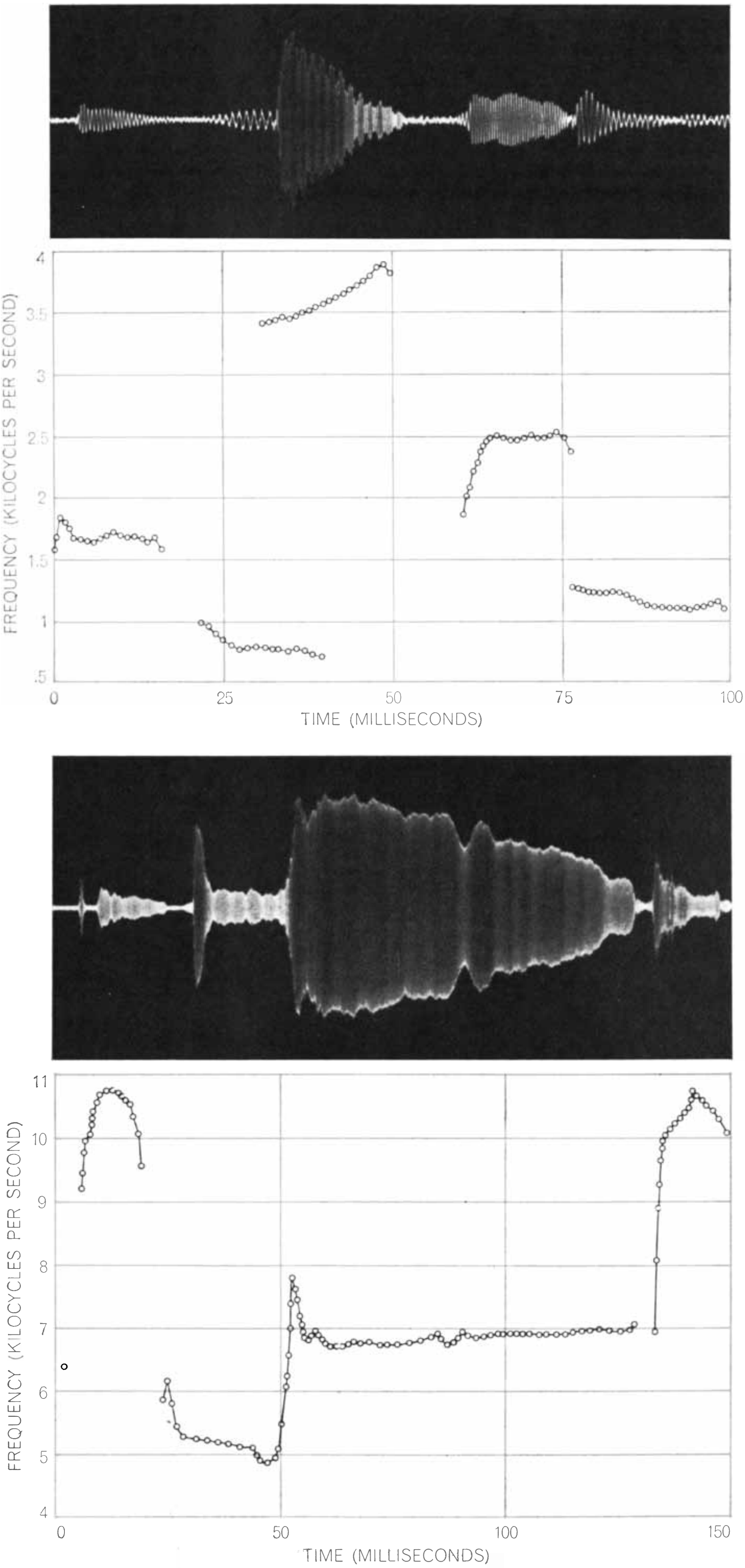
So far I have discussed only those phrases in bird song for which the basic wave form is sinusoidal without significant harmonic content. As we shall now see, harmonics do occur in bird song, but there is no broad generalization relating to the presence or absence of harmonics in the songs of the several bird families. One might say with some confidence that songs of the Passeriformes in which harmonics appear are relatively rare, whereas in the songs and calls of birds in other families phrases with substantial harmonic content are comparatively common. Each statement will, however, have numerous exceptions.

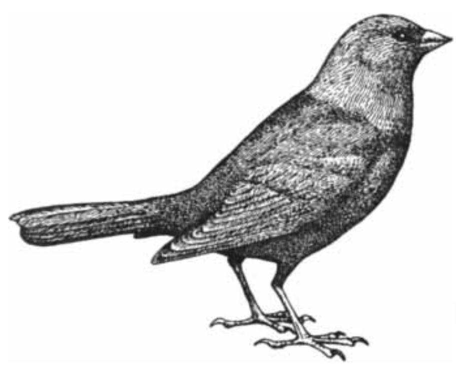
It can nonetheless be said with considerable assurance that for any given species there is a threshold frequency below which harmonics occur and above which one hears a phrase without significant harmonic content. This threshold varies widely for different species, from a value near 4,000 cycles per second for


NOW
NOW
NOW
NOW
NOW
NOW
NOW
NOW
NOW
NOW
NOW

the blue-gray gnatcatcher to below 500 cycles per second for the barred owl. Such threshold frequencies can be determined only when a bird sings over a range of frequencies that embrace the threshold. This circumstance is uncommon, because the majority of the Passeriformes sing only in the frequency range giving rise to phrases free of harmonics, and birds of other families sing only in the range giving rise to harmonic phrases. Indeed, whether or not a bird passes through the harmonic threshold in its songs or calls appears to be a matter of choice; there is no physiological reason that would prevent a crow from singing a harmonic-free phrase, or a wood warbler, by way of contrast, from singing a phrase with substantial harmonic content.

As an example of the development of harmonic spectra as the fundamental descends below the threshold frequency, I offer a glissando sung by a smooth-billed ani that embraces nearly three octaves (from 485 to 3,500 cycles per second). As the frequency rises dominance shifts from the fourth harmonic through the third and second harmonics to the fundamental frequency. The transition frequencies (the frequency at which the relative amplitudes of adjacent harmonics are equal) are 1,600 cycles per second for the fundamental and the second harmonic, 950 cycles per second for the second and third harmonics and 560 cycles per second for the third and fourth harmonics. In all the many cases I have examined in which the fundamental passes through the threshold, curves of relative amplitude plotted against frequency show a small frequency interval embracing each transition between





NONREPETITIVE MODULATION is seen in the courtship song of the brown-headed cowbird. In the oscillogram and frequency analysis at top is the phrase described as "glug"; in the corresponding records at bottom is the phrase "gleeee." Both of the phrases cover a large range of frequencies.

harmonics and a larger interval during which a particular harmonic contains a relatively large fraction of the acoustical energy. I have found no glissandi for which the dominant harmonic is higher than the fourth, but there are many calls in which the fundamental is constant, and in which the associated harmonic spectra show similar characteristics, that is, a dominant harmonic with adjacent harmonics falling off rapidly in relative amplitude.

To understand how harmonic spectra with these characteristics can be generated in the syrinx, let us return once again to my simplified model of the organ. Imagine the tension in the tympanic membrane gradually being reduced, with the membrane at its vibrational peak approaching the opposing bronchial wall (or the external labium) more and more closely. The point will come when the bronchial wall will *constrain* the membrane, forcing it to depart from a pure sinusoidal vibration. At this point the second harmonic will become evident, increasing in amplitude as membrane tension falls and the constraint of the opposing bronchial wall influences an increasing percentage of the period of vibration (the period of the fundamental). As the process continues, the amplitude of the fundamental will fall as the amplitude of the second harmonic rises. As membrane tension decreases still further, the second harmonic will become constrained and the third harmonic will become dominant. At this point the membrane can be visualized as being in a state of *rippling* vibration, with a fundamental fixed by membrane tension and the associated harmonic spectrum dictated by the constraints imposed by the passage within which the membrane is vibrating.

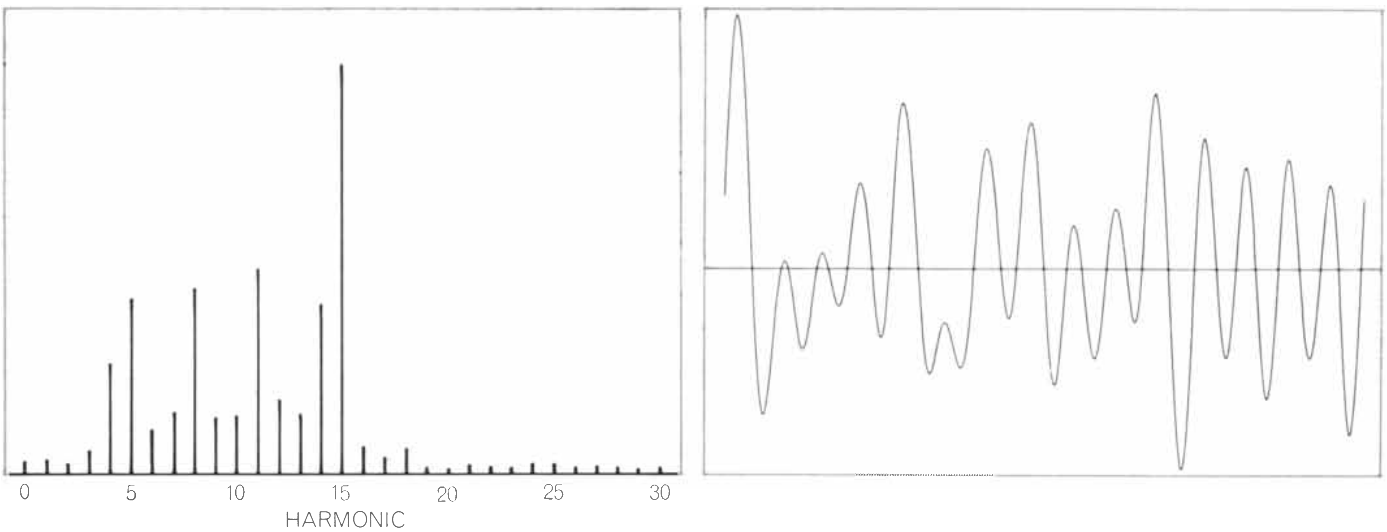
Among the ducks and geese harmonic spectra are found with many terms and with amplitudes showing no particular pattern. Such spectra must be associated with a form of membrane vibration resembling the pulses produced in the human glottis, that is, a high-amplitude pulse followed by a series of less violent ripples of variable period.

Throughout the harmonic domain, whatever the species and whatever the wave form or harmonic spectrum may be, the evidence shows quite clearly that there is no tracheal modulation; the call is produced in the syrinx and passes unchanged through the trachea.

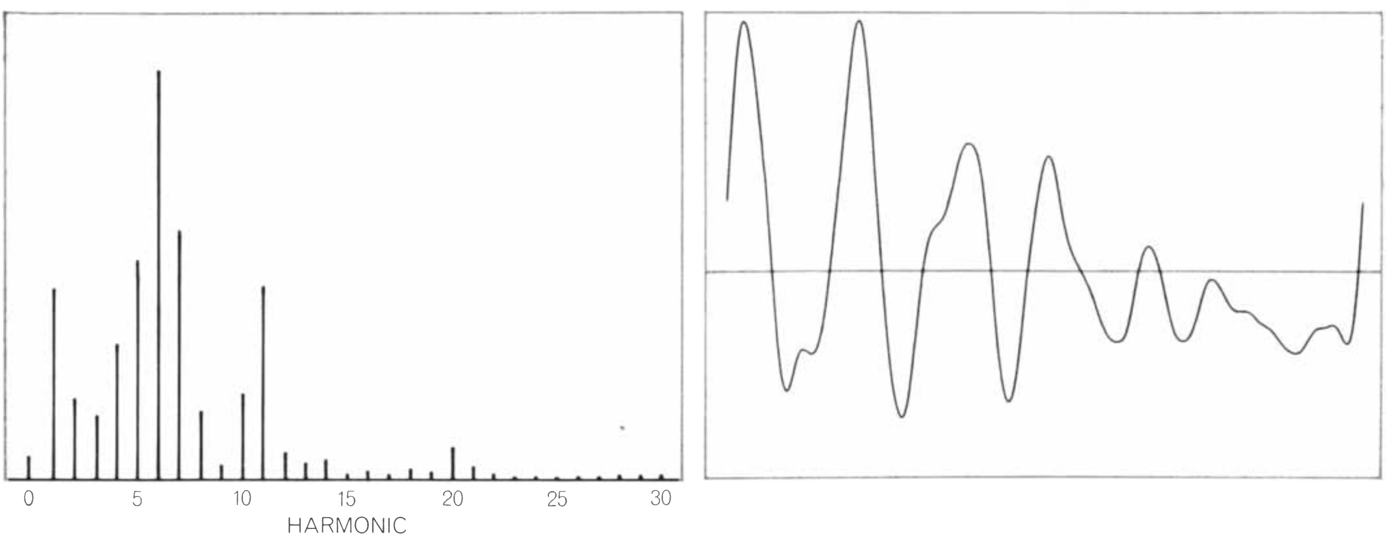
We come now to those birds that produce a more or less convincing imitation of the human voice. The question is whether such birds bring into play some completely new physiological or acoustical mechanism, or whether they achieve their imitations using the same mechanisms that produce their normal songs. I have selected the Indian hill myna as the standard of reference because its imitation of the human voice is excellent; it might well deceive the listener into thinking a person is speaking. A phrase was selected from the myna repertoire and was examined in detail in comparison with the same phrase spoken by me.

The reader will recall that the human voice originates as a series of puffs of air emanating from the glottis, the period between puffs corresponding to the fundamental frequency (80 to 180 puffs per second for the male voice). This acoustical disturbance is modulated in passing through our oral cavities so as to reinforce certain frequencies and attenuate others. Areas of reinforcement are called "formants" and are characteristic of particular vowel sounds. The wave form and the corresponding harmonic spectrum for the vowel "a" as enunciated by the myna and by me are shown in the illustrations below. The sound was unmistakably "ah" for both of us, but the associated wave forms are entirely different.

When the acoustical disturbance originating at the glottis stimulates a resonator (the oral cavities) to vibrate at a natural frequency, the generated wave form will have the frequency of the resonator, and the amplitude of the wave will decay exponentially over a period corresponding to that of the fundamental. One can readily isolate the formants for the myna and me corresponding to the vowel "a" and see if the resulting wave form fits these criteria. We find that the amplitude for my formants does indeed decay exponentially within the fundamental period; for the myna there is no decay and



INDIAN HILL MYNA SAYS "AH" as in "Charlie." At left is the harmonic spectrum of the sound. At right is the wave form of the sound, which does not decay in amplitude as it would if it were generated by a resonant system such as the human vocal apparatus.



AUTHOR SAYS "AH" as in "Charlie." The wave form shows a decay in amplitude. The two "ahs" sound alike, but one is generated by a nonresonant mechanism and the other by a resonant one. This is a fundamental difference between the vocal apparatus of birds and the vocal apparatus of men. The vertical lines in the harmonic spectra represent the relative intensity of sound at each harmonic.

only random change in amplitude. This is the most convincing evidence we can offer that resonators are not involved in the myna "imitation."

How, then, is the imitation produced? It is important first to realize that the myna need not reproduce the human wave form at all. Since the human ear is not sensitive to phase, the myna need produce only an approximation of the amplitudes of the several harmonics, and this can be done with literally an infinity of wave forms. The ability to produce a separate harmonic spectrum with each of the two acoustical sources should be adequate to produce a reasonably good imitation, particularly since it has been shown that much of the human voice spectrum is redundant even if the criterion is recognition of the speech of a particular individual.

Let me summarize. The physiology and acoustics of bird vocalization are unique in the animal kingdom. Sound is produced at the syrinx in an air stream modulated by an elastic membrane vibrating in a restricted passage bounded by the walls of the bronchus. This source-generated acoustical disturbance appears not to be modified in its passage through the trachea. The syrinx contains two independently controllable sources, one in each bronchus, enabling the bird to produce two notes or phrases simultaneously. Harmonics arise below a threshold frequency by mechanical constraints on the vibrating membrane, forcing a departure from a purely sinusoidal wave form. The source-generated sounds can be modulated in frequency or in amplitude or (more usually) in both with extraordinary rapidity, so rapidly that human ears cannot perceive the modulators as such, receiving instead impressions of notes of varying quality or timbre.

I end this account by pointing out that it has been in effect a scientific detective story, with conclusions reached by analyzing the evidence in the bird songs themselves. The criminal did not confess in the last chapter, and the evidence must remain circumstantial, without direct proof. Had I the deductive powers of an Albert Campion, a Gideon Fell or any of the other erudite detectives of fiction, coupled with the persuasiveness of Perry Mason, I might have done better. In any event I have developed a model, highly convincing to me, and I shall patiently await experimental evidence that will raise my spirits if the answer is yes but will not be too devastating if it should be no.

MATHEMATICAL GAMES

A new pencil-and-paper game based on inductive reasoning

by Martin Gardner

Many games and pastimes have flimsy analogies with induction, that strange procedure by which scientists observe that some ostriches have long necks and conclude that all unobserved ostriches also have long necks. In poker and bridge, for instance, players use observational clues to frame probable hypotheses about an opponent's hand. A cryptographer guesses that a certain "pattern word," say BRBQFBQF, is NONSENSE, then tests this inductive conjecture by trying the letters elsewhere in the message. An old parlor entertainment involves passing a pair of scissors around and around a circle of players. As each person transfers the scissors he says "Crossed" or "Uncrossed." Those acquainted with the secret rule tell a player when he says the wrong word, and the joke continues until everyone has guessed the rule inductively. The scissors' blades are a red herring; a player should say "Crossed" if and only if his *legs* are crossed.

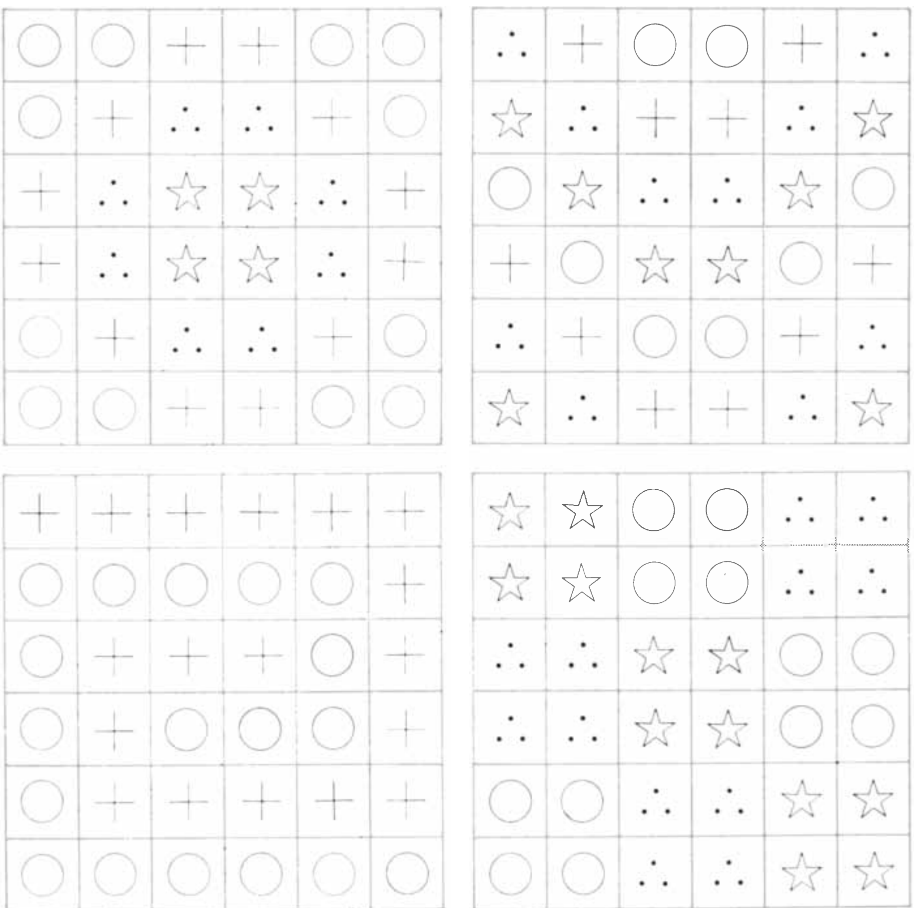
Familiar games such as Battleship and Jotto have slightly stronger analogies with scientific method, but the first full-fledged induction game was Eleusis, a card game invented by Robert Abbott and first explained in this department in June, 1959. (Fuller details are in *Abbott's New Card Games,* a Stein and Day hard-cover book in 1963 and a Funk & Wagnalls paperback in 1969.) Eleusis intrigued many mathematicians—notably Martin D. Kruskal of Princeton University, who worked out an excellent variant that he described in 1962 in a privately issued booklet, *Delphi—a Game of Inductive Reasoning.*

In Eleusis and Delphi a secret rule, specifying the order in which single cards may be played, corresponds to a law of nature. Players try to guess the rule inductively and then (like scientists) test their conjectures. This month I shall explain a new type of induction game called Patterns, devised by Sidney Sackson and included in his delightful book *A Gamut of Games,* to be published this fall by Random House.

Patterns is a pencil-and-paper game that can be played by any number of people, although preferably no more than six. It differs markedly from Eleusis and Delphi, but it shares with them such a striking similarity to scientific method that many thorny problems about induction that have needled philosophers of science ever since David Hume showed induction has no logical justification have pleasant analogues in the game.

Each player draws a square six-by-six grid on a sheet of paper. A player called the Designer (the role of Designer passes to another player with each new game) secretly fills in his 36 cells by drawing in each cell one of four different symbols. Sackson suggests the four shown in the illustration at the left, but any other four may be used. The Designer, who can be regarded as Nature, the Universe or the Deity, is free to mark the cells as he likes; they may form a strong or a weakly ordered pattern, a partially ordered pattern or no pattern at all. However (and here Sackson adopts the brilliant original idea of Abbott's), the method of scoring is such as to impel the Designer to create a pattern, or a regularity of nature, that is easy to discover for at least one player and yet difficult enough to be missed by at least one other player.

Four typical patterns given in Sackson's book are arranged roughly in order of difficulty [*see illustration at left*]. All have some type of visual symmetry, but nonsymmetrical forms of order can be used if the players are mathematically sophisticated. For example, a Designer might take the cells in sequence, left to right and top to bottom, putting a plus sign in each cell whose number is prime and a star in all the remaining cells. The basis for ordering the Master Pattern is intimately bound up with the Designer's estimate of the abilities of the other players because, as we shall see, he makes his



Patterns for Sidney Sackson's induction game, all showing forms of symmetry

SCIENCE/SCOPE

The TOW anti-tank missile scored "ten for ten" in a recent test by the U.S. Army Missile Command at Redstone Arsenal, Ala. All ten missiles struck targets more than a mile away "with pinpoint accuracy." They were part of the first TOW production hardware delivered to the Army by Hughes. Wire-guided TOW, which can be fired from a ground emplacement or from helicopters and a variety of vehicles, can knock out field fortifications or destroy any known enemy armor.

Ion beam "sputtering" -- the life-limiting erosion that plagues designers of vacuum tubes and ion-propulsion engines -- has been turned into a precision tool for microelectronic fabrication by Hughes research scientists. Their ion-beam micromachining technique can remove and create patterns in any substance and permits the use of new materials. The ion beam's directionality makes it superior to wet chemical etches for removing material in very fine patterns (line widths down to 1/2-micron have been achieved).

The prototype of an advanced radar -- forerunner of what could be the world's most powerful for defense against missile attack in the 1970s and beyond -- is now undergoing system tests at Hughes. It is being built under a multi-million-dollar Advanced Research Projects Agency/U.S. Army contract administered by the U.S. Air Force's Rome Air Development Center. Though the prototype will be only 1/50 the power of the proposed long-range system called ADAR (for Advanced Design Array Radar), it will be the most powerful radar yet built by Hughes.

The first Phoenix missile system trainer, delivered recently to the U.S. Navy at the Pt. Mugu, Calif., Naval Missile Center, will be used to train missile control officers for the Navy's new F-14A fighter. The Hughes-built simulator eliminates costly, time-consuming in-flight training. Using computers, tapes, and displays, it simulates a complete mission profile, from target acquisition to lock-on and missile launch.

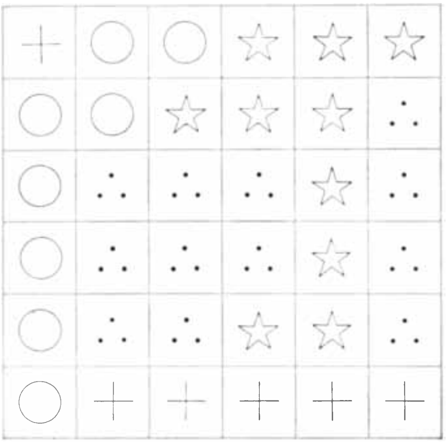
Hughes needs experienced engineers: Microcircuit, digital communication system analysis, computer systems, digital systems test, signal processing, circuit design, microwave solid state, radar systems. Also: real-time and weapon system programmers. A B.S. degree, two years of related experience, and U.S. citizenship are required. Please write: Mr. J. C. Cox, Hughes Aircraft Company, P.O. Box 90515, Los Angeles 90009. Hughes is an equal opportunity employer.

The U.S. Air Force's new Maverick missile passed its first air-launched test at Edwards AFB, Calif., recently -- just over a year after it went into development at Hughes. It was launched from an F-4, but is also designed for use on the A-7D and other aircraft. Mission of the air-to-ground Maverick is to attack small, hard, tactical targets, such as tanks and field fortifications. Automatic TV guidance will enable it to track its target after launch without further help from the launching aircraft.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY



How is this pattern ordered?

highest score when one player does very well and another very poorly. Can the reader discern the simple basis for the nonsymmetrical ordering shown above before it is given next month?

The Designer puts his sheet face down on the table. Any player may now make inquiries by drawing on his own grid a small slant line in the lower left-hand corner of any cell about which he seeks information. His sheet is passed face down to the Designer, who must enter the correct symbol in each cell in question. There are no turns. A player may ask for information whenever he wants, and there is no limit to the number of cells about which he may inquire. Each request represents an observation of nature—or an experiment, which is simply a controlled way of making special observations; cells filled in by the Designer correspond to the results of such observations. A player could ask for information about all 36 cells and obtain the entire pattern at once, but this is not to his advantage because, as we shall learn, it would give him a score of zero.

When a player believes he has guessed the Master Pattern, he draws symbols in all his untested cells. To make it easy to identify these inductions, guessed symbols are enclosed in parentheses. If a player decides he cannot guess the pattern, he may drop out of the game with a zero score. This is sometimes advisable because it prevents him from making a minus score and also because it inflicts a penalty on the Designer.

After all players have either filled in all 36 cells or dropped out of the game, the Designer turns his Master Pattern face up. Each player checks his guesses against the Master Pattern, scoring +1 for every correct symbol, −1 for every incorrect symbol. The sum is his final score. If he made a small number of inquiries and correctly guessed all or most of the entire pattern, his score will be high. If he has more wrong than right guesses, his score is negative. High scorers are the brilliant (or sometimes lucky) scientists; poor scorers are the mediocre, impulsive (or sometimes unlucky) scientists who rush poorly confirmed theories into print. Dropouts correspond to the mediocre or overcautious scientists who prefer not to risk framing any conjecture at all.

The Designer's score is twice the difference between the best and the worst scores of the others. His score is reduced if there are dropouts. Five points are subtracted for one dropout, 10 for each additional dropout. Sackson gives the following examples of games with a Designer (*D*) and players *A, B, C:*

If *A* scores 18, *B* scores 15 and *C* scores 14, *D*'s score is 8, or twice the difference between 18 and 14.

If *A* scores 18, *B* scores 15 and *C* scores −2, *D*'s score is 40, or twice the difference between 18 and −2.

If *A* scores 12, *B* scores 7 and *C* drops out with a score of 0, *D*'s score is 19, or twice the difference between 12 and 0, with five points deducted for the single dropout.

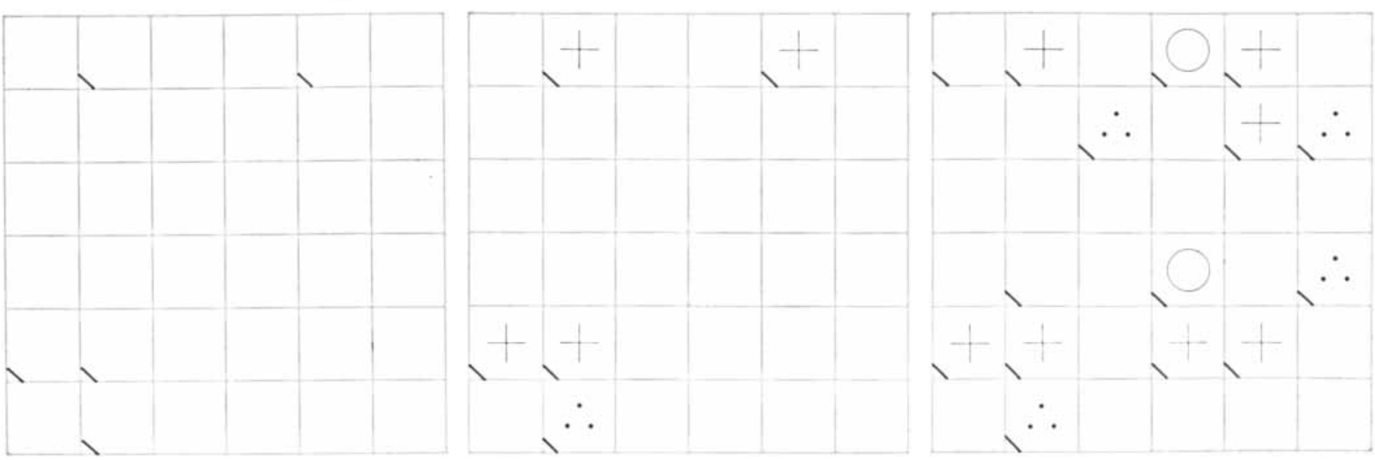
If *A* scores 12 and *B* and *C* both give up, *D* scores 9. This is twice the difference between 12 and 0, with five points deducted for the first dropout, 10 for the second.

If all three players drop out, *D*'s score is −25. His basic score is 0, with 25 points subtracted for the three dropouts.

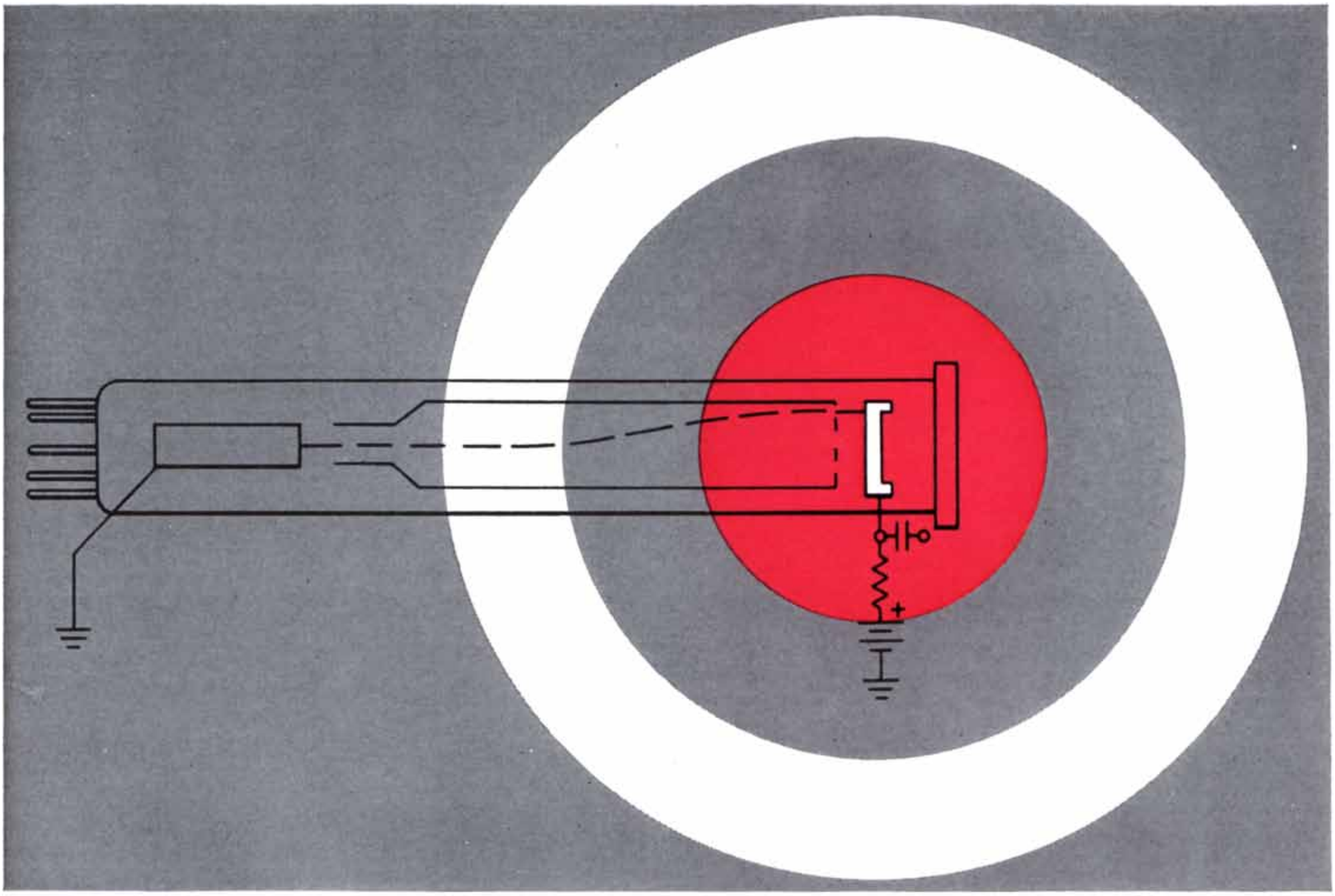
An actual game played by Sackson suggests how a good player reasons [*see illustration below*]. The five initial inquiries probe the grid for evidence of symmetry [*left*]. The sheet is returned with the five symbols filled in [*middle*]. A series of additional inquiries brings more information [*right*]. It looks as if the pattern is symmetrical around the diagonal axis from top left to bottom right. Since no stars have appeared, Sackson induces that they are absent from the pattern.

Now comes that crucial moment, so little understood, for the intuitive hunch or the enlightened guess, the step that symbolizes the framing of a hypothesis by an informed, creative scientist. Sackson guesses that the top left-hand corner cell contains a circle, that the three cells flanking it all have plus marks and that, continuing down the diagonal, the pluses are flanked by three-spot symbols, the pattern repeating itself with larger borders of the same three symbols in the same order. To test this conjecture with as few new inquiries as possible Sackson asks for information on only two more cells, the two cells shown empty but with slant lines on the grid at the right in the illustration below.

If those cells do not contain circles, his conjecture is false. As the philosopher Karl Popper maintains, the "strongest"



Three stages in probing for the Master Pattern



A solid state target in a camera tube

Bell System PICTUREPHONE® service will need small, reliable TV camera tubes for use in offices and homes, where lighting ranges from dim to very bright. Conventional vidicon tubes are unsuitable, so Bell Labs developed a new kind.

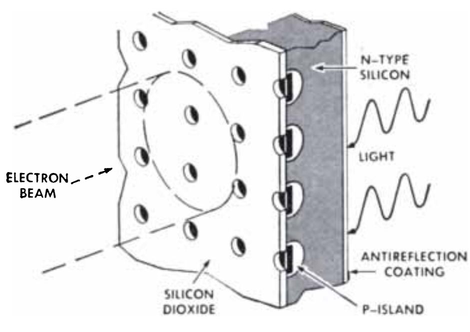
The heart of the new tube is a light-sensitive target containing nearly 700,000 silicon photodiodes in an area less than a half inch square. They are made by diffusing boron, a p-type impurity, through a silicon-dioxide mask into n-type silicon.

A scanning electron beam charges the p material negatively, reverse-biasing the diodes. Holes, created by incident light, are collected by the electric field at the p-n junctions, and individual diodes discharge by an amount proportional to the local light intensity. Recharging of the diodes by the scanning electron beam produces a varying current . . . the output signal.

Among the tube's advantages:

Its target tolerates high-temperature baking . . . a processing step to improve reliability. Conventional vidicon targets cannot stand this.

Silicon's high thermal conductivity and chemical stability help make the new tube immune to "burn-in" (degradation of performance from continuous exposure to a fixed image, very bright light, or a strong scanning electron beam).



The time between a change in target illumination and a like change in output is much shorter with the new target. This improves response to fast-changing scenes.

The light-sensitive face of the new target is optically flat. So, a multilayer antireflection coating can be applied for better sensitivity and minimum received-picture "halo".

Silicon targets have relatively uniform response through the visible and near-infrared—from 4,000 to 9,000 A. Quantum efficiency (electrons per photon) exceeds 0.5. So, these targets have at least 10 times the sensitivity of a standard vidicon camera tube in incandescent light.

This new camera tube is in the latest model PICTUREPHONE set, now undergoing field trials.

From the Research and Development Unit of the Bell System—

Bell Labs

conjecture is the one that is easiest to falsify, and Popper considers this the equivalent of the "simplest" conjecture. In Sackson's game the strongest (and simplest) conjecture is that every cell contains the same symbol, say a star. It is strong because a single inquiry about *any* cell, answered by anything but a star, falsifies it. The weakest conjecture is that each cell contains one of the four symbols. Such a hypothesis can be completely confirmed. Since no inquiry can falsify it, however, it is a true but useless hypothesis, empty of all empirical content because it tells one nothing about the Master Pattern.

The circles turn out to be where Sackson expected them. This increases what the philosopher Rudolf Carnap calls the "degree of confirmation" of Sackson's hypothesis in relation to the total evidence he has bearing on it. Sackson decides to take the inductive plunge and "publish" his conjecture. He fills in the empty cells of his grid. When his pattern is compared with the Master Pattern [*see illustration below*], a count of the guessed symbols (in parentheses) shows that Sackson has 20 right and one wrong, for a score of 19.

The single star Sackson missed is unexpected, but it is typical of the surprises Nature often springs. Science is a complicated game in which the universe seems to possess an uncanny kind of order, an order that it is possible for humans to discover in part, but not easily. The more one studies the history of the game of science, the more one has the eerie feeling that the universe is trying to maximize its score. A splendid recent example is the independent discovery by Murray Gell-Mann and Yuval Ne'eman of the "eightfold way." This is a symmetry pattern, defined by a continuous group structure, into which all the elementary particles seem to fit. As soon as enough information had accumulated the pattern was simple enough to be spotted by two physicists, and yet it remained complicated enough to be missed by all the other players.

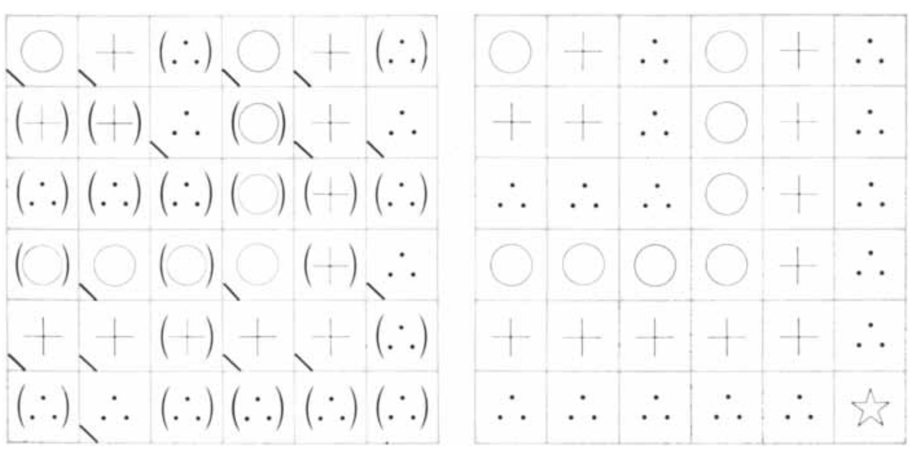
Sackson, the inventor of Patterns, is a professional engineer who works on steel bridges and buildings. Collecting, studying and inventing games has been his lifelong avocation. He owns what is probably the largest private collection of modern proprietary games, books on games, and notes obtained by painstaking research in the world's great libraries and museums. He has invented hundreds of games. The first, he discloses in his book, was invented when he was in the first grade; it had to do with circling words on a page and joining them in chains. The first board game he owned was Uncle Wiggily, a track game that is still on the market. He immediately modified it by altering its rules and substituting toy soldiers for rabbits to make it a war game.

All Sackson's marketed games emphasize intellectual skill rather than luck. Acquire, a game based on the theme of investing in hotel chains, has been his best-selling item. His other commercial games include The Case of the Elusive Assassin (a logic game based on Venn diagrams), Focus, Bazaar and two card games using special decks, Venture and Monad, both issued this year. Tam-Bit, Take Five, Odd or Even, Tempo and Interplay are among other Sackson board games that will go on sale in 1970.

A Gamut of Games is unique in that almost every one of its 38 games will be completely unfamiliar to any reader. All can be played with equipment that is easily acquired or constructed: cards, dice, dominoes, counters and checkerboards. Twenty-two are Sackson originals. The others are either creations of Sackson's game-inventor friends or old, forgotten games that deserve revival. No two readers will, of course, have the same reaction to every game. I particularly like Knight Chase, played on a chessboard with one black and one white knight and 30 small counters. It is the invention of Alexander Randolph, Czechoslovakian-born but now living in Japan, who has several excellent games on the U.S. market: Oh-Wah-Ree (based on the African game of mancala), Twixt and Breakthru. Another mathematically appealing game (which Sackson found in an 1890 book) is Plank, a version of ticktacktoe played with 12 tricolored cardboard strips. The reader will find a valuable bonus in the book's final section: brief reviews of more than 200 of the best adult games now on sale in this country.

Sackson's informal text is interspersed with personal anecdotes and snippets of surprising historical data. Until I read his book I did not know that the 17th-century poet Sir John Suckling invented cribbage, or that Monopoly, the most successful of all proprietary board games, is derived from The Landlord's Game, which was patented in 1904 by one Lizzie J. Magie and was intended to teach Henry George's single-tax theory. Sackson reproduces the patent drawing of the Magie board; the similarity to Monopoly is obvious.

Marketed board games, Sackson points out, tend to reflect major events and interests of the time. Although he does not mention it, an amusing example of this is The Money Game, a card game invented by Sir Norman Angell, who received the Nobel peace prize for 1933. The special cards and miniature money for this stock-market-speculation game were packaged with a 204-page explanatory book issued by E. P. Dutton, with puffs on the jacket by Walter Lippmann, John Dewey and noted economists. (A book by "Adam Smith" with the same title and topic has recently been a best seller.) Why is Angell's Money Game so grimly amusing? The year of its publication was 1929.



Player's grid (left) *compared with Master Pattern* (right)

Last month's questions are answered as follows:

1. The anagram of "moon starers" is "astronomers."

2. The solution to SPIRO × 7 = AGNEW is 14,076 × 7 = 98,532. This unique solution was first published by E. P. Starke in *The American Mathematical Monthly,* Volume 53, January, 1946, page 45. In *Mathematics Magazine,* Volume 42, March, 1969, pages 102–103, David E. Daykin gives two tables of computer re-

how to slash $30,000 per month  from your computer overhead

DatagraphiX systems save a fortune in machine rentals, staff time, computer time, print time, 87% less paper/carbon consumables, 99% less storage. Compared to impact printing, DatagraphiX output is roughly 27 times faster, takes 1/18th the computer time, cuts 7/8ths of your supplies cost. If you do multi-part printing in any volume, your paper consumption alone can run into 6 figures. One user switched to DatagraphiX Micromation after discovering the purchase price of his paper had exceeded $300,000 for one year.

Micromation also provides multiple economies in retention and retrieval. Translating computed data into report formats and shrinking storage requirements to 1/525th. Offering access to thousands of facts within a matter of seconds from screen display inquiry stations. Instant replay of the facts in time to effect better decisions, improved profits. Hard copies on demand.

For high volume production printing, paper copies can be produced from data film at 5,200 pages per hour. That means you could turn out 20,000 bank statements in a lunch hour. Or thousands of direct mail pieces on preprinted, multi-color forms.

Only one company offers the complete family of machine systems; service centers; Kalvar dry film processing; all associated supplies; systems and software support; worldwide maintenance. Discover what Micromation can do for you. Contact our local office or National Sales Manager, James P. Whitfield.

DatagraphiX
micromation systems

Stromberg DatagraphiX Inc., a General Dynamics subsidiary, P.O. Box 2449, San Diego, California 92112 (714) 283-1038 TWX: (910) 335-2058 sults that list the number of different solutions of $A = kB$, in all number systems with bases 3 through 15, where k is any number from 2 through 14 and A and B are numbers that together contain all 10 digits once each or all nine digits (0 excluded) once each. Solutions in which A or B begin with 0 are not considered. In Dr. Matrix' problem (base 10 and all 10 digits) there are 48 solutions for $k = 2$, six for $k = 3$, eight for $k = 4$, 12 for $k = 5$, none for $k = 6$, one for $k = 7$, 16 for $k = 8$ and three for $k = 9$.

3. The earliest A.D. year with the property explained last month for 1969 is the year 19. The square of 19 is 361, the last two digits of which, upside down, are also 19. (The trivial case of year 1 was excluded.)

4. A year can have no more than four "perverse months"—months requiring six calendar rows—and no fewer than two. It can have no more than three Friday the 13th's and no fewer than one. As Kirby A. Baker has shown, the number of perverse months in a year plus the number of Friday the 13th's is four, except for years beginning on Sunday and leap years beginning on Saturday, in which cases the sum is five. The last year with four perverse months was 1944; the next will be 1972, with no more in this century.

	JAN 1 2 3
JAN 1 2 3	FEB 1
MAR 1 2 3	MAR 1 2 3
APR 1 2	APR 1 2
MAY 1 2 3	MAY 1 2 3
JUN 1 2	JUN 1 2
JUL 1 2 3	JUL 1 2 3
AUG 1 2 3	AUG 1 2 3
SEP 1 2	SEP 1 2
OCT 1 2 3	OCT 1 2 3
NOV 1 2	NOV 1 2
DEC 1 2 3	DEC 1 2 3

Pseudochronian year and leap year

To prove that four and two are the maximum and minimum numbers of perverse months in any year, Baker suggests the following exercise: Eliminate the last four weeks of each month. The remaining days (two or three per month except for February, which in leap years has one day and in nonleap years has no days) form the calendar used by inhabitants of a planet that Baker calls Pseudochron. Pseudochronian years are 29 or 30 days long, but their weeks have seven days, like ours.

Pseudochronian years have two schema [*see illustration on this page*]. In each case a Pseudochronian month starts and ends on the same days as a corresponding earth month for a corresponding year. For example, 1969 on the earth started and will end on Wednesday. Label the dates on the Pseudochronian nonleap year, beginning with Wednesday, and you will find the year ends on Wednesday. Moreover, each Pseudochronian month begins and ends on the same days as the corresponding earth months of 1969.

A perverse earth month corresponds to a Pseudochronian month that includes both Saturday and Sunday. If we try the seven possible ways to label the dates of a Pseudochronian nonleap year, starting the year with each of the seven days, we find that it can have no fewer than two perverse months and no more than three. A similar procedure with the Pseudochronian leap year gives no fewer than two perverse months, but if and only if the year begins with Saturday, there are four perverse months.

Since Friday the 13th occurs only in months starting with Sunday, a similar test will prove that no year can contain fewer than one Friday the 13th or more than three. I leave it to readers to establish the relationship between perverse months and Friday the 13th's.

A proof by B. H. Brown that the 13th is more likely to fall on Friday than on any other day appeared in *The American Mathematical Monthly*, Volume 40, December, 1933, page 607; a proof by S. R. Baxter (done at the age of 13) appeared in *The Mathematical Gazette,* Volume 53, May, 1969, pages 127–129.

Many short proofs of the maximum and minimum numbers of Friday the 13th's in a year have been published. For a good recent discussion of this problem see William T. Bailey, "Friday-the-Thirteenth," in *The Mathematics Teacher*, Volume 62, May, 1969, pages 363–364.

The Koerners aren't trying to save the world. Just a little piece of it.

Romero is a small farm village clinging to the mountains of Colombia, South America. Poverty is extreme. Illiteracy is almost total. Electricity, sanitary facilities and running water are nowhere to be seen, and small children haul water from the river three miles away. There is a school in the district, but convincing parents that education should take precedence over water-hauling is a hard argument to win.

The Fuentes are one of the village families. Poor even by Romero's sub-standards. The mud-brick house with dirt floors they live in isn't even theirs. The two-acre farm behind it provides all the income for the family of nine. Most of the time there isn't enough to go around. Yet, this has been a good year for the Fuentes. Their 7-year-old daughter Luz Marina has been able to stay in school.

This miraculous stroke of luck is because of a family living in New York, who are helping Luz Marina. They're the Koerners. Through Save The Children Federation, they are contributing $15.00 a month to help.

For Richard and Marianne Koerner the $15.00 a month is not an extreme sacrifice. Yet, with three robust boys of their own, there are many ways they could use the money. But Richard, whose job involves foreign travel, doesn't have to imagine the poverty the Fuentes endure. He's seen it. He knows, too, that Save The Children can do a remarkable number of things with the money.

First, Luz Marina's immediate school needs are taken care of. Secondly, funds are available to the family to carry out their self-help plan to raise hens. The eggs will help improve the family diet and increase the family income. Finally, a small portion of the Koerner's contribution, together with money from other sponsors, has been lent to the villagers. With their own hands and sweat, they're building a water storage tank and a pipeline. Hopefully, the new water supply will increase the crop yield and the village's income so that some day they will no longer need help.

Self-help. That's what Save The Children Federation is all about. Although contributions are tax-deductible as a charity, the aim is not merely to buy a child a few hot meals, a new coat, or shoes. Instead, your contribution is used to give people the boost they need to start helping themselves.

Sponsors are desperately needed for children in Korea, Vietnam, Latin America, Africa, Greece, the Middle East, as well as Appalachian and American Indian children. You can select the child's nationality and will receive a photo, regular progress reports and a chance to correspond and visit.

Richard Koerner hopes that one day, one of his business trips will enable him to meet the Fuentes. Many sponsors do visit their children, and tell us that it is one of the most gratifying experiences of their lives.

The Koerners know they can't save the world for $15 a month. Just a small corner of it. But, maybe that is the way to save the world. If there are enough people like the Koerners. How about you?

National Sponsors (partial list): Faith Baldwin, Hon. James A. Farley, Gene Kelly, Mrs. Eli Lilly, Paul Newman, Mrs. J. C. Penney, Frank Sinatra.

Save The Children Federation, founded in 1932, is registered with the U.S. State Department Advisory Committee on Voluntary Foreign Aid.

Save The Children Federation

NORWALK, CONNECTICUT 06852

I WISH TO CONTRIBUTE $180 ANNUALLY TO HELP A CHILD.
☐ WHERE THE NEED IS GREATEST ☐ LATIN AMERICA
☐ AMERICAN INDIAN ☐ APPALACHIA ☐ KOREA ☐ GREECE
☐ VIETNAM ☐ AFRICA ☐ MIDDLE EAST

ENCLOSED IS MY FIRST PAYMENT
☐ $15.00 MONTHLY ☐ $90.00 SEMI-ANNUALLY
☐ $45.00 QUARTERLY ☐ $180.00 ANNUALLY

I CAN'T SPONSOR A CHILD. ENCLOSED IS A CONTRIBUTION OF $________

☐ PLEASE SEND ME MORE INFORMATION.

NAME________

ADDRESS________

CITY________ STATE________ ZIP________

CONTRIBUTIONS ARE INCOME TAX DEDUCTIBLE SA 10/9

Education

That truck in our picture causes quite a stir as it rolls from town to town. Some kids think it is a traveling circus. Ah well.

Its destination is not the fairground but the local schoolyard. Here it opens up into a classroom for thirty students. It has heat, air conditioning, a videotape system, a computer and more sight-and-sound devices than most local schools ever have.

All these gadgets have a single purpose. To bury the old lecture-and-listen teaching methods under the ivy. And substitute involvement.

Funded by a grant from Standard Oil Company (New Jersey) and built by the Education Development Center in Newton, Mass., the rolling classroom is now out on its first serious job.

Only one Massachusetts school in five now has a kindergarten class. State law will soon establish kindergartens in all public schools. A thousand new teachers will be needed. The rolling classroom is helping to train them.

This is a modest project. But it typifies many educational innovations now being funded by Jersey, its affiliates, and the Esso Education Foundation. Some of these experiments are already silver linings in a thundery sky.

Harlem Prep School is one. Nearly all of its

on the move.

students are dropouts from the public school system. Thirty entered college last year. Seventy-five have been admitted this year.

Another program has come up with some answers to the psychological problems of bright students who mysteriously flunk out of college. Since the national flunking rate is estimated at 400,000 a year, this experiment may have effects out of all proportion to its size.

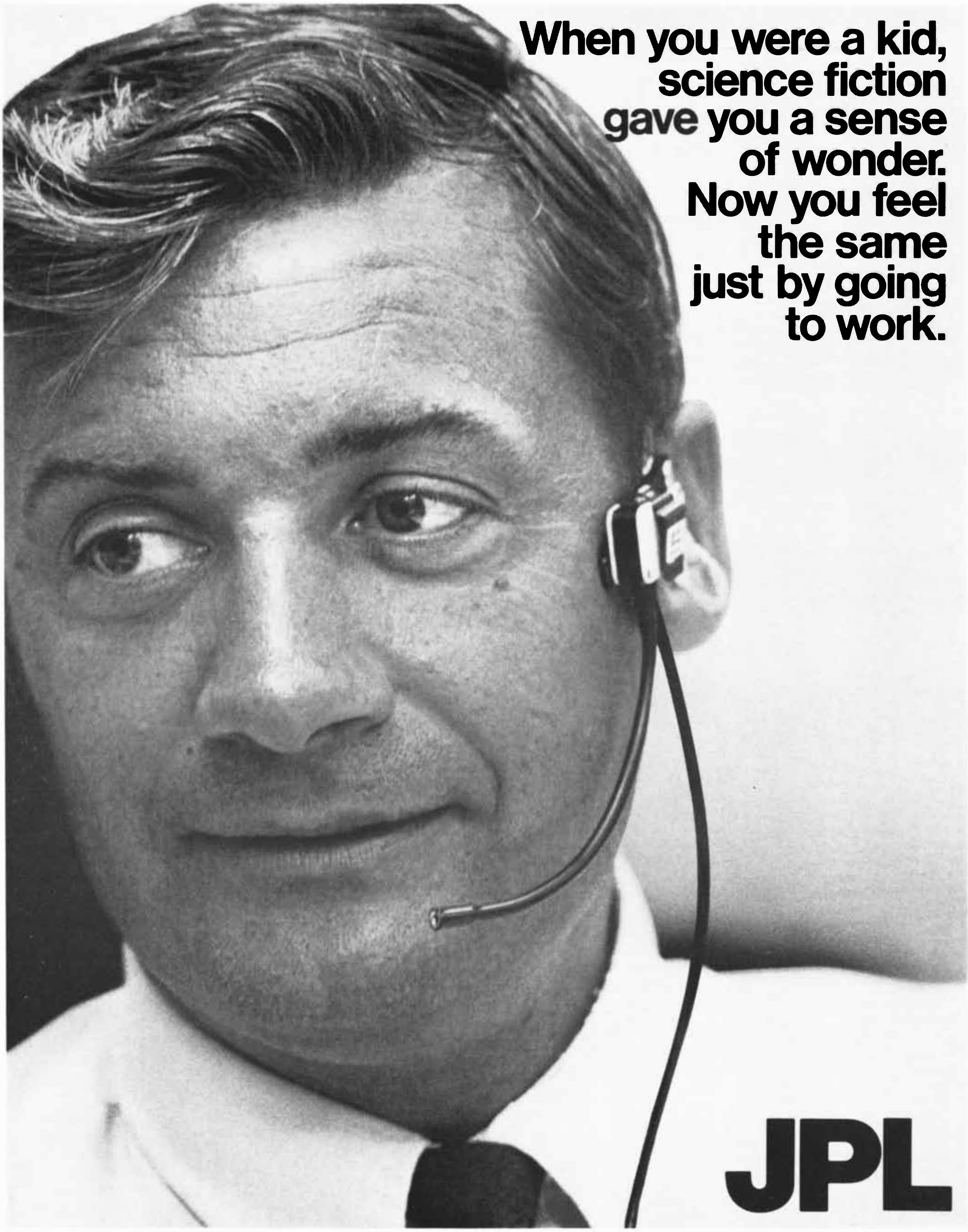
And there's one newborn venture that deserves a cheer from every campus. The Esso Education Foundation is now supporting a lively new magazine called "Change". It is dedicated to exploring new directions in higher education and hopes to find sensible solutions to the problems that underlie student unrest. Listen to the editor:

"We intend to be an irreverent foe of all that is arcane, banal and irrelevant in higher education. We intend to help return to higher education that long lost art—the plain English sentence."

Knowledge keeps no better than fish. The plain English sentence should speed its delivery enormously.

Standard Oil Company (New Jersey)

Esso



In fact, you've got what the great science-fiction heroes never had so good: the excitement of discovery. If you want to be part of the action, we can help.

Right now openings exist in: Spacecraft Structural Design · Temperature Control Materials & Processing · Application of Microelectronics & Transistors · Guidance & Control Systems · Electro Optics · Propulsion Systems Analysis · Space Vehicles Design · Trajectory Design & System Analysis · Systems Design & Integration · Deep Space Support Systems · Electromechanical Spacecraft Instrumentation · Scientific Programming · TV Image Processing and Electronic Packaging.

Send your resume to:
JET PROPULSION LABORATORY
California Institute of Technology
4802 Oak Grove Drive,
Pasadena, California 91103
Attn: Professional Staffing Dept. 11

"An equal opportunity employer." Jet Propulsion Laboratory is operated by the California Institute of Technology for the National Aeronautics and Space Administration.

THE AMATEUR SCIENTIST

Old refrigerators are salvaged to build a laboratory cooler and a gas liquefier

Conducted by C. L. Stong

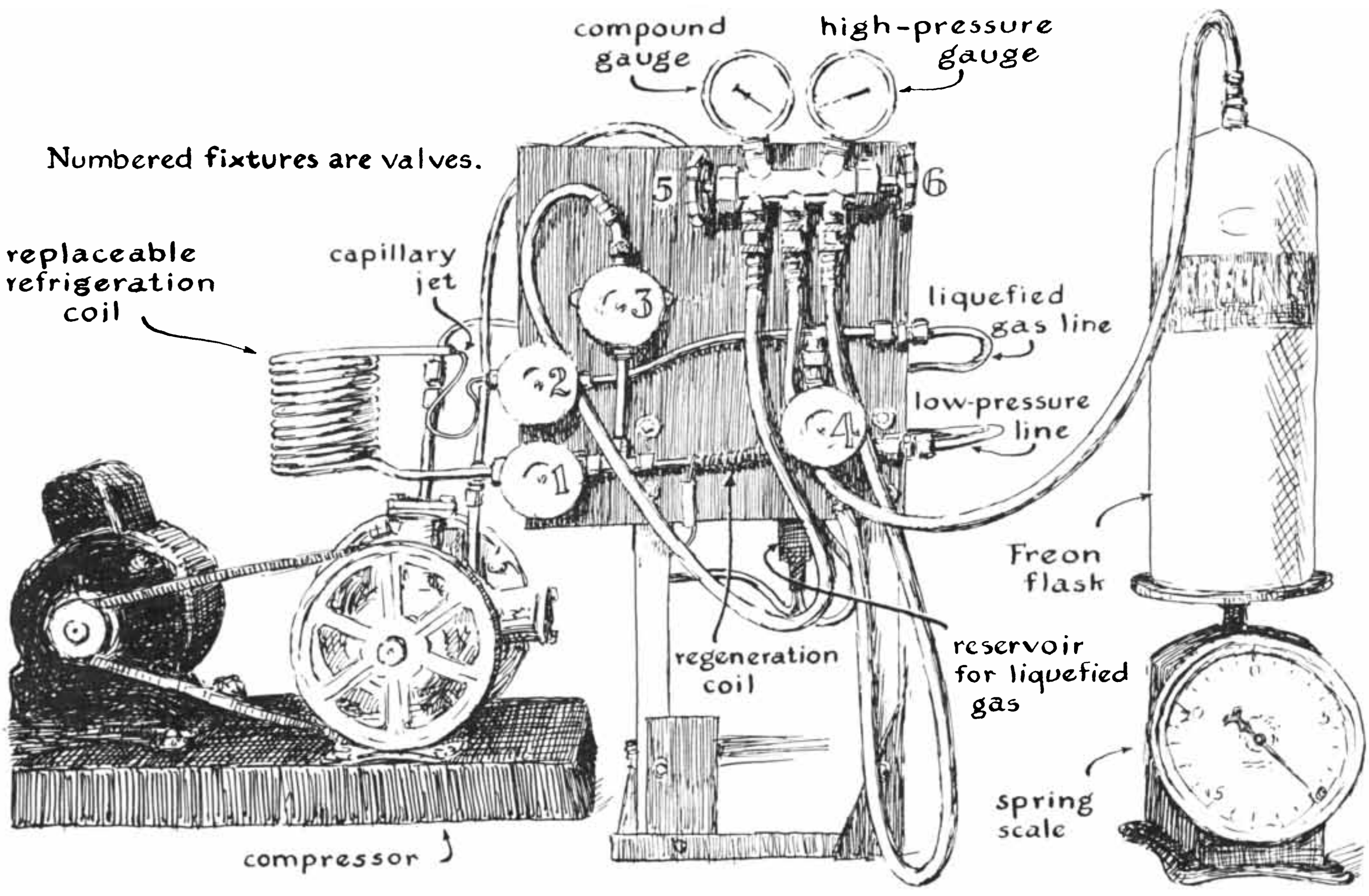
Many experiments and constructions involve the transfer of heat from one substance to another. Raising the temperature of a substance presents no problem. Portable heating devices abound in the form of gas burners, electric hot plates and so on. The removal of heat is another matter. Experimenters cool things with water, ice cubes, dry ice and occasionally with liquid nitrogen. All these substances are either inconvenient or costly.

I have long hankered for a portable cooling device that at the snap of a switch would absorb heat from any object of reasonable size or shape. Recently I built such a device from the works of an old electric refrigerator. It has turned out to be one of the handiest gadgets in my shop, and the experience of putting it together led to some interesting experiments.

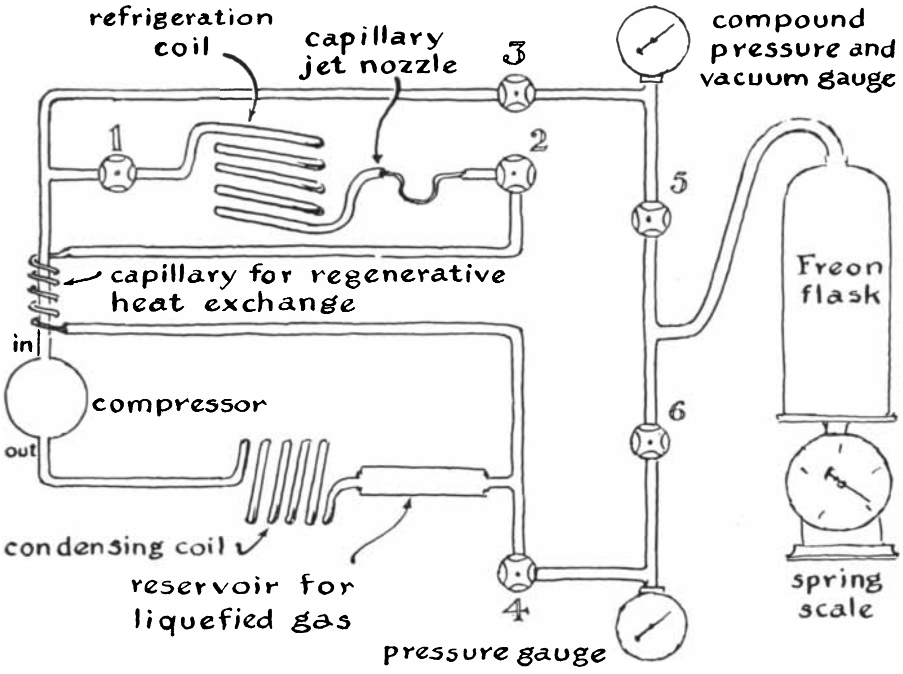
The electric refrigerator operates on a rather simple principle. A volatile fluid, the refrigerant, flows under high pressure from a flask through a nozzle into a coiled pipe, the freezing coil. The liquid absorbs heat from the coil and evaporates. The vapor is continuously carried away by a pump. Evaporation chills the coil, much as evaporating water cools the human body. The action continues as long as liquid refrigerant flows from the flask, which is known as the receiver.

Keeping the receiver filled with liquid is the job of the rest of the machine. Vapor from the freezing coil is compressed by the pump, thereby raising the vapor's temperature. The hot vapor flows through a second pipe, usually bent in the form of a grid, called the condenser. The condenser is cooled by the ambient air. Heat is transferred by the compressed vapor to the relatively cold condenser. The cooled vapor condenses into liquid on the wall of the tubing just as moisture condenses on cold objects during hot, humid weather. The condensed refrigerant trickles into the receiver for another trip through the machine.

The actions are nicely balanced. The rate at which the liquid refrigerant flows into the receiver matches the rate at which it drains from the receiver into the freezing coil, and the rate at which heat



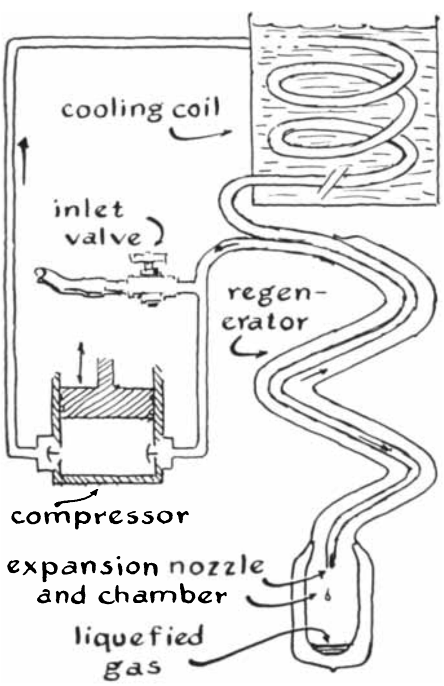
Elements of the portable cooling device



Cooling principle of the refrigerator

is absorbed by the freezing coil matches the rate at which the atmosphere absorbs heat from the condenser. In effect the machine pumps heat out of the freezing coil and into the air.

Freezing coils can be made in any reasonable size and shape from copper tubing, and they can be connected to the machine with flexible hoses. The freezing coils can be interchanged by fitting their ends with screw connections. The complete assembly need weigh no more than 60 pounds; when it is mounted on casters, it is easily portable.



Cooling principle of gas liquefier

You need not buy the mechanism. Cast-off refrigerators in good working condition are available in every city dump and often from one's neighbors. The essential mechanisms can be salvaged with ordinary hand tools. Usually the compressor and the condenser are assembled as a unit in the bottom of the cabinet. A pair of tubes connect this unit with the freezing compartment in the top of the refrigerator.

The smaller tube of the pair resembles a thick wire, but it is actually a capillary. It carries warm liquid refrigerant at high pressure from the receiver to the freezing coil. The capillary resists the flow of refrigerant: its length and bore determine the rate at which refrigerant flows through the machine. The larger tube carries cold vapor at low pressure from the freezer to the inlet of the compressor. The tubes are usually soldered together through most of their length, so that they function as a heat exchanger. Cold vapor in the low-pressure tube chills the warm liquid refrigerant in the high-pressure capillary and prevents this heat from entering the freezer—a clever trick for improving the efficiency of the machine.

To convert the machine into a portable cooler, cut the low-pressure tube and the capillary at the point where they enter the freezing coil and provide the cut ends with screw fittings for connection to freezing coils of your own design. Copper tubing in a range of diameters is available from hardware stores. Mount the compressor and condenser assembly on casters.

The job involves a few tricks, but it is not difficult. For example, the compressor and its motor are housed in a sealed tank partly filled with special refrigerator oil and with refrigerant at a pressure that varies with the temperature of the tank. If the refrigerant is Freon 12, as is usually the case, the vapor pressure will amount to about one pound per square inch per degree Fahrenheit, or roughly 75 pounds per square inch at a temperature of 75 degrees F. Freon dissolves in oil much as carbon dioxide dissolves in water. If the pressure is lowered abruptly, the oil foams for the same reason that carbonated water bubbles when a bottle is opened. If the tubes connected to the freezer are cut before the compressed vapor is released, foaming oil will squirt out and empty the compressor.

The loss of oil can be prevented by lowering the pressure slowly. This procedure requires the use of a tool known as a piercing valve, which is available from dealers in refrigeration supplies. Measure the diameter of the inlet and outlet tubes of the compressor and obtain a piercing valve that fits each pipe. Installation instructions come with the valves.

Equip the outlet of the piercing valve on the high-pressure side with a short hose and insert the open end of the hose into a quart bottle that contains an inch or so of water. Open the valve until the gas gently bubbles through the water. Pour the water out but leave the hose in the bottle to catch any oil that escapes. Gas may leak from the machine for a day or more. Do not hurry the job. When the machine is empty, measure the amount of oil that escaped. You will replace it later.

Cut the high-pressure and low-pressure tubes at the point where they enter the freezing coil. Do not use a hacksaw, because particles of copper may enter the tubes, find their way into the compressor and jam it. Cut the low-pressure tube with a pipe cutter of the wheel type. Cut the capillary by filing a deep nick and then bending the tube back and forth at the nick until it breaks. Promptly close the cut ends with adhesive tape.

Remove the compressor and condenser assembly from the cabinet and mount them on an improvised base fitted with casters. Obtain a 15-foot length of copper tubing with an outside diameter of 1/4 inch, cut a 10-foot length from it and wind the cut portion into a helical coil of convenient diameter, say six inches. Put a screw connector on one end. A short length of capillary tubing must be soldered in the other end to serve as a nozzle. I did it with a nail of slightly larger diameter than the capillary tube. I put the nail in the open end of the 1/4-inch tubing and, with a pair of pliers, squeezed the side of the tubing until the copper pinched tight around the nail. When the nail was withdrawn, the opening made a snug fit with the capillary.

Cut two inches of capillary from the machine, insert an inch of it in the pinched end of the coil and solder the two together. In the same manner solder the other end of the capillary into a three-inch length of 1/4-inch tubing and equip the other end of the 1/4-inch tubing with a screw fitting. Finally, put similar fittings on the cut ends of the low-pressure and high-pressure tubes of the machine.

The receiver of some machines is quite small and the required amount of refrigerant is correspondingly critical. I equip such machines with a second receiver. It is a six-inch length of 3/4-inch pipe closed with pipe caps. I drill the caps and thread the holes to accept tubing connectors.

The modified machine must be recharged with oil and refrigerant. Measure the amount of oil that escaped and put the same amount of fresh refrigeration oil in a jar with a wide mouth. Refrigeration oil is specially treated to remove waxes and minimize the tendency to foam when the oil is charged with refrigerant. Put a small, clean hose on the inlet tube of the compressor, which you have supplied with a screw fitting, insert the open end of the hose in the fresh oil and apply power to the motor momentarily. The oil will be sucked into the compressor.

Install the new freezing coil. Be sure to connect the nozzle end to the high-pressure tube (the one that includes the capillary). The nameplate of the machine specifies the kind and amount of refrigerant required, such as "Freon 12 #1" or "R-12 16 oz." Both indicate that the machine requires one pound of Freon 12. Certain old machines may specify sulfur dioxide or methyl chloride, and some special machines may require R-22, R-500 or R-502.

Freon 12 comes in one-pound and two-pound cans and in 12-pound flasks. It costs about $1.50 per pound. Access to canned refrigerant is had by means of a special piercing valve that costs about $4. Flasks are fitted with a Schroder valve, which is similar to the one in an automobile tire. Vapor is transferred from the containers through a high-pressure hose provided with screw fittings, one of which is designed to depress the stem of the Schroder valve.

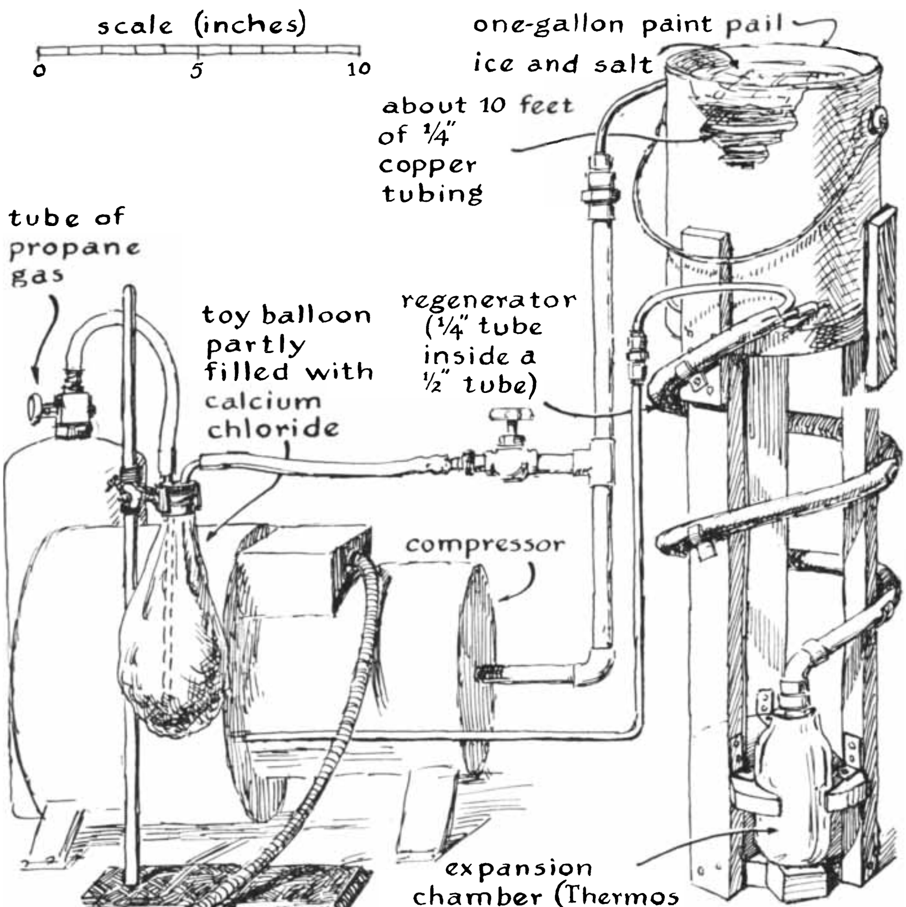
Close the piercing valve that was installed on the high-pressure side of the compressor. Attach the high-pressure hose to the piercing valve that you installed on the low-pressure side of the compressor. Connect the other end of the hose to the refrigerant dispenser. Place the can or flask, with the hose attached, on a spring balance and make a note of the gross weight. Start the compressor. Open the piercing valve on the low-pressure side, admitting refrigerant to the system. When the spring balance indicates a loss in weight of two ounces, shut off the supply of refrigerant and stop the compressor. Open the piercing valve on the high-pressure side and let the vapor escape slowly. This flushes most of the trapped air and moisture from the system.

When vapor stops leaking from the machine, close the high-pressure valve, start the machine and admit refrigerant as before. Continue until the machine has been charged with the specified amount. If all has gone well, the freezing coil will have accumulated a coating of frost. If you admit too much gas, frost will form on the low-pressure line, perhaps as far as the point at which it enters the compressor. No harm is done. Let vapor escape through the high-pressure valve until the frost line moves to within a foot or two of the freezing coil.

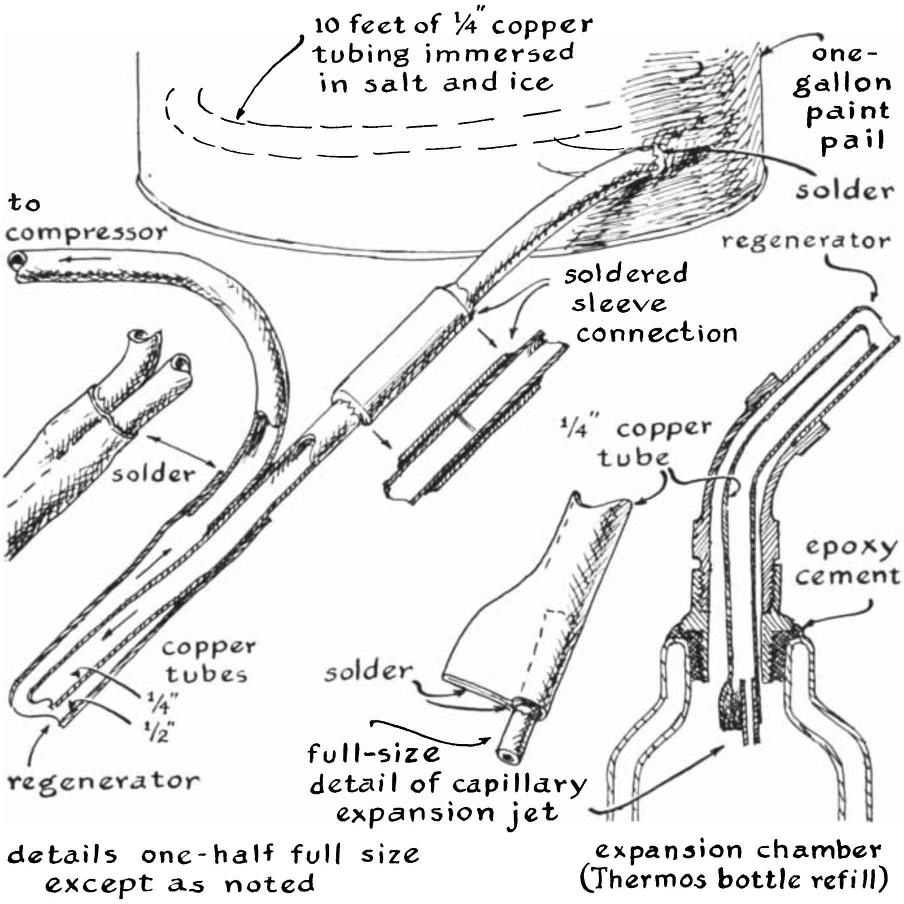
To prevent the wasteful absorption of heat from the air, it is advisable to insulate the freezing coils. For example, a convenient apparatus for chilling glass beakers can be made by winding the freezing coil on a tin can of appropriate size. Solder the coil to the can and place it in a wooden box along with a generous packing of rock wool. If you want to, you can fit the top of the box with a hinged lid. You can insert high-pressure hoses between the machine and the freezing coil. The hoses enable you to manipulate the freezers as you would a hand torch. If the hoses are more than three feet long, you may have to compensate for their volume by adding refrigerant.

The machine can be modified in various ways to make it more versatile, convenient or efficient. For instance, you can substitute 1/4-inch copper tubing for the capillary by installing a needle valve at the freezer end of the tube. The valve enables you to regulate the pressure in the freezing coil. Temperature varies with pressure. In the case of Freon 12 you can adjust the valve for any temperature down to −30 degrees F.

The same pressure-temperature rela-



Douglas Miller's apparatus for liquefying gas



Details of the regenerator and the expansion chamber

tion exists on the high-pressure side of the machine, where the temperature is determined in part by the effectiveness of the condenser. The condensers of domestic refrigerators have cooling fins and rely on air convection to remove heat. Normally they operate at roughly 90 degrees F., but the temperature of the vapor inside the pipe averages 20 to 30 degrees higher, and the corresponding pressure ranges from 135 to 155 pounds per square inch. By forcing air through the fins of the condenser with an electric fan the temperature of the vapor can be lowered 20 to 30 degrees. The pressure goes down correspondingly, thereby substantially reducing the load on the compressor. Much the same result can be achieved by connecting two condensers in series. Better yet, you can improvise and install a water-cooled condenser. The capacity of the machine can be increased by connecting two compressors in parallel. Don't overlook the fun of experimenting with refrigerants other than Freon 12.

When one makes modifications in the design of the machine, it is interesting to observe what happens inside the system as indicated by pressure measurements. Two systems of pressure measurement, known as gauge pressure and absolute pressure, are in use. For calibrating gauges zero pressure corresponds to atmospheric pressure, whereas in the absolute system zero indicates a perfect vacuum. Atmospheric pressure at sea level is assumed to be 14.7 pounds per square inch absolute and is equal to the weight of a column of mercury about 30 inches high. When air is fully exhausted from a tube that stands in a pool of mercury, the metal will actually rise to a height of 29.81 inches; in the approximate parlance of technicians a "30-inch vacuum" has been created. Subtract 14.7 from gauge pressure in pounds per square inch (abbreviated psig) to get absolute pressure (psia). The pressure-temperature relations of refrigerants are listed by reference texts in one system or the other but usually not in both.

An essential tool of the refrigeration technician is a pair of gauges that are calibrated in terms of gauge pressure and in the corresponding Fahrenheit temperature of Freon 12 and Freon 22. One gauge of the pair, used for measuring high pressures, is calibrated from 0 to 500 psig. The other, known as a compound gauge for measuring low pressures, is calibrated from 30 inches of vacuum through 0 to 80 psig. The instruments are mounted on a manifold that has a pair of two-way valves and three screw connections for three hoses. Two of the hoses provide continuous access to the gauges. By manipulating the valves either gauge or both can be connected simultaneously to the center hose. The hoses can be connected to the high-pressure and low-pressure sides of a sealed refrigerator by means of piercing valves, so that pressures can be observed while refrigerant is admitted to the machine through the third hose. Typically the reading on the high-pressure side is about 110 psig and on the low-pressure side about zero.

Valves can be installed at other points for isolating and interchanging various parts with minimum loss of refrigerant. For example, one may wish to use a needle valve to adjust the flow of liquid refrigerant to produce a certain temperature in the freezing coil and then, without substantial loss of refrigerant, substitute for the valve a capillary for automatically maintaining that temperature. I do this by first ascertaining the correct valve setting by experiment. The valve is removed without disturbing its setting and is connected to a water supply under low pressure. I count the number of drops that flow through the valve during a timed interval and then adjust the length of a capillary tube, by splicing or cutting, so that water at the same pressure flows through the capillary at the same rate.

The adjusted capillary is thoroughly dried and flushed with refrigerant before it is connected to the system. Dry Freon is relatively inert, but with water it becomes mildly corrosive. Moreover, moisture that is trapped in the system may freeze and block the flow of refrigerant.

Douglas Miller of Steubenville, Ohio, has made a device from similar parts for liquefying various gases. The basic principles of the machine are similar to those of a refrigerator, but the mechanism differs in that the gas is cooled by alternate compression and expansion until it condenses at atmospheric pressure and collects as liquid in an insulated container. Miller writes:

"My apparatus consists of a reservoir of gas to be liquefied, a deflated rubber balloon from which gas is admitted to the compressor at atmospheric pressure, a compressor capable of pumping about 900 cubic centimeters per second, a condenser cooled with a mixture of ice and salt, a heat exchanger known as the regenerator and a Dewar flask for collecting liquefied gas [*see illustration on preceding page*].

"The capacity of compressors from old refrigerators can be measured with sufficient accuracy by observing the rate at which air from the machine displaces water. I fill a gallon jug with water and invert it in a bucket containing about six inches of water. I connect a hose to the outlet of the compressor, insert the other end in the bottle and time the displacement interval with a stopwatch. A volume of one gallon is equal to 3,785 cubic centimeters.

"To make the condenser I wound a 10-foot length of 1/4-inch copper tubing into a helix that would fit inside a one-gallon paint bucket. The ends of the helix were slipped through holes near the top and bottom of the bucket and soldered.

"The regenerator consists of a pair of concentric copper tubes five feet long. Compressed gas that has been cooled by the condenser enters the center tube, which is 1/4 inch in diameter, and expands to atmospheric pressure through a nozzle at the other end. The nozzle is made of capillary tubing. The expanded gas absorbs heat from the closed Dewar flask and lowers the temperature of the flask. The cold, expanded gas returns to the compressor via the 1/2-inch outer tube of the concentric pair, where it absorbs additional heat from compressed gas then flowing to the nozzle. Thus the temperature of the compressed gas is lowered both by the condenser and by the regenerator.

"As the operation continues the Dewar flask ultimately reaches the low temperature at which the gas can exist in the liquid phase at atmospheric pressure. The regenerator will function as a straight concentric tube, but I made it in the form of a helix to conserve space. The condenser, regenerator and Dewar flask were assembled as a unit on a simple wood frame [*see top illustration on this page*].

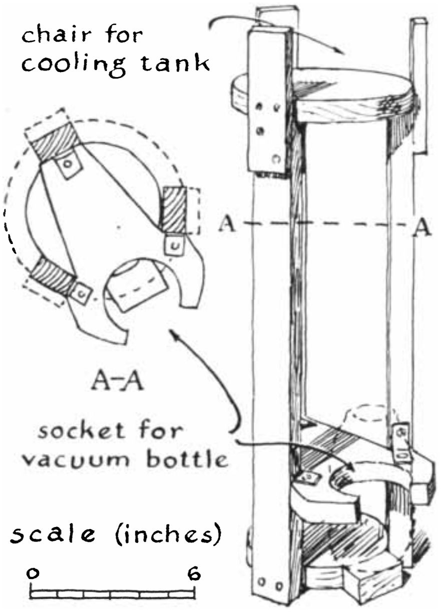
"Gas is admitted to the machine through the deflated rubber balloon, which contains a few ounces of calcium chloride to absorb water vapor. Unless the gas is dry, water vapor may freeze and clog the nozzle. The drying procedure is particularly essential in the case of gases prepared at home by the reaction of substances in solution. The balloon is closed and suspended by a rubber stopper that contains two holes. Gas enters the balloon through a plastic tube that extends to the bottom of the balloon, where it is surrounded by calcium chloride. Gas filters through the chemical and enters the system through a short tube in the second hole of the stopper. Thus the balloon performs the dual function of drying the gas and maintaining it at atmospheric pressure.

"The Dewar flask is closed by a snugly fitting stopper of plastic that I molded around the nozzle end of the regenerator. The plastic comes as a putty-like epoxy resin called 'Black Magic' that hardens in a few hours. It is available from shops that specialize in the repair of automobile bodies. I cleaned the tubing so that the material would adhere to the metal and greased the rim of the Dewar flask to release the plug after it hardened. The plug need not make an airtight seal with the flask, because the compressed gas expands to atmospheric pressure.

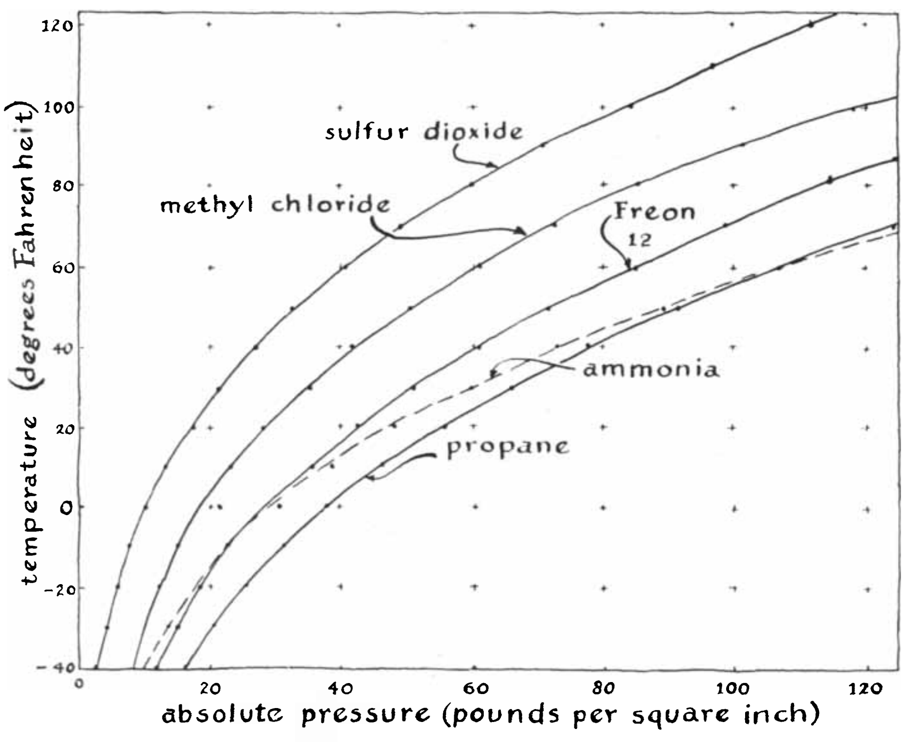
"The regenerator and condenser must be well insulated to prevent the absorption of heat from the atmosphere. I use a commercial insulation of fiber glass in strip form; it is manufactured for wrapping hot-water pipes. Suppliers of refrigeration materials stock similar insulation in the form of tape that is coated on one side with a pressure-sensitive adhesive. All tube joints should be checked for leaks before the machine is placed in operation.

"Although the system will liquefy a number of common gases, I suggest that the novice begin with one of the Freons, because these materials are odorless, nonflammable and nontoxic. Freon 12 is used in domestic refrigerators and hence is available in most communities. To operate the machine, fill the condenser with a mixture of rock salt and ice in the proportions used for freezing ice cream at home. Start the compressor and admit gas until the balloon puffs outward, but not enough to stretch the rubber. Then relax. The machine must pump heat from its various parts before it can chill the gas to the temperature at which it liquefies at atmospheric pressure. Liquid should begin to collect in the expansion flask within two hours. The balloon will begin to collapse as the gas liquefies. Admit more gas and continue. Stop the machine after three hours, remove the flask and inspect the contents.

"A variety of Freons are available for



Supporting rack for elements of liquefier



Pressure-temperature relations of five refrigerants experimentation. The starting material from which they are made is carbon tetrachloride, consisting of an atom of carbon bound to four atoms of chlorine. The structure of this molecule resembles that of methane, in which four atoms of hydrogen are bound to the atom of carbon. Carbon tetrachloride is made into Freon 12 by replacing two of the chlorine atoms with two atoms of fluorine. The chemical name of the resulting compound is dichlorodifluoromethane. Freon 11 is made by replacing only one of the chlorine atoms with fluorine. Similarly, three or four chlorine atoms can be replaced by fluorine atoms, and the resulting compounds can be mixed with still other preparations to produce vapors that liquefy at desired temperatures and absorb characteristic amounts of heat when they evaporate.

"In addition to the Freons I have liquefied propane, butane, sulfur dioxide, methyl chloride, ethyl chloride and ammonia. Propane and butane are extremely hazardous substances because they can combine with air to form a highly explosive mixture. Sulfur dioxide is a nonflammable and nonexplosive compound that, at atmospheric pressure and all temperatures above 14 degrees F., is a colorless gas with a strong, irritating odor. It corrodes copper but was a popular refrigerant in domestic refrigerators 25 years ago. If you intend to experiment with corrosive gases, substitute stainless steel for all copper parts. Methyl chloride, also formerly used in domestic refrigerators, is noncorrosive and nontoxic, although it can interfere with respiration if it is inhaled in large amounts. At atmospheric pressure it boils at −11.36 degrees F., and it can be chilled by expansion to form 'ice' at −132.7 degrees. Ethyl chloride is chemically similar but at atmospheric pressure it boils at 55.6 degrees F. Ammonia is somewhat flammable, and when it is combined with air in the correct proportion, it is explosive. It is toxic, attacks copper and is extremely soluble in water. On the other hand, ammonia absorbs a large amount of heat when it evaporates. For example, when a pound of ammonia vaporizes at five degrees F., it absorbs 565 B.T.U.'s of heat. (One B.T.U. of heat raises one pound of water one degree.) In contrast, one pound of Freon 12 that evaporates at five degrees F. absorbs only 69 B.T.U.'s. For this reason refrigeration machines that use ammonia as the refrigerant can be made substantially smaller than those that use Freon 12. Some characteristics of the gases with which I have experimented are plotted in the accompanying graph [*bottom of preceding page*].

BOOKS

Prehistoric reflections on the universe, chemistry without air, and other matters

by Philip Morrison

HAMLET'S MILL: AN ESSAY ON MYTH AND THE FRAME OF TIME, by Giorgio de Santillana and Hertha von Dechend. Gambit Incorporated ($10). In our times it is mainly amateurs who recognize particular stars. The experts have it all in books and phographic plates; the ship's captain is more likely to think about the jitter on his loran screen than about the constellations. The current revival of astrology is touching and repellent by turns. Yet from Cape Kennedy the morning papers now and again write of the bright star Canopus, somehow fixed on, and once in a while lost by, the automatic star-seekers of the latest space vehicle. Why Canopus? That star has long been the *gubernaculum*, or steering oar, of the starship Argo, just as it is marked in the first printed star map of 1603. To be sure, the designers of space vehicles may never have heard of this tradition, and it was surely not the basis of their selecting Canopus. Still, they are not entirely independent of it; their rational choice and the old tradition, which we can be sure arises in part from reason, rest on the same fundamental fact. The bright point of Canopus defines a line of sight nearly perpendicular to the plane of the ecliptic, where the lines of sight to the sun and all the planets lie. That makes the star useful for fixing an axis in space. The same fact makes Canopus (also called Ponderosus and in Arabic "the weight") the plumb bob of the sky, a fitting steersman's instrument.

The pole star is motionless to the eye of a man; it is steadfast for his lifetime. It moves nevertheless. It moves in the slow precessional motion detectable by our instruments; it moved for the classical discoverer of the precession of the equinoxes, Hipparchus, who had instruments of stone and bronze and the powerful instrument of Greek geometry. Above all, and this is the central theme of this marvelous and exasperating book, the pole and the entire sky tilt in the sight, not of any one man, but of a culture of men linked over millenniums by oral tradition. Before the invention of writing 5,000 or 6,000 years ago men were not dull or unobservant or barbarous. Genius flowered from time to time. One contemporary anthropological style, stemming from the school of Claude Lévi-Strauss, likes to think of great unlettered minds of the past expanding the logical matrices implicit in kinship schemes to construct traditional arrangements we can barely grasp today. Esoteric world systems are related in small African tribes.

Hamlet's Mill, following not a single tree to its roots but a banyan grove of correspondences, takes up a complex skein of myth. The authors, two learned historians of science, present case after case, allusion after allusion, in support of the general view that before men could write they thought deeply as they watched the order of the heavens. Hamlet is no mere orphan prince; Shakespeare had a source. That source, Saxo Grammaticus, is here, and in his tale the figure grows more heroic and more mythic. Even Saxo had his source, in oral myth. His Amleth the Dane was Icelandic Amlodhi, whose mill was the sea, grinding rock to sand. It is not even the classical tales of the Roman kings, where philologists at the turn of this century grounded the whole story, that begin it. Indeed, the steering oar and the mill and many another curious item can be traced to the Finnish epic the Kalevala, to the Persian hero Kai Khusrau and to Aryan myth. In the human warmth, glory and tragedy of all men's epics there is embedded a technical account of the "frame of time." Kai Khusrau, for example, was not only a great sovereign of Iran but also king of the zodiac, "from Pisces downward to the Bull's head."

A golden age may be a perpetual component of man's wishes, a part of his unconscious, but there was a genuinely luminous period around 6000 B.C., not on earth but in the sky. At that time the area of the sky where the plane of the ecliptic crosses the celestial equator (the equinox) was occupied both by the Milky Way and the bright stars in the belt of Orion. In *Hamlet's Mill* numbers, animal names, a "Star Menagerie" and all manner of other clues are pursued. We voyage from Kepler, Dante and Plato back to the Old Testament, the epic of Gilgamesh and the Milky Way tale of the Arawak. The book presents scores of works of art of the most puzzling kind (mighty churns of the Sea of Milk and Scorpion ladies) and almost 40 learned appendixes.

It is natural that so rich and complex a first unriddling is flawed. It is less necessary, but it is true, that the authors cannot conceal their impatience with translators and clumsy metaphysical assumptions. There are scholars who, purporting to disdain all metaphysics, hold that social evolution is simple and direct, from an unthinking barbarism through Aristotle straight through to modern times. It is hard to deny that translation from the dead tongues is an art that does not often go with a knowledge of positional astronomy. The word *Himmel* is given in the standard dictionary as the translation of no fewer than 38 distinct Egyptian hieroglyphs. That must represent a large loss of what well could be technical astronomical information. For Sumerian the case is no better. The book is polemic, even cocky; it will make a tempest in the inkpots. It nonetheless has the ring of noble metal, although it is only a bent key to the first of many gates.

"This book reflects the gradually deepening conviction that, first of all, respect is due these fathers of ours," says the preface. Plato explained in the *Phaedrus* that trust in writing was bound to discourage the use of memory, so that the invention of writing was a mixed blessing. Readers "will therefore seem to know many things, when they are for the most part ignorant and hard to get along with, since they are not wise, but only appear wise." Here is a book for the wise, however it may appear.

THE MANIPULATION OF AIR-SENSITIVE COMPOUNDS, by D. F. Shriver. McGraw-Hill Book Company ($17.50). Between the wars every experimental physics laboratory echoed with the hollow beat of the mechanical vacuum pump. The pride of the physicist was apt to be the skill of the departmental glassblower. High energy, ultrahigh vacuum, solid-state physics and the steady growth in electronic complexity has led most of physics into or beyond industrial-scale vacuum systems, and into experiments that rest on pulse-handling technique much more than on low-vapor-pressure sealing waxes. In those days the chemist was usually at work with solutions; his laboratory was a bench and sinks, a fume hood, thermometers, flames, glassware and a potentiometer.

Times have changed. The inorganic chemist and the modern preparatory chemist are likely to occupy laboratories whose bench space is dominated by a trellis of iron rods, over which grows a great vine of glass fruit in a wild variety of bulbs and pods. It is now the chemistry department that boasts of an adroit glassblower; even its graduate students can do a little of it. The beat of the pump is heard today in the chemistry laboratory.

Here is a knowing and reasoned account of how they keep the air out of chemical experiments. Fluorine and noble-gas chemistry, boron hydrides and metal hydrides, free radicals and other recent excitements of the chemist are based on study of substances so reactive that they must be made and used free of the air.

If quantitative results with small amounts are aimed at, air must be rigorously excluded. For this purpose the vacuum line is the only practical method. Most of the techniques for the conventional glass vacuum lines (fluorine chemists must often use metal systems, made of nickel or nickel alloy) appear here as modern outgrowths of the stratagems developed by Alfred E. Stock, a pioneer hydride chemist between the world wars. The method is of course to carry out all reactions and measurements in apparatus attached to a line from which all the air has been removed down to some 10^{-9} atmosphere. The substances are transferred mainly as gases, using liquid nitrogen to condense a substance here or an electrical heater to volatilize it there. Pressure measurements allow close control of amounts. Reciprocating mercury displacement pumps, versions of a 19th-century design, with alternating air pressure and vacuum to change the mercury level, serve to transfer and to meter those gases that, like hydrogen, will not condense.

Valves are a hard problem. Stopcocks, with their turning parts, need greased seals. They leak. Stock devised leakless valves that used mercury to float glass plugs into their seats. (Stock suffered from mercury poisoning!) Bellows valves sound more modern. They work, although they cannot easily be cleaned. Joints are often fitted tapers of two ground-glass parts, or ball-and-socket fits. They are tight when waxed, just as in the old days, and some good waxes are listed by trade name and source. More modern joints, which can be opened and yet resist heat and solvents as even a fluorocarbon wax cannot, are sealed by clamped O rings made of neoprene or nitrile rubber, on the idea of the red rubber ring of the Mason jar. The sizes, shapes, materials and receiving grooves of O rings form a large body of precise technical lore; much of it is given here. O rings are all but indispensable for moderate-cost metal vacuum systems. For fluorine-handling over a period of time no elastic materials at all can be used, and joints are then made with metal alone; they are sealed by high pressure applied to areas of direct metal-to-metal contact, through careful engineering design and construction of the entire heavy screw-clamped fitting.

For larger samples and less reactive material one can work not in a vacuum line but with bench-top apparatus such as flasks and test tubes, designed to keep the materials in contact with a moderately inert gas. There are several ingenious systems for this style of work. One is to use thin rubber closures to seal the test tube kept filled with nitrogen. Transfer is made by using a hypodermic syringe, purged of air in a nitrogen line, to pierce the rubber, remove the sample and inject it into another sealed tube. A variant of this scheme is the "pop bottle" technique: capping the containers to hold in a considerable excess pressure of nitrogen. The excess pressure keeps air out, and transfers are made easily with hypodermic syringes. This scheme will not work for solids without other aids.

The third system described here is the glove box. Now the entire laboratory resides in its own atmosphere, and the chemist reaches into that world for his work only through plastic gloves. Such an apparatus can be made on every scale; there is a design for a small glove box that can be pumped out to fill small test tubes with reactive solids. Bigger ones can hardly be evacuated; the load of atmospheric pressure on a big area is costly to support. There the inert gas has to occupy the volume slowly, pouring in to displace air over a long time.

The limit of the technique is the slow diffusion of air and water through the glove material around the chemist's sweating hand into the inert world within. This establishes a need for constant repurification of the atmosphere within, at a rate determined by the reactivity of the materials. Big, expensive laboratories for specialized purposes, such as the chemistry and machining of plutonium, represent the end point of this technique.

There are appendixes on hazards (plenty of them), on glassblowing and on materials, both metals and plastics. You can learn how to weld polyethylene, using nitrogen gas fed to the seam over an electric heating coil and a welding rod of the same plastic. (Air won't work—the surfaces oxidize.) The book should be of real value to the chemists it is aimed at, and a pleasure to examine for anyone who has a sense of nostalgia for the diverse ingenuities of the experimental laboratory, before automation seized control.

Music by Computers, edited by Heinz Von Foerster and James W. Beauchamp. John Wiley and Sons, Inc. ($14.95). Cybernetic Serendipity: The Computer and the Arts, edited by Jasia Reichardt. Frederick A. Praeger, Publishers ($8.95). These are two candid books, one a symposium report on music and the other an enlarged exhibition catalogue. The music making is discussed and explained; moreover, the music itself is presented to the ears, on four small discs fitted into a pocket. The eye is offered reproductions (small ones, to be sure) of many genres of computer visual art. The reader may therefore judge and not merely imagine.

Computers can play chess these days. They are indifferently skillful at it; not being grand masters, they aspire to reach the semifinals of a statewide tournament. The combinatorial richness of a chess game seems comparable to that of a haiku; one might expect that the artistic effort of computers (or rather those humans who contrive their programs) might achieve the level, say, of good grade-school writers of haiku. That seems no bad critical assessment of the work presented here, if one may speak by metaphor across many mediums.

There are often two distinct roles in art; in music it is quite clear nowadays that they are distinct. We seldom expect composer and performer to be one. Visual arts differ, yet the roles are there. In poetry, however, long since digitized and impersonally presented, we find only one role: poet. The computer is now a fair, if slow, performer on some musical instruments. Asked to synthesize their sounds, it (and Professor Beauchamp at Urbana) can do rather well. One disc proudly displays the computer at work, being by turn oboe, flute and cornet. (Drums, no.) Its input gear listens to a performer, samples a second of tone 30,000 times and records the amplitude of each sample. These numbers are Fourier-analyzed by an approximate routine that imagines that the separate harmonic components vary slowly compared with audio rates. It follows these harmonic strengths, against time, for some dozens of harmonics, and does various smoothings and interpolations to cover up its finite grasp. Then it will recompute the numbers on request, to be redone into voltages for taped sound.

Most computer sounds are far less conventional. There is an entire family of programs around an original version called MUSIC4, which stores one cycle each of many wave forms, each expressed by harmonic analysis. The composer calls up a fundamental frequency; the program gives him repeated strings of the cycle, computed for that pitch, and modulated with another signal he can call on for changing harmonics with time. He can choose noninteger harmonic mixes to yield novel sounds. There are ways of reducing the regularities with judicious noise, producing tremolos, vibrato, random errors in pitch and even a harmonic content varying nonlinearly with loudness, all attributes of human performers. The range of timbre, loudness and attack is wider in this mode; there are drums too! Over it all is our preset experience; one composer-programmer remarks rather wistfully that some of his best-planned effects elicit quiet asides from listening experts, who helpfully point out to him that his amplifiers are clipping badly. The computers are good performers, playing, as they should, unknown instruments with interesting but somehow disturbing sounds.

The computer also composes. One of the discs presents a piece of an elaborate composition in progress, scored mainly for live string players. Professor Hiller has made a set of stochastic routines, which choose what to do next by random means from a subtly conditioned set of possibilities: "A choice dependent on the difference between the differences between the differences... between the previous five choices." With computers composers gain power, whereas performers lose it.

In the visual arts the computer is a tyro at performance. About all it can do is draw with a bland, even line or to print repeatedly from type over a carbon ribbon. It can also make a television image. Super-Lissajous figures are the genre there. (Back to art school, computer. Rembrandt had wide-band peripherals, not to mention any African carver.) Its strength is in composition. Here it makes its best efforts, for example drawing line segments of random length at random positions that reassemble in a few steps into the drawing of a hummingbird they set out to be. Drawings of this kind make an attractive set. In literature, the dance, graphics and film animation what the computer can do is make various fixed transformations on random sequences, following the spirit of the aphorism Schönberg derived from Beethoven: "Everything in music must be at once surprising and expected." That tension is certainly present in the computer treatment of all the arts. It is, however, too thin and too taut. The quality remains about an order of magnitude below serious notice. The work is always cold. The greatest success so far appears to lie in the analogue closed-loop "sculpture," such as John Lifton's projection dome. Music and viewers' speech within the dome caused a "light show" to vary responsively around the onlooker. There is no photograph of the result, but the comment by the viewers was encouraging.

The computer has grown up able to extrapolate Newtonian orbits and to withhold the right amounts from daily payrolls. High art is demonstrably still beyond it. One should report that all the musicians cited here are serious and professional; the visual artists are not at all like that. There are one or two howlers, and quite a few put-ons.

WILDLIFE IN DANGER, by James Fisher, Noel Simon, Jack Vincent *et al.* The Viking Press ($12.95). The Survival Service Commission of the International Union for Conservation of Nature and Natural Resources distributes worldwide a loose-leaf Red Data Book, whose changing data sheets list and describe those species of animals and plants anywhere that are in danger of extinction. Here we have a popularized version of the Red Data Book, as of late 1966, with some species reports updated into 1968. Its British designers have made a handsome hard-cover book, with very attractive illustrations of most of the organisms, about half of them reproductions of paintings in color. They have not, however, made a book that is easy to use; only the long and good index provides access to the entry for a species.

Most of the known extinguished species, like those listed as being currently in peril, are large organisms, mainly mammals and birds. All species, like their individual members, are mortal; the issue is the present high rate of species death. Since 1600, a reasonable base line, the statistical mean life of bird and mammal species has dropped from hundreds of thousands of years to about 10,000 or 15,000 years. Our hunter ancestors, particularly the men who came across the Bering Sea with their "sophisticated flints," were formidable too; they may in their day have cut the life-span of American species by a factor of 10. Man has been the most skilled of predators for a long time. Hunting remains the main cause of danger to species, although disruption of habitat, with agricultural and industrial man competing with other species, is a close and gaining second.

Hunters are not species-killers by necessity or out of self-interest. Controls, seasons and royal preserves have made successful conservationists of them not only in modern times but also traditionally. The sudden breakdown of such old social arrangements is one of the greatest dangers. The white rhinoceros of the Congo was decimated after the sudden occupation of its park reserve in 1963 by a heavily armed force of rebel soldiery. The courageous Philippine dwarf buffalo, the tamarau, has been severely endangered of late by cattlemen who fly into its range preserve and illegally hunt it, often using spotlights and automatic weapons. "Since [World War II] large quantities of modern firearms have come into general circulation" in the Philippines. Elsewhere the decline of the sole surviving species of wild horse dates "from the acquisition of modern firearms by Chinese and Mongolian hunters."

There is another side. The polar bear is hunted these days from planes. All five Arctic nations, however, have now joined in a circumpolar scheme to regulate the hunt. The California sea otter, strictly protected, has come back from assumed extinction. Przewalski's horse and the European bison have been saved by their rearing in zoos, and special preserves for free-living herds of these beauties are newly established or soon expected. The bird-plumage trade has been stopped, but only just in time. The ivory-billed woodpecker of our deep South, each pair of which needs a few square miles of mature forest with large oak and gum trees, already seemed doomed by relentless logging a generation ago. After a decade without one sighting, in 1967 some 10 pairs were found in Big Thicket, an East Texas pine

By arrangement with The Nobel Foundation . . .

THE COMPLETE COLLECTED

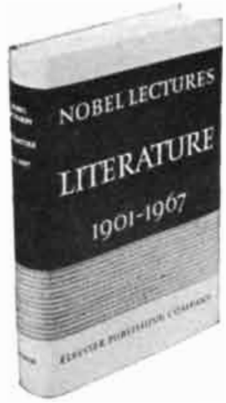
NOBEL PRIZE LECTURES

AVAILABLE FOR THE FIRST TIME IN THE ENGLISH LANGUAGE
ARRANGED IN CHRONOLOGICAL ORDER . . . ACCORDING TO PRIZE

NOW AVAILABLE:

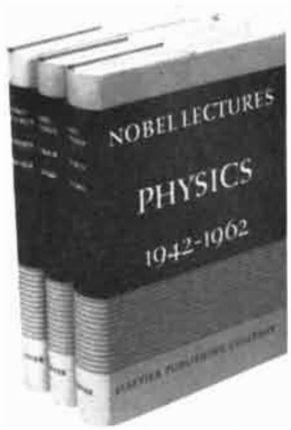
NOBEL LECTURES IN LITERATURE

1901-1967 EDITED BY HORST FRENZ, 652 pages. $14.50

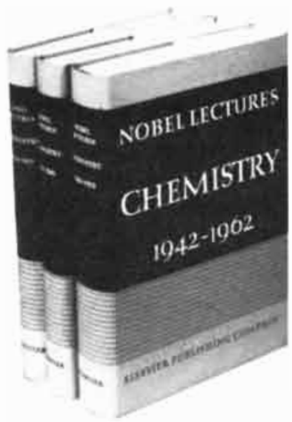


A treasury of unforgettable highlights in twentieth century literature. Now, for the first time in one volume, complete authoritative details of all of the Nobel Literature Prize awards from their inception in 1901 through 1967. You are there to witness the joy, reactions, excitement through the presentation speeches, acceptance speeches, banquet remarks, the winners' lectures—even the remarks of those who rejected the prizes. Complete with bibliographical and biographical details. A memorable work for pleasurable reading as well as reference . . . *"This is a book that should exclude all other reading for a few days."—Washington Evening Star.*

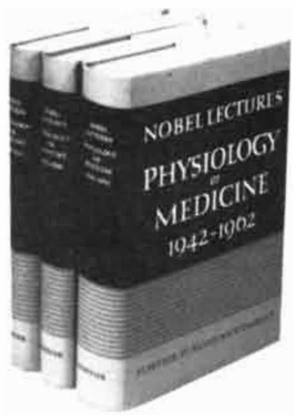
THESE THREE SERIES ALSO AVAILABLE:



NOBEL LECTURES IN PHYSICS. *Edited by the Nobel Foundation.* Three volumes: 1901–1921, 1922–1941, 1942–1962. *Approximately 1500 pages.* $85.00



NOBEL LECTURES IN CHEMISTRY. *Edited by the Nobel Foundation.* Three volumes: 1901–1921, 1922–1941, 1942–1962. *Approximately 1600 pages.* $85.00



NOBEL LECTURES IN PHYSIOLOGY/MEDICINE. *Edited by the Nobel Foundation.* Three volumes: 1901–1921, 1922–1941, 1942–1962. *Approximately 1800 pages.* $85.00

DUE IN 1970: NOBEL PRIZE LECTURES IN PEACE, 1901-1967

Order through your local bookseller . . . or direct from the publisher

AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE NEW YORK, NEW YORK 10017

swamp. The bird "survives by the skin of its ivory bill." The whales may yet survive by their very martyrdom: they may become too scarce to hunt!

Once man himself is not endangered by overkill weaponry, disorder, short-sighted profit taking and outworn fashions and superstitions, the other organisms that share this world are likely to pull through. It is pretty clear that an order that is good for man will turn out to be good for most, but not all, species of his fellow travelers.

ENGINEERS AND ENGINEERING IN THE RENAISSANCE, by William Barclay Parsons. The M.I.T. Press ($12.50). The posthumous work of a distinguished practicing engineer, this volume first appeared in 1939. It is reprinted almost unchanged (an orotund introduction by Nicholas Murray Butler has been judiciously omitted, along with a couple of chapters on general history) with its rich store of original works quoted at length and its splendid illustrations. The nearest rival the book has is the appropriate volume from the great five-volume history of technology by Charles J. Singer; Parsons is the closer to the primary sources. The topics include Da Vinci, machines and mechanism (not clocks), mining, the engineering of cities, of rivers, canals and harbors, of bridges and domes. Here is the famous account of how Domenico Fontana moved the obelisk to St. Peter's, and the story of the design and construction of the Rialto Bridge in Venice, the Pont Neuf over the Seine, the Santa Trinità across the Arno, the great domes of Florence and of Rome. Tourists with the engineer's eye will find this book a fine preparation for a summer in Europe.

LUCRETIUS: THE WAY THINGS ARE. THE DE RERUM NATURA OF TITUS LUCRETIUS CARUS, translated by Rolfe Humphries. Indiana University Press ($6.50). The convinced, urgent, skeptical long argument of a poem by the Roman Epicurean is put into clear metrical verse by a latter-day poet. The title itself shows the aptness of the translation: "I try to learn about the way things are/And set my findings down in Latin verse." Lucretius could be absurd, led by his para-religious fervors, but he repays study almost everywhere, particularly for those who, unlike Mr. Humphries, claim expertise in the modern outcome of the world view set going by old Democritus, "that hero of the mind." The text followed is that of the Loeb Classical Library; there are helpful—often acute—notes and a critical introductory essay.

INDEX OF ADVERTISERS

NOVEMBER 1969

BIBLIOGRAPHY

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

ACUTE RESPIRATORY FAILURE

Physiology of Respiration: An Introductory Text. Julius H. Comroe, Jr. Year Book Medical Publishers Incorporated, 1965.

Respiratory Care. H. H. Bendixen, L. D. Egbert, J. Hedley-Whyte, M. B. Laver and H. Pontoppidan. The C. V. Mosby Company, 1965.

Oxygen Therapy. J. Hedley-Whyte and P. M. Winter in *Clinical Pharmacology and Therapeutics,* Vol. 8, No. 5, pages 696–737; September–October, 1967.

AMORPHOUS-SEMICONDUCTOR SWITCHING

Modern Aspects of the Vitreous State: Vol. III. Edited by J. D. Mackenzie. Butterworths, 1964.

Quantum Electron Theory of Amorphous Conductors. A. I. Gubanov. Consultants Bureau, 1965.

Electrical Conduction in Thin Insulating Films. J. G. Simmons in *Endeavour,* Vol. 27, No. 102, pages 138–143; September, 1968.

Reversible Electrical Switching Phenomena in Disordered Structures. S. R. Ovshinsky in *Physical Review Letters,* Vol. 21, No. 20, pages 1450–1453; November 11, 1968.

Conduction and Switching in Noncrystalline Materials. N. F. Mott in *Contemporary Physics,* Vol. 10, No. 2, pages 125–138; March, 1969.

Electrical Switching Phenomena in Transition Metal Glasses under the Influence of High Electric Fields. C. F. Drake, I. F. Scanlan and A. Engel in *Physica Status Solidi,* Vol. 32, No. 1, pages 193–208; March 1, 1969.

A Simple Band Model for Amorphous Semiconducting Alloys. Morrel H. Cohen, H. Fritzsche and S. R. Ovshinsky in *Physical Review Letters,* Vol. 22, No. 20, pages 1065–1068; May 19, 1969.

EARLY MAN IN THE WEST INDIES

An Archeological Chronology of Venezuela. J. M. Cruxent and Irving Rouse in *Social Science Monographs, VI: Vols. I and II.* Pan American Union, 1958, 1959.

The Entry of Man into the West Indies. Irving Rouse in *Yale University Publications in Anthropology,* No. 61; 1960.

Prehistory of the West Indies. Irving Rouse in *Science,* Vol. 144, No. 3618, pages 499–513; May 1, 1964.

Paleohunters in America: Origins and Diffusion. Hansjürgen Müller-Beck in *Science,* Vol. 152, No. 3726, pages 1191–1210; May 27, 1966.

MAGNETIC RECORDING

The Physics of Magnetic Recording. C. D. Mee. John Wiley & Sons, Inc., 1964.

IEEE Transactions on Magnetics. Vol. MAG-4, No. 3, pages 269–308; September, 1968.

Reviewing Slow Motion Disc Principles. Lee Stratton in *Broadcast Engineering,* Vol. 11, No. 2, pages 14–18; February, 1969.

Design of a General Videofile Information System. Michael O. Felix in *Computer Design,* Vol. 8, No. 9, pages 58–62; September, 1969.

NON-EUCLIDEAN GEOMETRY BEFORE EUCLID

Non-Euclidean Geometry: A Critical and Historical Study of Its Development. Roberto Bonola. The Open Court Publishing Company, 1912.

A History of Greek Mathematics: Vols. I and II. Sir Thomas Heath. Oxford University Press, 1960.

Das Parallelenproblem im Corpus Aristotelicum. Imre Tóth in *Archive for History of Exact Sciences,* Vol. 3, No. 4–5, pages 249–422; March 17, 1967.

THE ORIGIN OF THE OCEANIC RIDGES

The Floors of the Oceans, I: The North Atlantic. Bruce C. Heezen, Marie Tharp and Maurice Ewing. The Geological Society of America Special Paper 65, 1959.

History of Ocean Basins. H. H. Hess in *Petrologic Studies: A Volume in Honor of A. F. Buddington,* edited by A. E. J. Engel, Harold L. James and B. F. Leonard. The Geological Society of America, 1962.

Magnetic Anomalies over Oceanic Ridges. F. J. Vine and D. H. Matthews in *Nature,* Vol. 199, No. 4897, pages 947–949; September 7, 1963.

Marine Geology of the Pacific. H. W. Menard. McGraw-Hill Book Company, 1964.

Principles of Physical Geology: New and Fully Revised Edition. Arthur Holmes. Thomas Nelson and Sons Ltd, 1965.

THE RECEPTOR SITE FOR A BACTERIAL VIRUS

Studies on the Chemical Basis of the Phage Conversion of O-Antigens in the E-Group Salmonellae. P. W. Robbins and T. Uchida in *Biochemistry,* Vol. I, No. 2, pages 323–335; March, 1962.

Mechanism of ε^{15} Conversion Studied with a Bacterial Mutant. Richard Losick and P. W. Robbins in *Journal of Molecular Biology,* Vol. 30, No. 3, pages 445–455; December 28, 1967.

Mechanism of ε^{15} Conversion Studied with Bacteriophage Mutants. D. Bray and P. W. Robbins in *Journal of Molecular Biology,* Vol. 30, No. 3, pages 457–475; December 28, 1967.

Biosynthesis of Cell Wall Lipopolysaccharide in Gram-Negative Enteric Bacteria. Hiroshi Nikaido in *Advances in Enzymology and Related Areas of Molecular Biology: Vol. XXXI,* edited by F. F. Nord. Interscience Publishers, 1968.

Isolation of a Trypsin-sensitive Inhibitor of O-Antigen Synthesis Involved in Lysogenic Conversion by Bacteriophage ε^{15}. Richard Losick in *Journal of Molecular Biology,* Vol. 42, No. 2, pages 237–246; June 14, 1969.

HOW BIRDS SING

Bird Song: Acoustics and Physiology. Crawford H. Greenewalt. Smithsonian Institution Press, 1968.

MATHEMATICAL GAMES

A History of Board-Games Other than Chess. H. J. R. Murray. Oxford University Press, 1952.

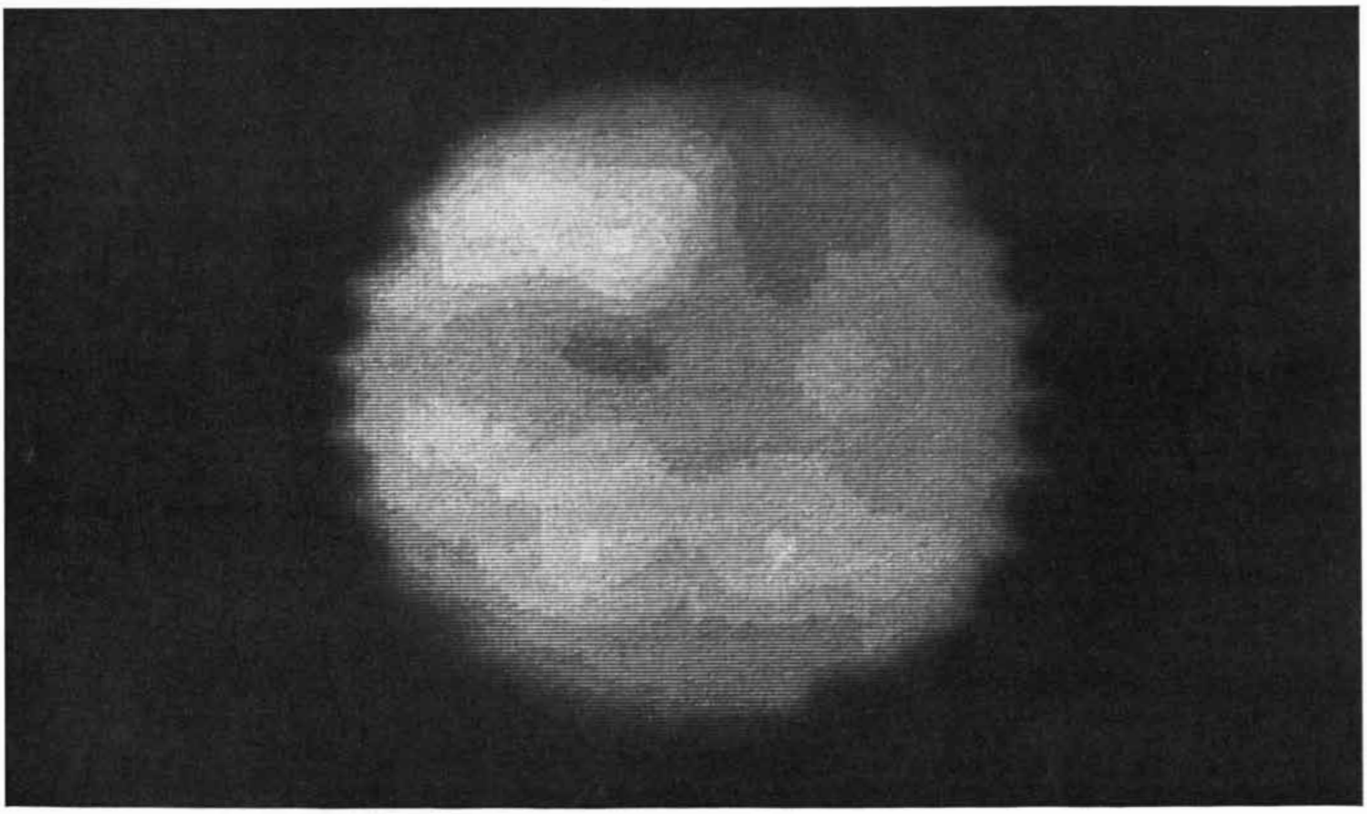
Board and Table Games from Many Civilizations. R. C. Bell. Oxford University Press, 1960.

Board and Table Games 2 from Many Civilizations. R. C. Bell. Oxford University Press, 1969.

THE AMATEUR SCIENTIST

Refrigeration Engineering. H. J. Macintire and F. W. Hutchinson. John Wiley & Sons, Inc., 1950.

Principles of Refrigeration. Roy J. Dossat. John Wiley & Sons, Inc., 1961.



CalComp puts the sun out in 3.1 minutes flat.

It's not doomsday, exactly, but it's almost as weird.

What you're looking at is a *drawing* of what a radio telescope sees when it looks at the sun.

A computer-controlled CalComp 835 plotter drew it on microfilm for The Aerospace Corporation in Los Angeles.

The people at Aerospace need a large quantity of sun complexion maps like this one for their study of the sun's radio burst mechanism.

And by using our 835 plotter, they get them fast.

In fact, the 835 can draw at a rate of 100,000 plot increments a second. Because it draws with modulated light. On microfilm.

(With a choice of five fully automatic microfilm camera options.)

What's more, CalComp microfilm drawings can be viewed separately, like ink and paper drawings.

Or in sequence for time-lapse motion picture studies.

CalComp is the leader in computer graphics.

With sales, service and software in 34 cities. And the greatest variety of plotters in the world.

Including compact drum plotters. Giant flatbed plotters. And in our 800 series, plotters that draw not only on microfilm, but on microfilm *and* paper.

So if you have a computer that can't draw, call your nearest CalComp man today.

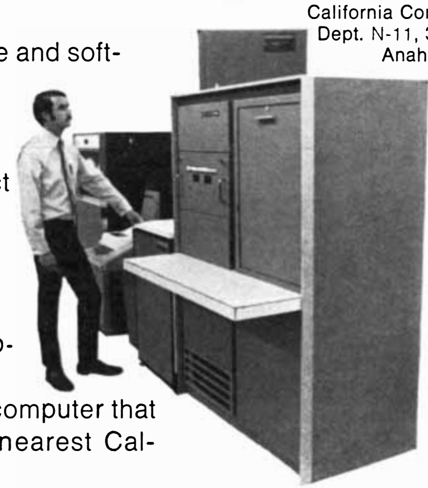
He can help you with just about any graphic problem under the sun.

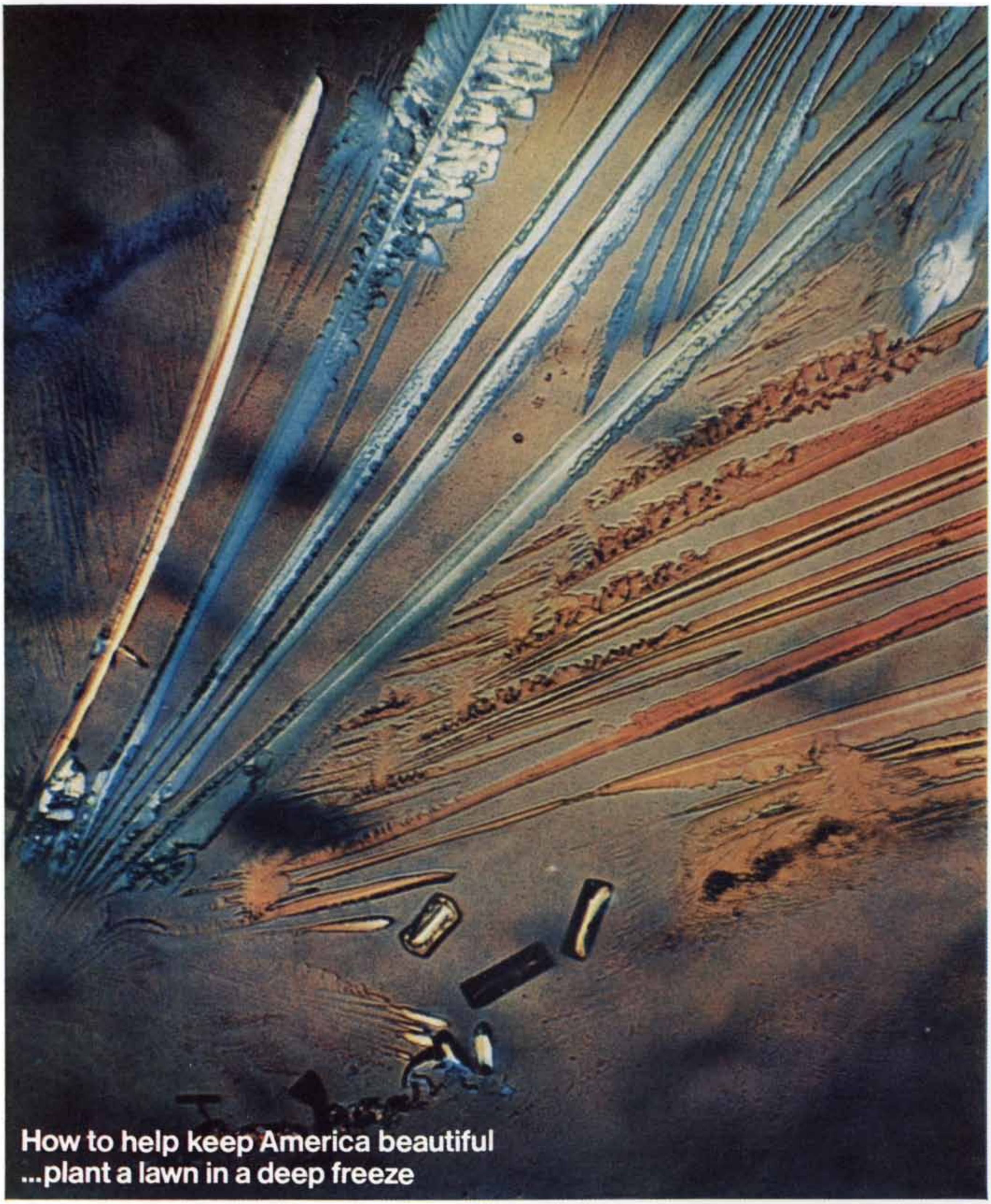
Including the sun.

CALCOMP

TEACH YOUR COMPUTER TO DRAW.

California Computer Products, Inc.
Dept. N-11, 305 N. Muller Avenue
Anaheim, California 92803





The North Alaskan Tundra is topped by a fragile mat of green growth. If you break it, heat from the never-fading Midnight Sun slowly turns the frozen mud into a swamp. Water runs. And irreparable erosion could occur. For that land has little ability to regrow its ground cover. This year, ARCO began an experiment to allow the Tundra to heal itself. Thirteen strains of winter-hardy grass have been seeded into a 20-acre Arctic lawn. And a special fertilizer, developed and made by an ARCO plant, will help it grow - - all so that the tracks of man, as he searches and develops the riches of the wilds, will not mar or destroy as well. It's just one of the many steps ARCO is taking to see to it that the world we live in is just a little bit better than when we started.

the new **AtlanticRichfield**Company **ARCO**