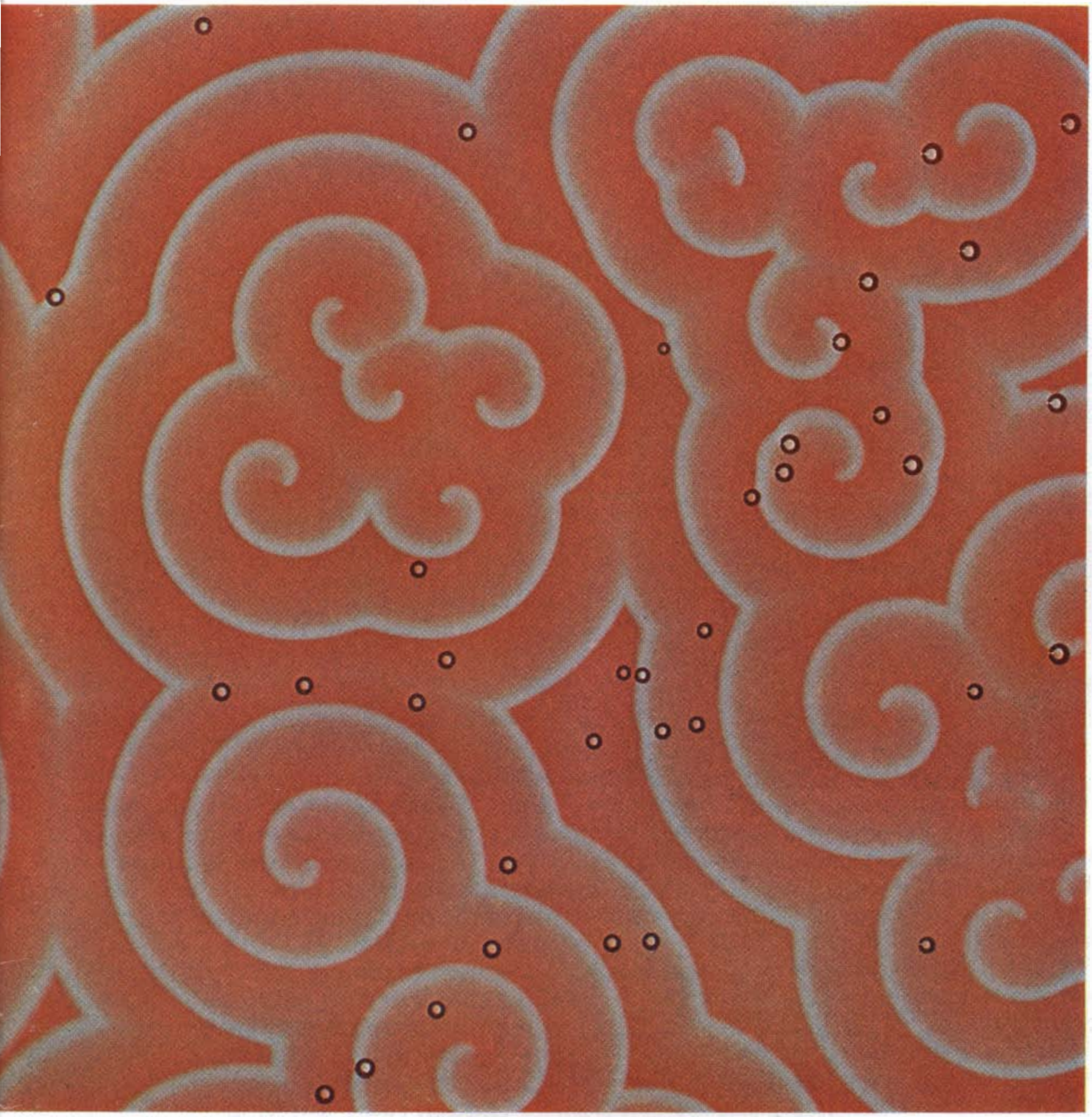
SCIENTIFIC AMERICAN



ROTATING CHEMICAL REACTIONS

ONE DOLLAR

June 1974

The Dodge Colt.

For a little car, it's a lot of car.

Sure, all Dodge Colts are economical to drive. Our 4-cylinder engine sees to that.

In fact, in a mileage test conducted and sanctioned by the United States Auto Club (USAC) in January, Colt got substantially better mileage than either Vega or Mustang II.

If that's all there were to Colt, it would seem like plenty. But Colt gives you much more.

For example, Dodge Colts have standard features such as front disc brakes, four-on-the-floor, and an adjustable steering column. Plus a few other things listed below.

The '74 Dodge Colt. As you can see, for a little car, it is a lot of car.

Here're just a few of the features standard on Dodge Colts.

- **Thrifty four-cylinder engine.**
- **Front disc brakes.**
- **Four-speed transmission.**
- **Adjustable steering column.**
- **Flow-through ventilation.**
- **Reclining bucket seats (except coupe).**
- **Hidden radio antenna (except wagons).**

Dodge

Dodge Trucks

CHRYSLER
MOTORS CORPORATION

Henry Lewis has the knack of surrounding himself with the best.

Like his New Jersey Symphony. Birgit Nilsson. Marilyn Horne. Luciano Pavarotti. Pioneer.

What with recording sessions and concerts you'd imagine conductor Henry Lewis might want to take a vacation from music. Not so. In between he catches up with his own listening on his Pioneer high fidelity system.

Henry Lewis takes his music seriously. But he also likes to have fun with it. His Pioneer TX-9100 stereo tuner and SA-9100 integrated stereo amplifier enable him to continue his musical adventures at home.

Pioneer is renowned for producing the finest tuners made. The TX-9100 substantiates this with remarkable specifications like 90dB selectivity, 1.5uV FM sensitivity and a 1dB capture ratio. Audio magazine summed it up perfectly with, "You can't buy better audible performance than is achievable with Pioneer's TX-9100 at *any price*."

Only Pioneer's SA-9100 integrated amplifier could possibly

match the performance of this excellent tuner. As High Fidelity magazine put it, ". . . It's performance is so exceptional, and the so many extras in the way of switching options, and so on, so eminently useful, that we find it the most exciting piece of audio hardware." Small wonder with features like the unique twin stepped tone controls that custom tailor listening to an enormous variety of tonal variations. And a power output of 60+60 watts RMS (8 ohms, both channels driven) with an amazingly low 0.1% distortion from 20 Hz to 20,000 Hz.

Pioneer's TX-9100 and SA-9100 assure conductor Henry Lewis the same absolute control of his music at home as he can achieve with his baton on the podium.

TX-9100 - $349.95;
SA-9100 - $449.95, including walnut cabinets.

U.S. Pioneer Electronics Corp., 75 Oxford Drive, Moonachie, New Jersey 07074 / West: 13300 S. Estrella, Los Angeles 90248 / Midwest: 1500 Greenleaf, Elk Grove Village, Ill. 60007 / Canada: S. H. Parker Co.

PIONEER
when you want something better

Our Green Lite℠* Service satisfied the world's toughest critic.

Not even Don Rickles can find fault with our Green Lite* Service.

It means your rental agreement is filled out and waiting when you pick up the car. All you do is show your driver's license, authorized credit card and sign.

If you're in a hurry, when you return the car, you don't have to wait in line. Charge customers can take advantage of our express check-in box.†

We also feature GM cars, and give S&H Green Stamps on all U.S. rentals. And at most locations you get two extra hours free with our 26 hour check-in day.

Sign up for National's Green Lite Service. It's so good, we left Rickles speechless.

Fill out questions 1 through 20 providing all information requested (please print).
Mail to National Car Rental, 5501 Green Valley Drive, Minneapolis, Minnesota 55437.

NATIONAL CAR RENTAL

When it comes to service we mean business.

1 FIRST NAME AND INITIAL

2 LAST NAME

3 HOME ADDRESS

4 CITY (AND PROVINCE)

5 STATE (OR COUNTRY) 6 ZIP

7 COMPANY NAME

8 COMPANY ADDRESS

9 CITY (AND PROVINCE)

10 STATE (OR COUNTRY) 11 ZIP

12 IF WE BILL YOU DIRECTLY WHERE SHOULD WE SEND THE BILL? ☐ HOME ☐ OFFICE

13 DRIVERS LICENSE NUMBER

14 ISSUING STATE

15 CAR PREFERENCE
- ☐ LUXURY (CADILLAC OR EQUIVALENT)
- ☐ STANDARD (IMPALA OR EQUIVALENT)
- ☐ COMPACT (NOVA OR EQUIVALENT)

16 DO YOU NORMALLY PARTICIPATE IN COLLISION DAMAGE WAIVER? YES ☐ NO ☐

17 DO YOU NORMALLY PARTICIPATE IN PERSONAL ACCIDENT INSURANCE? YES ☐ NO ☐

PLEASE INDICATE BELOW WHICH CREDIT CARD HONORED BY NATIONAL YOU PREFER TO USE. SELECT ONLY ONE AND INDICATE CARD NUMBER.

NUMBER (INCLUDE ALL LETTERS AND NUMBERS)

19 NATIONAL CAR RENTAL CARD NUMBER

20 AMERICAN EXPRESS [AX] ______

CARTE BLANCHE [BL] ______

DINERS CLUB / AMERICAN TORCH CLUB [DC] ______

BANKAMERICARD [BA] ______

MASTER CHARGE [MC] ______

TWA GETAWAY [TG] ______

AIR TRAVEL [AT] ______

TILDEN/HERTZ/AVIS [THA] ______

OTHER [OT] Name ______

Number ______

DATE ______

SIGNATURE ______

6-SA

†Available at participating locations. © National Car Rental System, Inc. 1974 (in Canada it's Tilden Rent-a-car; in Europe, Africa and the Middle East it's Europcar.) *Green Lite and the design of the semaphore are service marks of National Car Rental System, Inc.

Established 1845 SCIENTIFIC AMERICAN June 1974 Volume 230 Number 6

ARTICLES

18 PSYCHIATRISTS AND THE ADVERSARY PROCESS, by David L. Bazelon The facts of an expert witness should be subject to cross-examination.

24 FUSION POWER BY LASER IMPLOSION, by John L. Emmett, John Nuckolls and Lowell Wood Implosion is necessary if the fuel is to ignite.

38 POPULATION CYCLES IN RODENTS, by Judith H. Myers and Charles J. Krebs They seem to result from changes in a population's genetic makeup.

58 NEUROTRANSMITTERS, by Julius Axelrod Nerve cells communicate by means of substances released by nerve-fiber endings.

72 THE STEADY STATE OF THE EARTH'S CRUST, ATMOSPHERE AND OCEANS, by Raymond Siever Their geochemical cycles are in balance.

82 ROTATING CHEMICAL REACTIONS, by Arthur T. Winfree Certain colorful reactions organize themselves in space and time as spiral waves.

96 ICE-AGE HUNTERS OF THE UKRAINE, by Richard G. Klein In a rigorous Pleistocene climate these men built shelters out of mammoth bones.

106 WINES, GRAPE VINES AND CLIMATE, by Philip Wagner Grapes of the more idiosyncratic wines grow under the more extreme conditions.

DEPARTMENTS

8 LETTERS

12 50 AND 100 YEARS AGO

16 THE AUTHORS

48 SCIENCE AND THE CITIZEN

116 MATHEMATICAL GAMES

122 THE AMATEUR SCIENTIST

129 BOOKS

136 BIBLIOGRAPHY

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $12 PER YEAR, U.S., POSSESSIONS AND CANADA; $15 PER YEAR, ALL OTHER COUNTRIES. EUROPEAN SUBSCRIPTION COPIES MAILED C/O PUBLICATIONS DISTRIBUTION SERVICE, ST. KATELIJNEVEST 14, B2000 ANTWERP, BELGIË.

THE COVER

The photograph on the cover shows the surface of a shallow liquid in which a sequence of chemical reactions is proceeding in an array of spiral patterns (see "Rotating Chemical Reactions," by Arthur T. Winfree, page 82). The liquid, which has been named the Z reagent, is a mixture of bromine compounds, malonic acid and the indicator dye iron phenanthroline. The Z reagent is mostly red; blue waves of chemical activity propagate through it at a rate of several millimeters per minute, forming spirals or rings. Where two of the blue waves meet they annihilate each other. The small black circles are bubbles of carbon dioxide, a product of the reaction.

THE ILLUSTRATIONS

Cover photograph by F. W. Goro

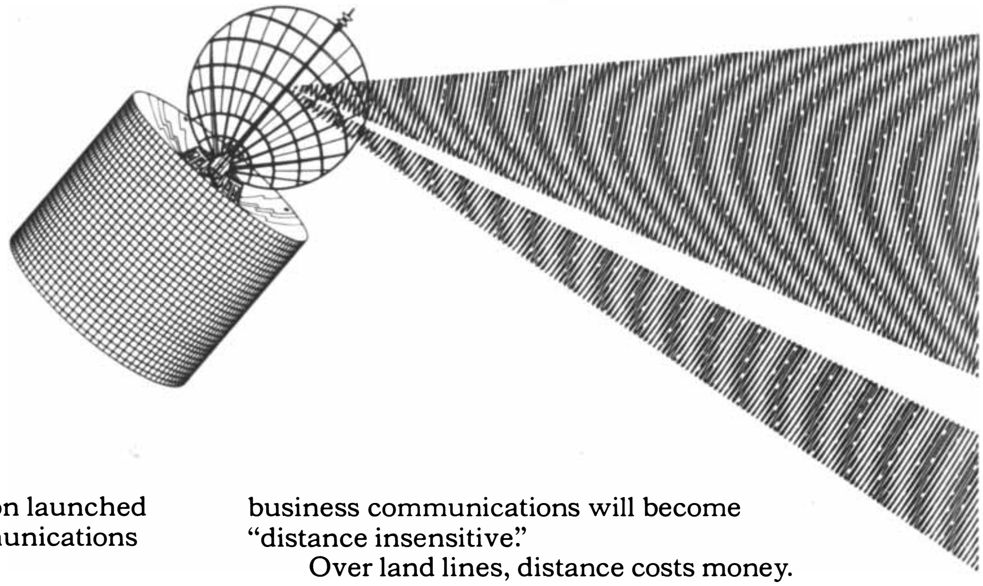
Page	Source
19–21	Association of the Bar of the City of New York
23	U.S. Court of Appeals for the District of Columbia
25	Lawrence Livermore Laboratory
26–29	George V. Kelvin
30	Lawrence Livermore Laboratory
32–37	George V. Kelvin
39–41	Lorelle M. Raboni
42	Judith H. Myers and Charles J. Krebs, University of British Columbia
43–45	Lorelle M. Raboni
46	Judith H. Myers and Charles J. Krebs, University of British Columbia
58	David Jacobowitz, National Institute of Mental Health
60	Ilil Arbel
61	Floyd Bloom, National Institute of Mental Health
62–71	Ilil Arbel
72–79	Alan D. Iselin
83–84	F. W. Goro
86–94	Dan Todd
97–105	Tom Prentiss
106	William A. Garnett
108–113	Eric G. Hieber Associates, Inc.
114	*Progrès Agricole et Viticole*
117	Ilil Arbel
118	Photo Researchers, Inc.
119	Dondero Photography
122–126	Jerome Kuhl

This hotel costs twenty-three million dollars. It soars above a private 15 acre parkland. It sits 20 minutes from the international airport and 15 from the downtown of one of the rising great cities of the world. It has 406 rooms, each a third larger than the North American average. Each room has a view. Over 90% have a balcony from which to view the view. It has a great ballroom, four licensed restaurants, two generous banquet rooms and ten meeting rooms. It is designed, built and serviced for the international traveller, visiting alone or in a group of 800. It will open in June, 1974.

THE PRINCE HOTEL TORONTO

PRINCE HOTELS INTERNATIONAL
York Mills Road at Don Mills Road
Box 9900, Don Mills, Ontario, M3C 3B7
Telephone 444-2511 Telex: 06-966725

Western Union puts a messenger into orbit.



On April 13, Western Union launched America's first domestic communications satellite—Westar I.

In June, another space age messenger—Westar II—will be launched. Shortly thereafter, both satellites will become fully operational along with Western Union's five earth stations.

Each satellite increases our long-haul capacity many times.

The two domestic satellites will vastly increase Western Union's information-moving capacity in all forms of voice, video, data and graphic communication. Each Westar satellite, for example, can relay more than 8 million words per second. And for data transmission, Westar I or II can handle six hundred million bits per second.

In essence, these satellites will become superhighways of communications that will greatly affect the way business is conducted.

A 45,000-mile shortcut from New York to Los Angeles.

What the new satellites mean is that business communications will become "distance insensitive."

Over land lines, distance costs money. But with satellite transmission, costs no longer automatically increase with each mile a message travels.

For example, in long-haul communications over 600 miles, medium and small-size businesses will be able to obtain rates that only large companies could get before. And satellite relays provide point-to-multiple-point communications, regardless of location.

Also, existing Western Union networks stand to benefit with the satellite transmission. Already, our Telex/TWX network can connect you quickly with 135,000 terminals in North America and 400,000 additional terminals worldwide.

When the Westar satellite system becomes fully operational, Telex/TWX will become an even more efficient way to send and receive important information.

And later this year our high-speed electronic mail service called Mailgram® will be moved over Westar.

LETTERS

Sirs:

The luminous and comprehensive article by F. Keith Hall on wood pulp [SCIENTIFIC AMERICAN, April] did not cover one subject that is of increasing importance in a number of connections. Why is it that some papers disintegrate in a very few decades whereas others last seemingly forever? Librarians and archivists in recent years have been given to understand that this is almost entirely a matter of residual acidity in the paper, and that therefore "acid-free" papers and "acid-free" containers for documents and files are the solution that should be promoted by all good men in every level of industry and society. Is there merit in this idea? And if there is, what are the researchers doing to ensure that printing papers, at least, are acid-free?

Whatever the answers to these questions may be, it seems clear that the rapid deterioration of papers we see in libraries suggests a serious danger in the use of wood-pulp products in construction and other industries. How long can the paper bridge to which Dr. Hall refers in his article be expected to remain safe?

FREDERICK K. HENRICH

Documents and Microforms Librarian
Lockwood Memorial Library
State University of New York
Buffalo

Sirs:

Mr. Henrich is undoubtedly correct in assuming that acidic hydrolytic degradation is primarily responsible for the degradation of book papers on prolonged storage, certainly over the first several decades of a book's life. Storage in buffered or counteracidic conditions is desirable.

Conditions of low humidity and low temperature are of enormous help in reducing degradation; the classic example of the durability of cellulose in low-humidity (but not low-temperature) conditions is the linens in Egyptian sarcophagi. These materials are weak, but they are still intact after a period of several millenniums.

Since cellulose is inherently biodegradable by several different mechanisms, avoidance of exposure to light is desirable. An oxidative reaction takes place related to the presence of carbonyl (–CHO) groups in the cellulose. This gives rise to chromophores that cause the paper to yellow and become tender as a result of depolymerization of the cellulose molecules.

The original quality of the paper is also important. Pure wood cellulose is as degradation-resistant as the cotton or linen cellulose used in the old rag papers. However, groundwood, containing lignin, and other less pure papermaking fibers must be avoided if permanence is desired. Since permanence is not a prime target for most types of paper (and indeed may even be a hindrance in the newspapers and packaging materials that litter public places), one feels that the best solution to achieve permanence in library books is to print special editions on durable paper for our archives, as is done with newspapers. Also there is considerable interest in the industry in plastic papers based on polypropylene, polyethylene or polystyrene for archival, legal or map papers. The plastic papers are quite durable.

With regard to the final question about the paper bridge, it was quite durable and safe in the low humidity of the Arizona desert, where it was installed across a ravine as a demonstration of its remarkable strength. It would not have withstood a thorough soaking. Nevertheless, those same lignocellulosic papers impregnated with suitable water-resistant cross-linking resins have since been shown to be extremely durable even when totally immersed in water. Several developments based on this technology are now reaching fruition.

F. KEITH HALL

Corporate R & D Division
International Paper Company
Tuxedo Park, N.Y.

Sirs:

In July, 1936, an enormous quantity of Newton's manuscripts was sold by auction at Sotheby's in London. These manuscripts comprised virtually all Newton's papers on alchemy, theology and Biblical chronology, together with his Mint papers and many letters. The sale was a disaster for Newton scholarship, for while his purely physical and mathematical papers are housed together in the University Library, Cambridge, all his other papers are now grievously dispersed.

Both because of its intrinsic value and because of its rarity, Sotheby's has decided to reprint the catalogue prepared for the 1936 sale. I have been invited to edit the reprint and am preparing a finding list giving the present locations of the manuscripts all over the world. This finding list will be appended to the reprinted catalogue.

About 85 percent of the manuscripts have been located, but the remaining 15 percent are proving very intractable. As it is of importance for future Newton scholarship that the finding list should be as complete as possible, I would be grateful indeed for any information on the present location of Newton manuscripts, particularly those in smaller libraries or in private possession.

P. E. SPARGO

53 Greenside Road
Greenside
Johannesburg
South Africa

Scientific American, June, 1974; Vol. 230, No. 6. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

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

Advertising correspondence should be addressed to Harry T. Morris, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

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

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

☐ Please enter change of address.

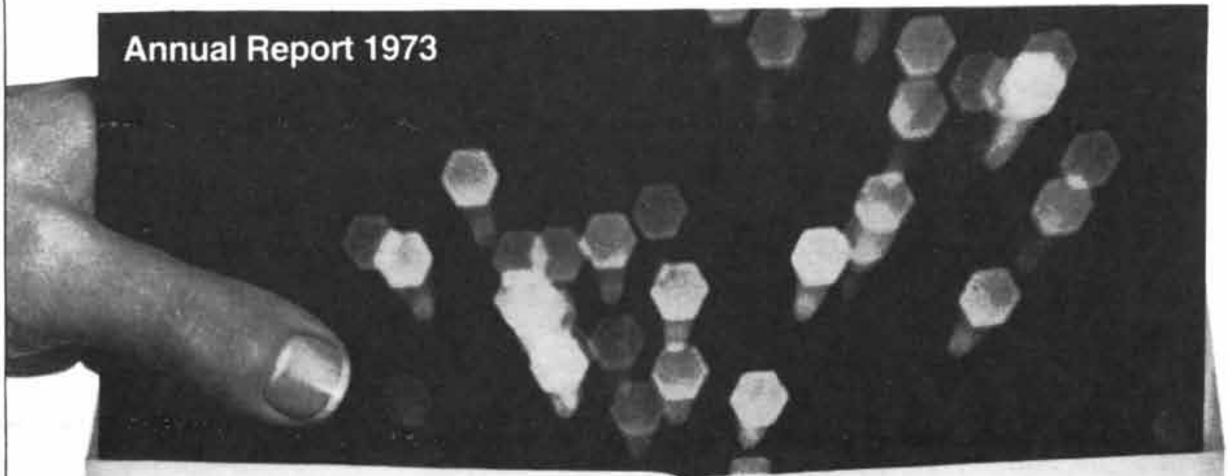
☐ Please enter subscription.
U.S. and Canada:
☐ 1 year, $12 ☐ 2 years, $22 ☐ 3 years, $30
All other countries:
☐ 1 year, $15 ☐ 2 years, $27 ☐ 3 years, $37

NAME

NEW ADDRESS

OLD ADDRESS

We're taking the wraps off 1980 with our 1973 Annual Report.



In the year 1973 we instituted Project 80—our new plan to more than double our present volume over the balance of this decade.

This will be accomplished by expanding present businesses and introducing new products and services.

Attainment of our Project 80 goals would, just as importantly, make possible steady and increasing employment. (Our employees have increased to 438,000 from 428,000 at the end of 1972.)

In 1973 our company prepared for the future with a record $852 million investment in new plants and facilities worldwide, as well as with research, development and engineering programs which reached $400 million, a portion funded by customers.

In addition, we entered 1974 with a record manufacturing backlog of $4.1 billion, equal to seven months of manufacturing sales.

More importantly, 1973 was the 14th consecutive year in which we reached record levels in sales, income, and earnings per share before extraordinary items.

Our worldwide sales and revenues were $10.2 billion—a 19 percent increase over the restated 1972 figure of $8.6 billion. This does not include insurance premiums earned and finance income of $2.6 billion, an increase of 16 percent over 1972.

Consolidated income, before extraordinary items, was a record $521 million, a gain of 9 percent over 1972 income, before extraordinary items, of $477 million.

This was equal to $4.17 per common and common equivalent share, an increase of 10 percent over 1972 per share earnings of $3.79, before extraordinary gains of $6.5 million and $6.7 million in 1973 and 1972 respectively.

On balance, 1973 was a year in which there were record accomplishments and actions in virtually all our business areas which give us a head start toward our dual Project 80 goals: business growth and greater service to people throughout the world.

For a copy of our 1973 Annual Report, which also outlines our Project 80, write to: Director of Investor Relations, International Telephone and Telegraph Corporation, 320 Park Avenue, New York, N.Y. 10022.

The best ideas are the ideas that help people. ITT

A new GC that tells you what its measurements mean.

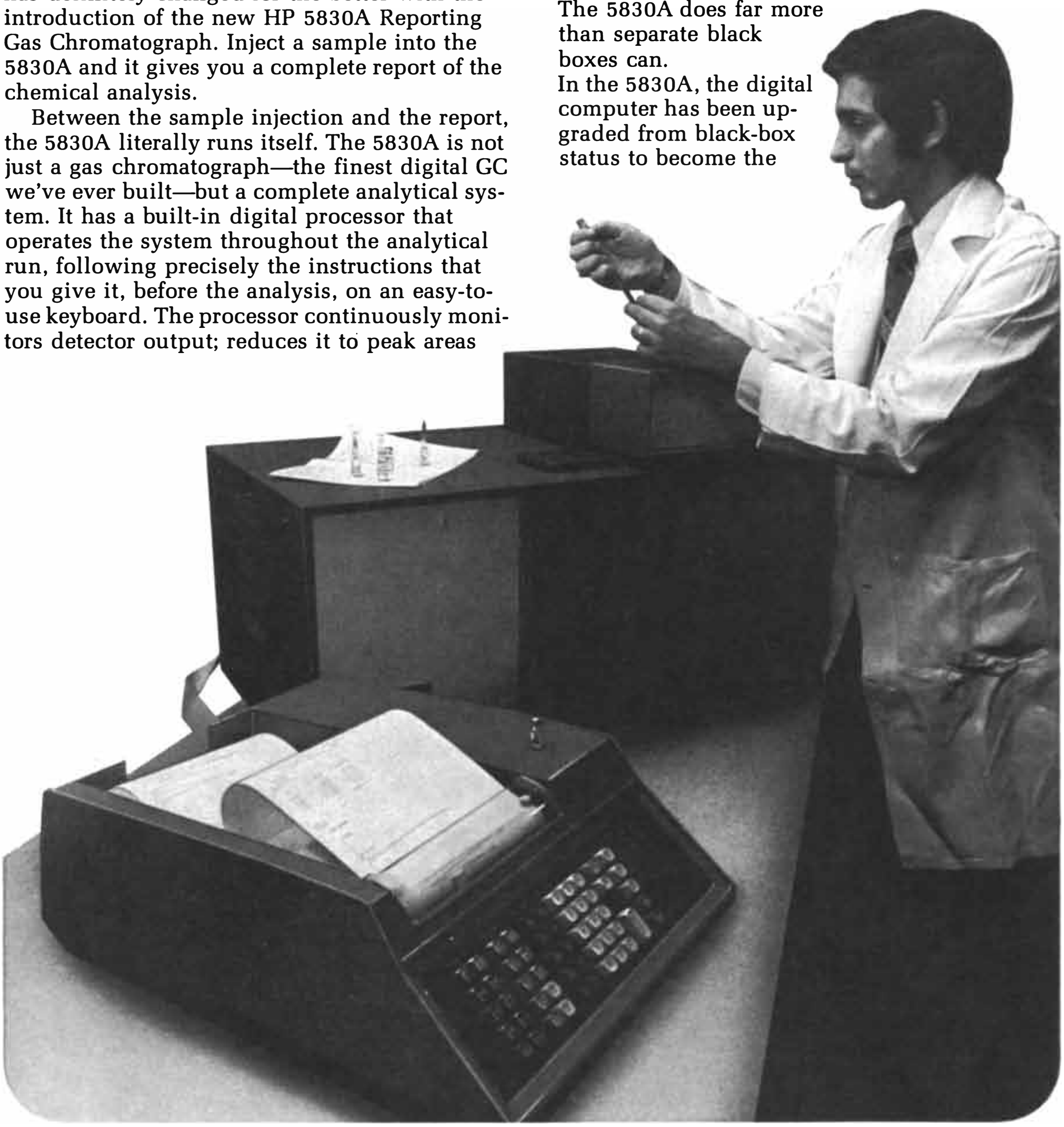
The gas chromatograph is easily the most widely used instrument in the history of chemical analysis. Chemists, biologists, and scores of other scientists use the GC to get answers to all sorts of analytical problems — *after* they've recorded its signal output, integrated individual component peaks, and reduced the quantitative data.

This rather involved process of getting the analytical answer (from the basic measurement) has definitely changed for the better with the introduction of the new HP 5830A Reporting Gas Chromatograph. Inject a sample into the 5830A and it gives you a complete report of the chemical analysis.

Between the sample injection and the report, the 5830A literally runs itself. The 5830A is not just a gas chromatograph—the finest digital GC we've ever built—but a complete analytical system. It has a built-in digital processor that operates the system throughout the analytical run, following precisely the instructions that you give it, before the analysis, on an easy-to-use keyboard. The processor continuously monitors detector output; reduces it to peak areas and times according to a sophisticated integration algorithm; identifies the sample components and computes their concentrations; and generates the chromatogram. Then it presents a complete analytical report, including a list of analysis conditions, as well as the chromatogram and the component concentrations on its own built-in printer/plotter.

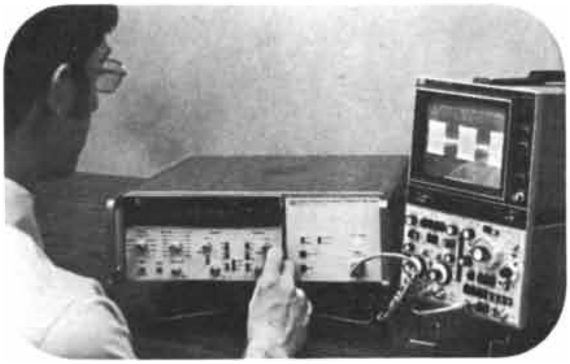
No longer do you have to assemble a series of expensive black boxes to do the job, and still wind up doing much of the work manually. The 5830A does it all for you, automatically.

And that's only the beginning. The 5830A does far more than separate black boxes can. In the 5830A, the digital computer has been upgraded from black-box status to become the



intelligent control center of the entire analytical system. From the keyboard, you can instruct the 5830A to alter conditions automatically, throughout the analysis, at times that you specify beforehand. You can give it up to 100 such time-programmed commands for each analysis—to make changes in GC conditions, detector operation, chromatogram presentation, integration and data reduction parameters, operation of accessories and even of external instruments and devices.

Is it any wonder we call the 5830A the Analytical Answer Machine? Prices start at $7995.*



The new HP 5345A counter with optional frequency converter plug-in automatically measures CW or very brief pulsed signals well into the microwave region. Here, the counter measures the pulsed signal displayed on the scope.

A radically new performance standard for frequency and time measurements.

Based on a new HP-designed-and-built family of monolithic IC's capable of sub-nanosecond switching, the HP 5345A plug-in electronic counter fundamentally improves frequency and time measurements. That's why we don't hesitate to call it "radically" new; to say anything less would amount to calling a mountain a molehill.

In only one second, you can measure any frequency from 1 Hz to 500 MHz with 9-digit resolution; or sample a signal up to 5000 times per second, with 4-digit resolution. You can resolve one-shot time interval measurements to 2 nanoseconds; or automatically average repetitive measurements to get 1-picosecond resolution, while HP's unique random dithering technique assures you of a valid answer. You can measure 100-mV signals using standard 10 megohm oscilloscope probes; or use 50-ohm input cables.

You can measure RF burst frequencies and the carriers of microwave pulses, accurately and automatically. Using a gate time of <50 nsec, you can get inside a pulse and count the carrier frequency profile or use external gating to do frequency averaging within narrow pulses, as brief as 50 nsec.

With the new fully automatic 4.0 GHz frequency converter plug-in, you can make CW measurements that are immune to very high levels of FM — and also automatically lock onto microwave pulses as short as 250 nsec.

You can make full-bandwidth ratio measurements directly since both channels have sensitive, wideband input amplifiers. In digital communications, for example, you can compare a very fast clock with an error signal as small as 2 nsec, occurring at 10 mV levels and widely spaced intervals.

With the optional third plug-in input C channel, you can examine an isolated portion of a bit stream in high-speed digital circuits by using the A and B channels to start and stop the measurement.

When equipped with the HP Interface (ASCII) Bus option, you can easily combine the 5345A with HP calculators and programmable accessories (as detailed in a series of HP application notes) for automatic operation, data reduction, statistical analysis, and self-calibration.

Cost of the 5345A counter is $3450;* automatic frequency converter adds $1950;* C-channel plug-in adds $850;* Interface bus (ASCII input/output) adds $800.*

For more information on these products write to us. Hewlett-Packard, 1503 Page Mill Road, Palo Alto, California 94304

*Domestic USA prices only 00430

Mail to: Hewlett-Packard, 1503 Page Mill Road, Palo Alto, CA 94304
Please send me information on
( ) HP 5830A Reporting Gas Chromatograph
( ) HP 5345A Electronic Counter

Name__________ Title__________
Company__________
Address__________
City__________ State______ Zip______

HEWLETT hp PACKARD

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

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JUNE, 1924: "The modern view of the crust of the earth is that at a certain considerable depth there is a vast accumulation of semi-plastic material, solid and unyielding for all ordinary purposes but yielding slowly like a fluid to very long continued forces. In this magma the continents are believed to be floating, solid and stable enough to all appearance and yet liable to slow and regular movements. Some parts might rise; opposite parts might sink; this sort of slow disturbance is known to be going on. Still more extraordinary consequences have also been suggested as resulting from this floatation theory. For example, a great floating mass of land might break up and a part of it drift away from the rest. The shape of America looks as if it once fitted on to Africa and Europe; the idea is that it once formed part of that great continent."

"Nearly all the recent and important advances in muscle physiology have resulted from a study of the phenomenon of fatigue, in particular the extreme athletic fatigue that results rapidly from a very violent effort. Recovery from this fatigue is possible only in the presence of oxygen. It was natural to suppose that the oxygen was used to oxidize some product, the presence of which acted unfavorably on the muscle. Lactic acid was known to occur in muscle, and was found to be increased by exercise and diminished or abolished by recovery in the presence of oxygen. Furthermore, there appeared to be a certain definite maximum beyond which the lactic acid content of the muscle could not be driven, even by the most vigorous stimulation; clearly this corresponds to the maximum effort a muscle can make."

"To Kant and to Laplace we owe the first rational hypothesis of stellar evolution, but they had at their command insufficient data to establish their nebular hypothesis on a sound basis. Within the past few years astronomers have accumulated a great mass of material that gives a more secure foundation for a consistent theory of stellar evolution. From the laws of physics we know the general trend of the changes that must take place in any star. If at first it is a highly rarefied and diffuse mass of vapor at a low temperature, gravitational forces will cause its gradual condensation, and this will produce an elevation of temperature. After the density reaches a certain stage condensation will proceed more slowly. When the heat lost by radiation exceeds the heat produced, the star will cool, but condensation and cooling will be extremely slow and may be prolonged by internal development of heat due to superchemical transformations made possible by enormously high temperatures and pressures. These may cause the formation of one element from another."

"Formerly geologic time was measured in millions, later in hundreds of millions, of years; now it may have to be reckoned in thousands of millions of years. This revision of ideas has been suggested by the discovery that flowering plants, which had been supposed to be relatively modern, grew in the coal age. Previous study had given the information that plant life of the coal era was of a low order, consisting of ferns, club mosses and plants of the horsetail family. The discovery now of the fossil of a highly developed flowering plant in a specimen of American coal upsets all previous conception of the age of the world. 'Over a thousand million years must have elapsed,' according to A. C. Noé, paleobotanist of the University of Chicago, 'between the origin of plant life on earth and the stage of evolution reached at the time when the coal was laid down.' "

SCIENTIFIC AMERICAN

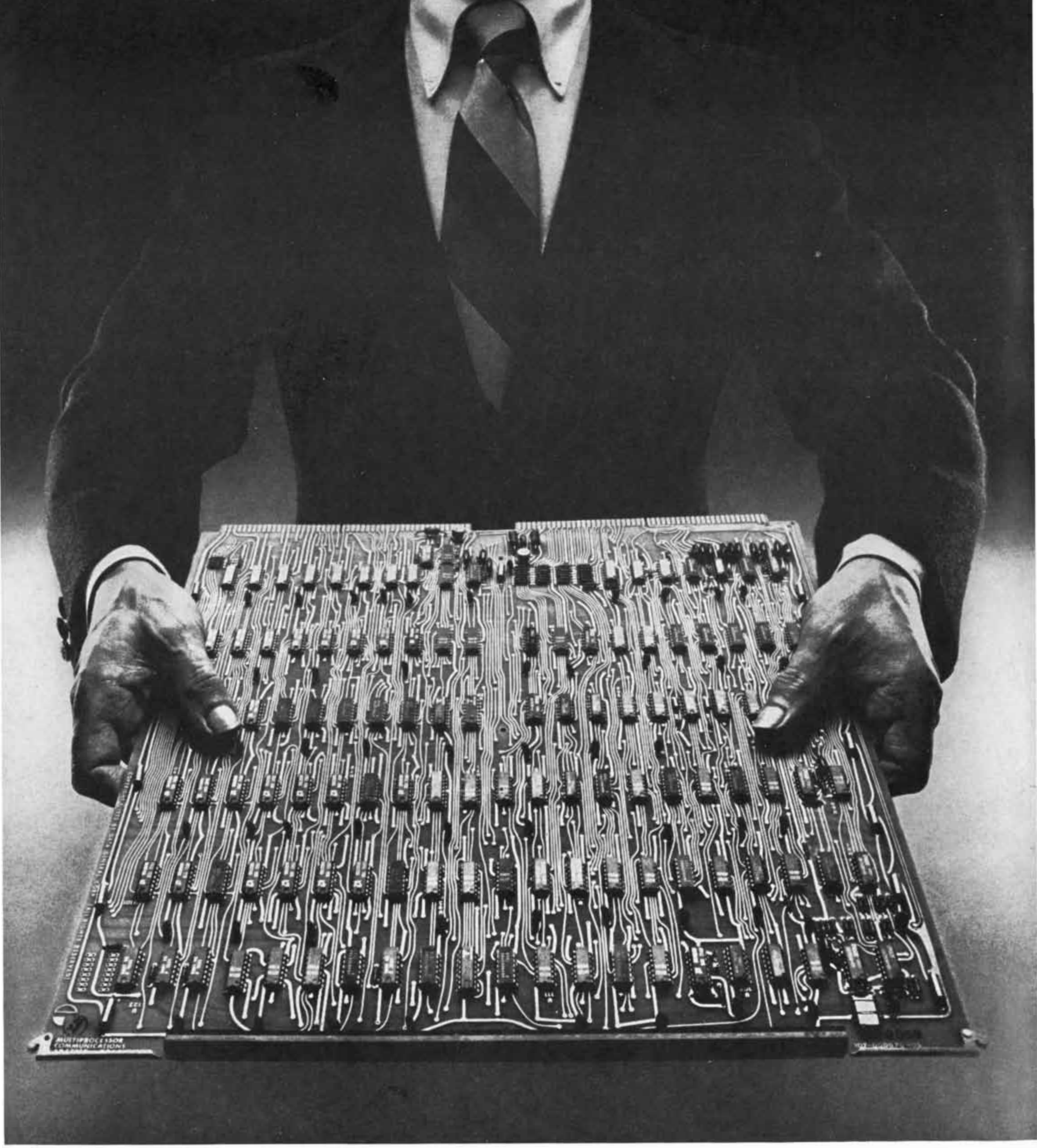
JUNE, 1874: "The theory of quantivalence, by which the modern chemistry differs so radically from the science laid down in the old text books, is based on close comparisons concerning the nature of divers chemical combinations. These have taught that each elementary atom possesses a certain definite number of bonds, by which alone it can combine with other atoms. Thus oxygen will combine with two hydrogen atoms to form water; this is expressed in the ordinary way by H_2O, but after the new method by H–O–H, indicating that the oxygen atom has two bonds, whereas each hydrogen atom is attached by only one bond. This definite quantivalence of atoms is one of their most important inherent properties. The old-fashioned authors did not question how the elementary substances were united in a compound; now it is considered of the utmost importance to determine the exact manner in which the atoms are united to build a molecular structure. A compound may be totally changed by simply changing the relative position of the atoms in the molecule, which itself may change without any alteration in the number or quality of the atoms."

"The 11 principal nations of the world, Great Britain, the United States, France, Germany, Belgium, Austria, Russia, Italy, Spain, Holland and Sweden, have more than doubled their aggregate commerce in less than 20 years. The foreign trade of these countries amounted in 1855 to \$4,251,700,000 and in 1872 to \$9,272,000,000, showing in 17 years an increase of slightly more than five billion dollars, or 118.5 per cent. The increase in population during the above period was 40,177,000 souls, or 14.8 per cent. During the first mentioned year the commerce per capita was \$15.62, in the last year \$29.76, an increase of \$14.14 to each person."

"M. Béchamp has isolated the red coloring matter of blood, which shows the presence of iron."

"It appears from the soundings made by the *Challenger* expedition that the Gulf Stream, or Florida Current, is a limited river of superheated water, of which the breadth is about 60 miles near Sandy Hook, N.J., whereas near Halifax, Nova Scotia, it has separated into divergent streams forming a sort of delta. Its depth is nowhere more than 100 fathoms; at less than double that depth we come into what is clearly polar water."

"Croce-Spinelli, in his recent balloon ascension to an elevation of 25,000 feet, finds by spectroscopic observation that the lines in the solar spectrum ascribed to the vapor of water are due to the terrestrial and not to the solar atmosphere. When the former, by reason of the elevation, is greatly eliminated, the bands also in like proportion decrease. It may therefore be considered that in the sun there is no watery vapor, at least in appreciable quantity, and that consequently the temperature of that body is not yet sufficiently lowered to allow water to form."



We set the industry standard the day we started making computers.

THE AUTHORS

DAVID L. BAZELON ("Psychiatrists and the Adversary Process") is chief judge of the U.S. Court of Appeals for the District of Columbia Circuit. He received his law degree from Northwestern University in 1931 and practiced law privately until he was appointed assistant U.S. attorney for the Northern District of Illinois in 1935. From 1940 to 1946 he was in private practice again, serving thereafter as an assistant attorney general in the U.S. Department of Justice until his appointment to the Court of Appeals in 1949. He has been chief judge of the court since 1962.

JOHN L. EMMETT, JOHN NUCKOLLS and LOWELL WOOD ("Fusion Power by Laser Implosion") are at the Lawrence Livermore Laboratory; Emmett is head of the laser fusion division, Nuckolls is associate leader of a nuclear explosion design division and Wood is a staff physicist. Emmett, who did his undergraduate work at the California Institute of Technology, received his Ph.D. (in physics) from Stanford University in 1967. He then joined the Naval Research Laboratory, where he became head of the laser-physics branch. He went to Livermore in 1972 because, he says, he was attracted by "the horrendous difficulties and fantastic possibilities of laser fusion." In his spare time he enjoys skiing and making liqueur. Nuckolls was graduated from Wheaton College in 1953 and received his master's degree (in physics) at Columbia University in 1955, thereupon joining the Livermore laboratory. Wood, whose degrees are from the University of California at Los Angeles, describes himself as "a full-time dilettante" at Livermore. "I am probably most notorious for buying the largest computer ever built (the IBM STRETCH) when the Government declared it surplus and installing it on the ground floor of the barn in whose loft I have my quarters. On the rare occasions when I forsake the designing of starships, the quest for the 'magic bullet' for cancer and similar activities at the laboratory, I take pleasure in skiing, chasing solar eclipses, backpacking, aerial acrobatics and sports-car games."

JUDITH H. MYERS and CHARLES J. KREBS ("Population Cycles in Rodents") are at the University of British Columbia; Myers is assistant professor in the Institute of Animal Resource Ecology and the Department of Plant Science and Krebs is associate professor of zoology and animal resource ecology. Myers received her bachelor's degree at Chatham College, her master's degree at Tufts University and her Ph.D. from Indiana University. She describes her avocational interests by season: "Summer, doing fieldwork in the sunshine; winter, population ecology of snowflakes on the ski slopes." Krebs obtained his bachelor's degree at the University of Minnesota and his master's degree and doctorate from the University of British Columbia. He taught at Indiana for six years before he took up his present work. He too describes his outside interests by season: "Summer, heading off to the Yukon for field studies blissfully away from the city and the chaos of academic life; winter, improving my techniques on the ski slopes, including those of skiing."

JULIUS AXELROD ("Neurotransmitters") is chief of the section on pharmacology in the laboratory of clinical science of the National Institute of Mental Health. He writes: "I was born on the lower East Side of New York City and was graduated from City College in 1933. I wanted to become a doctor but it was extremely difficult for a City College graduate to be accepted in medical school then. After working in a food-testing laboratory for about 12 years I had by a stroke of luck an opportunity to do research at Goldwater Memorial Hospital on the fate of drugs in the body. In 1950 I started work at the National Heart Institute. I received a Ph.D. from George Washington University in 1955 and joined the National Institute of Mental Health in the position I still hold. I have no major interests outside my work but am an avid reader and like to listen to music and go to the movies." Axelrod shared the Nobel prize for physiology and medicine in 1970.

RAYMOND SIEVER ("The Steady State of the Earth's Crust, Atmosphere and Oceans") is professor of geology at Harvard University. Following his graduation from the University of Chicago in 1943 he spent several years with the Geological Survey of Illinois while obtaining his master's degree (1947) and his Ph.D. (1950) at the university. He has been at Harvard since 1956. "Doing science and teaching," he writes, "frequently interferes with the important activity of longest duration of my life—playing piano. I have played every day since I was seven years old. As this is written I am practicing a Beethoven concerto to play with the local amateur symphony orchestra. I am still trying to learn contemporary music and have got up to Aaron Copland so far."

ARTHUR T. WINFREE ("Rotating Chemical Reactions") is associate professor of biological science at Purdue University. As an undergraduate at Cornell University he studied engineering physics, but he then moved toward biology and biophysics, receiving his Ph.D. (in biophysics) from Johns Hopkins University in 1970. He taught theoretical biology at the University of Chicago from 1969 to 1972. Apart from his work he is interested in flying.

RICHARD G. KLEIN ("Ice-Age Hunters of the Ukraine") is associate professor of anthropology at the University of Chicago. His graduate degrees are from that institution (his master's degree in 1964 and his Ph.D. in 1966); he did his undergraduate work at the University of Michigan. He taught at the University of Wisconsin at Milwaukee, Northwestern University and the University of Washington before returning to Chicago. "I am remarkably short on hobbies," he writes. "I do read novels, go to movies, take walks et cetera, but beyond that I spend all my time on my work. Perhaps the fact that my work involves considerable travel and is a mixture of library, laboratory and outdoor activities makes my obsessiveness a bit less pathological than it would be otherwise. (I'm not sure my wife would agree.)"

PHILIP WAGNER ("Wines, Grape Vines and Climate") is the retired editor of the Baltimore *Sun;* he continues active in journalism by writing twice a week a syndicated newspaper column on public affairs, based mainly on Washington. "Work with grapes and wine has been a second vocation for me for many years," he writes. "It is an ideal arrangement because the contrast is so complete. I keep my two occupations entirely separate." His work in viticulture, which he has done in close collaboration with his wife, has involved grape growing, experiments with new varieties, propagation of vines and wine making. The Wagners established and continue to operate the Boordy Vineyard in Maryland; since then regional Boordy Vineyard wineries have been established in New York and the Yakima Valley of Washington in association with Seneca Foods. Wagner wishes to acknowledge the assistance of Dolores Ballou, who translated the work of A. M. Negrul and other Russian publications.

Established 1845 SCIENTIFIC AMERICAN June 1974 Volume 230 Number 6

Psychiatrists and the Adversary Process

How should the determination of individual responsibility be made in a criminal case? It is argued here that the facts and reasoning of expert witnesses should be subjected to close cross-examination

by David L. Bazelon

In our society part of the task of scrutinizing the decisional process involving experts in a great many disciplines has fallen to the judiciary. With public issues increasingly conditioned by scientific discovery and technological change, courts often confront technical questions with legal and moral implications for society. As a judge on the United States Court of Appeals for the District of Columbia Circuit for the past 25 years, I have been exposed to almost every sensitive scientific and medical question that has legal, moral and ethical implications for our society. Arguments before our court, where the Federal Government is often a party, have ranged over the spectrum from abortion and blood transfusions to the underground-nuclear-explosion experiment in the Aleutian Islands and the safety of nuclear reactors. It is my duty to approach these questions not as a surgeon, a physicist, an ecologist or any other technical expert does but as one charged with monitoring the decisional process. In scrutinizing the decision-making process of experts I have seen our familiar judicial procedures attempt to bring the most arcane sciences and technologies under public surveillance.

Psychiatry, I suppose, is the ultimate wizardry. My experience has shown that in no case is it more difficult to elicit productive and reliable expert testimony than in cases that call on the knowledge and practice of psychiatry. Such cases turn up in every jurisdiction. They raise fundamental questions: Who can be morally convicted of a crime? Who can be ordered into a hospital for compulsory treatment, and for how long? What kinds of treatment can be imposed without the consent of the patient? These questions engage the overriding question of the balance of power between the state and the individual. The effort by the courts to strike that balance requires the knowledge and expertise of the experts in the behavioral sciences, particularly psychiatry.

The discipline of psychiatry has direct relevance to cases involving human behavior. One might hope that psychiatrists would open up their reservoirs of knowledge in the courtroom. Unfortunately in my experience they try to limit their testimony to conclusory statements couched in psychiatric terminology. Thereafter they take shelter in a defensive resistance to questions about the facts that are or ought to be in their possession. They thus refuse to submit their opinions to the scrutiny that the adversary process demands.

Psychiatrists are not alone, of course, in their failure to comprehend the nature and the importance of the adversary process. A word of explanation about what I mean when I use the term is in order here. The adversary process is the central feature of the system of legal institutions and procedures set up by our society to resolve controversies that arise between contending interests, values and ideologies. The adversity—to use the word in its old dictionary meaning—is supplied not by the process but by the parties to the conflict; the adversary process is merely the decisional mechanism for resolving their conflict. Decisions must be reached even in the absence of any source of perfect information or wisdom. We therefore arrange an orderly contest of the parties in the courtroom, in which adversary roles are assigned to reflect the reality of the underlying dispute. Those of us who are engaged in conducting these proceedings have an awed awareness of the risk of arriving at an imperfect decision often with enormous consequences for the individual and for our society.

Cases always present conflicts over both facts and values since they arise on petitions for redress of grievances. To find facts we rely on an exhaustive inquiry. Parties and counsel must make the best case for themselves; they must check and correct the material offered by their adversary. Specific rules make the system at once skeptical and objective. These rules of evidence presuppose that men are biased and that their testimony is invariably shaped by their background, personality, interests and values. Cross-examination challenges witnesses' veracity, accuracy and bias. The inquiry is conducted in the presence of fact finders—the jury—chosen by measures that it is hoped will ensure their impartiality. Counsel are expected to be sensitive to their own conflicts of interest and to take care they do not serve more than one master.

We might be able to develop better rules to bring out facts in the courtroom and to dispel the excesses in the proceedings that excite hostile charges from those who are subjected to rigorous cross-examination. With full awareness of its weaknesses as well as its strengths, however, we still rely on the adversary process to uncover as many of the relevant facts as possible.

The law must also reconcile competing values that seem irreconcilable. In this task it does not seek final solutions; it recognizes the ongoing nature of deep-rooted conflicts. A judge reviews and develops criteria for resolving each case as it comes before him. The criteria are made known to the public in written opinions in which the competing values are ventilated. The court's decisions are never fixed and frozen; they are altered in response to new information, new understanding and new public demands. The law of itself does not provide wisdom. It offers a method for seeking wisdom.

In this system the expert witness has a special role. He is the only witness who is allowed to testify to a conclusion ("The defendant was intoxicated") as well as to the facts ("The defendant had two drinks last night"). Precisely because the expert testifies to a conclusion, and to one that almost inevitably favors one side, he must be open to cross-examination by the other side on the facts and premises on which he rests his conclusion. To the court and to the great aim of the proceeding he owes all the insight

***200 *DANIEL M'NAGHTEN'S CASE.**

1843.

Murder. Evidence. Insanity.

The House of Lords has a right to require the Judges to answer abstract questions of existing law. (*a*)

Notwithstanding a party accused did an act, which was in itself criminal, under the influence of insane delusion, with a view of redressing or revenging some supposed grievance or injury, or of producing some public benefit, he is nevertheless punishable if he knew at the time that he was acting contrary to law.

That if the accused was conscious that the act was one which he ought not to do ; and if the act was at the same time contrary to law, he is punishable.[1] In all cases of this kind the jurors ought to be told that every man is presumed to be sane, and to possess a sufficient degree of reason to be responsible for his crimes, until the contrary be proved to their satisfaction ; and that to establish a defence on the ground of insanity, it must be clearly proved that at the time of committing the act the party accused was labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing, or as not to know that what he was doing was wrong.[2]

That a party labouring under a partial delusion must be considered in the same situation, as to responsibility, as if the facts, in respect to which the delusion exists, were real.[3]

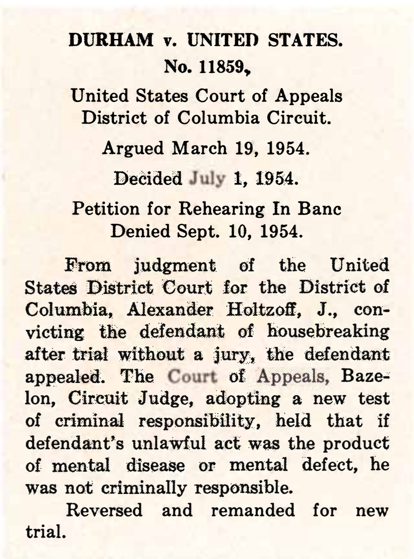
That where an accused person is supposed to be insane, a medical man, who has been present in Court and heard the evidence, may be asked, as a matter of science, whether the facts stated by the witnesses, supposing them to be true, show a state of mind incapable of distinguishing between right and wrong.[4]

M'NAGHTEN RULES, which have been the basis of the insanity plea as a defense in criminal cases throughout most of the English-speaking world, were the result of questions of law propounded by the House of Lords to the 15 criminal judges of England in 1843. The original case was the murder trial of Daniel M'Naghten (*top*), who shot and killed Edward Drummond, private secretary to the prime minister, Sir Robert Peel. M'Naghten believed he was being persecuted by the Tories and intended to shoot an individual he thought was bent on his destruction. The accused pleaded not guilty, and medical evidence was presented that he suffered from "morbid delusions." The verdict was not guilty on the ground of insanity. The case was made the subject of debate in the House of Lords, and the criminal judges were summoned to give their opinion on the law governing such cases. The official summary of their answers is reproduced here.

and information that his knowledge and training can offer.

My first exposure to psychiatry inside the courtroom concerned the expert opinions that psychiatrists were asked to render in criminal trials. In the early 1950's psychiatry and the law were deadlocked on the same issue of criminal responsibility, the so-called insanity defense. The traditional legal test for insanity, in use in virtually every American jurisdiction, was the M'Naghten test. According to this holding of a British court in 1843, the accused could not be found morally responsible if it was shown that he was suffering such "a defect of reason, from disease of the mind, as not to know the nature and the quality of the act he was doing, or, if he did know it, that he did not know he was doing what was wrong."

The M'Naghten rule thus betrayed its origins in a period when the dominant perception of human behavior was that people, as rational beings, made free choices informed by conscious consideration. The psychiatric profession was outspokenly critical of this model of the human psyche. It ignored, they said, modern dynamic understanding of man as an integrated personality, manifesting nonrational and irrational as well as rational, compulsive as well as volitional, behavior. The M'Naghten rule recognized only one aspect of that personality—cognitive reasoning—as the determinant of conduct.


DURHAM v. UNITED STATES.

No. 11859,

United States Court of Appeals District of Columbia Circuit.

Argued March 19, 1954.

Decided July 1, 1954.

Petition for Rehearing In Banc Denied Sept. 10, 1954.

From judgment of the United States District Court for the District of Columbia, Alexander Holtzoff, J., convicting the defendant of housebreaking after trial without a jury, the defendant appealed. The Court of Appeals, Bazelon, Circuit Judge, adopting a new test of criminal responsibility, held that if defendant's unlawful act was the product of mental disease or mental defect, he was not criminally responsible.

Reversed and remanded for new trial.

DURHAM RULE was formulated in 1954 by the United States Court of Appeals for the District of Columbia. It held that a defendant was not criminally responsible if the unlawful act "was the product of mental disease or mental defect." The case involved Monte Durham, who had been convicted of housebreaking. At his trial the defense asserted that Durham was of unsound mind at the time of the offense. In an appeal the defense argued that the trial court had not correctly applied the existing rules governing the insanity plea. The appeals court adopted a new test of criminal responsibility in order to provide juries with "wider horizons of knowledge concerning mental life."

Psychiatrists declared that, if the law would let them, they could give a more adequate account of realities. They complained that the M'Naghten test forced the physician to testify on whether or not the accused knew right from wrong and hence to decide the ultimate issue of moral responsibility, which should be left to the jury. Some urged their colleagues to refuse to speak to the question of "right or wrong, by virtue of reason alone." If psychiatrists were to be qualified as experts in criminal cases, they should be allowed to address the issue of responsibility in terms appropriate to their medical discipline.

Now, the law recognizes that the question of guilt or innocence is essentially a moral one. I believe the morality of a person's actions cannot be determined solely by abstract philosophical principles without regard to the facts that condition human behavior in the real world. To obtain these facts I wrote the opinion in the *Durham* case, in 1954, in which our court formulated a new test of criminal responsibility. *Durham* held that an accused man is not criminally responsible if it is shown that his unlawful act was the product of a mental disease or defect. *Durham* was not based on any notion that psychiatrists know everything there is to know about behavior. Its purpose was to bring into the courtroom the knowledge they do have and to restore to the jury its traditional function of applying "our inherited ideas of moral responsibility" to those accused of crime.

The response of the psychiatric profession to the new rule was enthusiastic. Here was the chance they had asked for to bring their expertise into the courtroom. Karl Menninger described the decision as "more revolutionary" than the *Brown* decision, also given in 1954, that outlawed racial segregation in the public schools. If the *Durham* "revolution" comes with the same "deliberate speed" as the desegregation ordered by the United States Supreme Court in *Brown,* however, only our great-grandchildren will be able to validate the Menninger appraisal.

Gregory Zilboorg had this to say about *Durham:* "[The *Durham* rule] might require of the psychiatrist that he study offenders and examine them clinically with much greater care than has often been the custom. . . . But at least . . . psychiatry is permitted to take the witness stand with all the dignity, medical and professional, which is due it. . . . The old unpleasantness of the 'adversary proceeding' in which psychiatric expert testimony has so often been engulfed . . . is bound to disappear."

Zilboorg's appraisal can be evaluated now. He was wrong. Psychiatrists continued to limit their testimony to conclusions—applying conclusory labels to the defendant—without explaining the origin, development or manifestation of a disease in terms comprehensible to the jury. They began to wage a war of words in the courtroom, arguing about whether a defendant had a "personality defect," a "personality problem," a "personality disorder," a "disease," an "illness" or simply a "type of personality." Even more disturbing, they began to speak conclusively to the question of whether the criminal act was the "product" of mental disorder. The word "product" in their testimony conveyed no more clinical meaning than "right" or "wrong."

From St. Elizabeths, the Federal Government's well-known mental hospital in the District of Columbia, from which many cases come before our court, there came this insight into the response of the profession to the *Durham* test.

"Our psychiatric staff became alarmed that the courts would equate personality disorders with the psychoses and because of this anxiety, we erroneously—although in good faith—decided to add the words 'without mental disorder' in parenthesis immediately following the diagnosis . . . of personality disorder."

What psychiatrists have not understood is that conclusory labels are no substitute in judicial proceedings for facts derived from disciplined investigation. Labeling a person "schizophrenic" does not make him so! Although the law must settle for an "educated guess," that guess is only as good as the investigation, the facts and the reasoning that underpin it.

The sterility of the profession's response to *Durham,* I now conclude, was due to the fact that its observance was bound to make the psychiatrist's task in the courtroom much more demanding than before. The late Winfred Overholser, superintendent of St. Elizabeths, once told me that the breadth of information I envisioned being placed before the jury would require from 50 to 100 man-hours of interviewing and investigation; he declared that a public hospital simply could not afford it. If that were the case, I replied, psychiatrists should frankly explain on the witness stand that their opinions are thus qualified by lack of time and resources. It was no service to the administration of justice for them to create the false impression that they had learned substantially all that could be known about someone on the basis of study they knew was inadequate.

The jury is equally entitled to know of differences of opinion and outright conflicts involved in psychiatric diagnosis. Consider this case reported by a leading forensic psychiatrist:

"A 17-year-old boy, confined as a patient in a state hospital, strangled another patient. He was sent to another hospital for medical-legal observation. The diagnostic choices lay between schizophrenia and schizoid personality disorder. Which diagnosis was made would naturally have great bearing on his criminal responsibility. When the case was presented to the staff conference, there was a great division of opinion.... After repeated medical conferences over a period of a year, it was agreed that further observation was futile, and that a diagnosis would have to be agreed on. The vote was five to four in favor of psychopathy [that is, personality disorder]. A report was then sent to the court stating that this diagnosis had been made, following a year's observation, and that he was now sane and could stand trial. The wording of the report clearly implied that the diagnosis was definitely established and that there was full agreement."

In the words of the psychiatrist reporting this incident: "What we have here, of course, is nothing else than the familiar star-chamber proceeding. The hospital staff usurped the function of the jury ... and the court was then deprived of the full evidence, the conflict of medical opinion."

Attempts by my court to obtain records or tapes of just such clinical conferences have consistently been opposed and thwarted by the psychiatric staff at St. Elizabeths. In asking the profession to open up its opinions and decisions, we were asking no more than that psychiatric expertise should submit to the process by which the shortcomings of all opinion evidence are tested. The status of court-appointed "impartial" expert is not open to any profession. The potential for bias, distortion and deviation from the truth is unavoidable. Like any other man, a physician acquires an emotional identification with an opinion that comes down on one side of a conflict; he has an inescapable, prideful conviction in the accuracy of his own findings. These realities belie the myth that a medical expert can speak from above or outside the legal adversary system.

In the end, after 18 years, I favored the abandonment of the *Durham* rule because in practice it had failed to take the issue of criminal responsibility away from the experts. Psychiatrists continued to testify to the naked conclusion instead of providing information about the accused so that the jury could render the ultimate moral judgment about blameworthiness. *Durham* had secured little improvement over M'Naghten.

In 1972, in the *Brawner* case, our court unanimously set aside the *Durham* rule. Essentially we replaced it with the insanity test that had been proposed by the American Law Institute:

"A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the wrongfulness of his conduct or to conform his conduct to the requirements of the law."

The *Brawner* formulation was designed in large part to end the expert's dominance over the question of moral responsibility. Psychiatrists will nonetheless still be able to take away the jury's function by presenting conclusory testimony. Thus they will testify that a defendant lacked capacity "as a result," just as under *Durham* they testified to whether the act was "the product" of a mental disorder.

Although no phrase will magically solve the problem of expert dominance, my own separate opinion in *Brawner* suggested the jury be instructed that a defendant is not responsible "if at the time of his unlawful conduct his mental or emotional processes or behavior controls were impaired to such an extent that he cannot justly be held responsible." This approach envisions that the jury will be provided with a broad range of information about the accused from a variety of sources including but not necessarily limited to psychiatrists. Other disciplines with special skills and knowledge in the field of human behavior would not be precluded from the opportunity to show the relevance of their data in the courtroom. Moreover, experts will be less likely to address the ultimate issue: whether the accused can be "justly held responsible." Even the promise of this approach, however, will be broken unless means are provided for the defendant, who is virtually always indigent, to obtain and present the required broad spectrum of information.

The form of words by which we call the insanity defense is not crucial, but the provision to the jury of all arguably relevant information about the accused's freedom of choice is. As I said in my opinion in *Brawner*, "while [we] generals are designing inspiring new insignia for the standard, the battle is being lost in the trenches."

In the *Jenkins* case, in 1962, spokesmen for the psychiatric profession betrayed a further misunderstanding of the role of the psychiatric witness that must

Vincent E. JENKINS, Appellant,

v.

UNITED STATES of America, Appellee.

No. 16306.

United States Court of Appeals District of Columbia Circuit.

On Rehearing in Banc.

Reargued Feb. 9, 1962.

Decided by Judgment Entered April 12, 1962

Opinion Rendered June 7, 1962.

Defendant was convicted of housebreaking with intent to commit an assault, assault with intent to rape and assault with a dangerous weapon. The defendant relied solely upon defense of insanity. From a judgment of the United States District Court for the District of Columbia, Edward M. Curren, J., the defendant appealed. The Court of Appeals, Bazelon, Circuit Judge, held, inter alia, that the court's sua sponte exclusion of psychiatrist's testimony concerning his changed diagnosis of accused because psychiatrist did not conduct a personal re-examination constituted prejudicial error requiring reversal for new trial, and that the determination of a psychologist's competence to render an expert opinion based on his findings as to presence or absence of mental disease or defect must depend upon nature and extent of his knowledge, and it does not depend upon his claim to title "psychologist", and that determination, after hearing, which may be conducted in presence of jury unless special circumstances warrant its exclusion, must be left in each case to traditional discretion of trial court subject to appellate review.

Reversed and remanded for new trial.

JENKINS CASE of 1962 raised the issue of whether or not a psychologist was competent to render an expert opinion on the presence or absence of mental disease or mental defect. The American Psychiatric Association argued that psychologists were not qualified because they lacked medical training. The majority opinion of the appeals court held that qualified psychologists could give expert opinion, and that "the critical factor in respect to admissibility is the actual experience of the witness and the probable probative value of his opinion."

have been widely shared among their colleagues. The trial court in this case had excluded the testimony of highly qualified and certified clinical psychologists, on the ground that "a psychologist is not competent to give a medical opinion." In the appeal to our court the American Psychiatric Association supported the lower court's decision. It asserted that in medical problems medical opinion can be the only guide. It chose to overlook the fact that the problem of criminal responsibility is not the exclusive terrain of psychiatry. I wrote the opinion for our court rejecting such guild mentality.

The "right to treatment" cases, in the 1960's, brought the issues under consideration here into sharper focus. In the *Rouse* case, in 1966, a patient confined without his consent sued for release or for adequate treatment of his disability. For our court I wrote that the plaintiff had a right to treatment, founded on a statute passed by Congress for the District of Columbia. This placed a more difficult question before us: How to establish whether this theoretical right was violated in actual practice. We held that adequacy of treatment is reviewable in court. In view of the inherent lack of certainty in psychiatric (as in other scientific) decisions, we held that the question of adequacy must be weighed on the basis of the "state of the art." We imposed no artificial criteria of success or failure, nor did we suggest we would overturn informed medical judgments. Independent experts should be enlisted, we said, to establish the parameters of acceptable treatment, and we urged that the American Psychiatric Association, which had published standards of medical care, be consulted.

The American Psychiatric Association responded to *Rouse* with an adamant statement of what I must call professional mystique: "The definition of treatment and the appraisal of its adequacy are matters for medical determination." This declaration ignored the explicit message in *Rouse* that the court does not presume to assess the quality of anyone's performance, unless that performance is patently arbitrary and capricious. As in all administrative law, the task of the court is to ensure that the administrative process itself controls abuse of discretion, that a factual record is established, that alternatives are considered and that reasons for decisions are set forth.

The instant opposition to *Rouse* was a clue to a deeper discordance between the professional and the judicial outlook. Plainly the profession was blind to or was concealing the conflict between the imposition of treatment and the human and civil rights of patients. In the view of most members medical decisions are by definition made in the best interests of the patient. If a physician says a man is sick, he must be sick; if he says the man must be treated or confined, that must be what is best for him. Such bootstrap reasoning comes under scrutiny only when, in the case of psychiatric prescriptions, it calls for involuntry treatment. The patient's interest in release, in less restrictive confinement or in adequate treatment cannot be matters solely "for medical determination."

Bringing these matters into court does not impose an artificial adversary relationship between the patient and his keepers; it reflects an adversity that already exists. This proposition comes as a surprise, of course, to the psychiatrist engaged exclusively in office practice and to his voluntary patient. In the public sector the adversity of interests that confronts the psychiatrist and his involuntary patient—although it does not encompass the entire relationship—must be recognized as an inescapable reality.

What is more, it should be recognized that the physician-patient relationship in the public sector is compromised by the interests of third parties and the pressures from hidden agendas. At the state hospital in Napa, Calif., the superintendent told me a few years ago in a public meeting that the staff had "Sacramento looking over its shoulder." Psychiatric decisions—to confine or release—are influenced by public outcry for "law and order." In some hospitals the shortage of beds and manpower has been known to override medical determinations; in Veterans Administration hospitals the need to fill empty beds also presses medical determinations, but in the opposite direction. Psychiatrists have justified fudging their testimony on "dangerousness"—a ground for involuntary confinement—when they were convinced that an individual was too sick to seek help voluntarily.

What is disturbing about these situations is not that they impute venality or frailty to merely human practitioners, nor that conflicting societal interests can dictate different and not necessarily the best medical results. It is rather that the psychiatric profession should resist facing these conflicts in the open. Serious legal challenges have been needed to surface its hidden agendas.

The hazard implicit in professional resistance to public scrutiny is well illustrated by the use in the U.S.S.R. of psychiatric facilities for the suppression of political dissenters. I had occasion to read the Russian case studies as a member of an *ad hoc* committee set up by the American Psychiatric Association. The studies showed how the medical model of "sickness" could be perverted to encompass judgment of what is socially and politically unacceptable behavior. On the record the physicians did not seem to be acting in their patients' best interests, or even in their own direct self-interest, but were using psychiatric terminology and techniques in the service of state policy. Yet when I was in the U.S.S.R., the Russian psychiatrists steadfastly insisted that they follow the medical model, much as their American colleagues are known to do.

As I read these case studies, however, I was impressed to realize how, in many analogous situations in this country, I had had occasion to find psychiatrists making decisions for motives and under pressures from outside their professional role. Whenever psychiatrists enter the public sector to apply their knowledge in the service of public institutions—the military, state hospitals, schools and penal institutions, to name only a few—they face conflicts between the therapeutic interests of their patients and the institutional interests of their employers. Needless to say I found the leadership of the profession in America not nearly as eager to investigate such conflicts in their own ranks as they were to look into evidence of malpractice in the U.S.S.R. To make a not very pleasant story short, when our *ad hoc* committee turned its attention to the American scene, at first with official approval, its charter was soon revoked. The need for examination remains. The American Bar Association and the Institute of Medicine of the National Academy of Sciences plan studies of these conflicts. The American Psychiatric Association apparently recognizes the importance of acting itself; it now has commissioned an inquiry by the Institute of Society, Ethics and the Life Sciences in Hastings, N.Y.

Such guild self-protection is not, of course, peculiar to the psychiatric profession. Business enterprises, labor unions and government agencies all exhibit the same penchant for privacy. The usual counter to the call for public scrutiny is the promise of "self-regulation." Whereas peer review is a much praised and not much observed principle of the medical profession, it has been largely foreign to the practice of psychiatry. Its practitioners work alone with their patients, behind ritually locked doors. For medical practitioners of all kinds peer review is now required, however, for the validation of professional performance paid for by Medicaid and Medicare funds. Curiously, this requirement was imposed not out of professional agreement that self-evaluation is in order but because of Congressional concern at rising medical costs.

For monitoring the performance of a profession there is no substitute, in the end, for the adversary process. This discussion has focused on psychiatry because decisions grounded on this discipline and on counsel from its practitioners are employed by the state to confine people against their will and to treat people in ways they do not ask for. As such the discipline is of concern to those of us whose judicial duty is to scrutinize governmental intrusions on liberty. Much of what I have said, however, applies equally well to the public surveillance of other highly specialized professions on which the operation of our complex civilization depends. Today every profession is being challenged by those who believe that trust should rest not on mystique but rather on what the public knows about its exercise of its expertise. Challenging the expert and digging into the facts behind his opinion is the lifeblood of our legal system, whether it is a psychiatrist characterizing a mental disturbance, a physicist testifying on the environmental impact of a nuclear power plant or a Detroit engineer insisting on the impossibility of meeting legislated automobile exhaust-emission standards by 1975. It is the only way a judge or a jury—or the public—can decide whom to trust.

BRAWNER DECISION by the United States Court of Appeals for the District of Columbia in 1972 set aside the Durham rule and replaced it with a rule set forth in the Model Penal Code of the American Law Institute. Archie W. Brawner had been convicted of second-degree murder. Expert witnesses at the trial had testified that Brawner suffered from a "psychologic brain syndrome associated with a convulsive disorder," but the defense and prosecution witnesses disagreed about whether or not there was a causal relation between the mental disorder and the alleged offense. One objective of the appeals court in setting out a new criterion of insanity was to "prevent the experts from exercising undue dominance over the jury.... While the expert may testify as to the existence or not of mental disease, and the causal relationship between such disease and the defendant's capacity to control, and appreciate the wrongfulness of, his conduct, he will be required to present the basis underlying his conclusions."

NOTE: On occasion, a syllabus by the court is released at the time the slip opinion is issued. The syllabus is not part of the opinion of the court, but has been prepared for the convenience of the reader.

United States Court of Appeals

FOR THE DISTRICT OF COLUMBIA CIRCUIT

No. 22,714

UNITED STATES OF AMERICA

v.

ARCHIE W. BRAWNER, APPELLANT

Appeal from the United States District Court
for the District of Columbia

On Rehearing En Banc

Decided June 23, 1972

Syllabus By The Court

This appeal, from a judgment following a jury conviction for second degree murder, concerns the defense of insanity. The court sets forth a new standard for the insanity defense. It remands to the District Court to consider whether, in view of the doctrine, the judgment appealed from should be retained or a new trial awarded. These are the principal features of the decision:

1. The court adopts as the criterion of insanity, for all trials beginning after today, the rule stated in § 4.01(1) of the Model Penal Code of the American Law Institute. That rule, which has been adopted in essence by the other Federal circuit courts of appeals, states: "A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity to appreciate the wrongfulness of his conduct or to conform his conduct to the requirements of the law." The rule of Durham v. United States, 94 U.S. App.D.C. 288, 214 F.2d 862 (1954), which excused an unlawful act if it was the product of a mental disease or defect, will no longer be in effect.

FUSION POWER BY LASER IMPLOSION

Laser-fusion schemes are based on the ignition of a pellet of fuel by focused laser beams. For the laser approach to succeed the fuel must be imploded to 10,000 times normal liquid density

by John L. Emmett, John Nuckolls and Lowell Wood

The first laser was demonstrated by Theodore H. Maiman in 1960. Less than a year later computer calculations were undertaken by Stirling A. Colgate, Ray E. Kidder, John Nuckolls, Ronald Zabawski and Edward Teller to try to find out what would happen when tiny deuterium-tritium pellets were imploded to thermonuclear conditions by intense beams of laser light. It was also proposed that the fusion microexplosions could be applied to the generation of power. These early computer calculations showed that efficient generation of fusion energy would not result from simple laser heating of thermonuclear fuel but that such generation could be achieved if lasers were used to implode the fuel to 10,000 times its normal liquid density. Although laser fusion was frequently discussed in the late 1960's, the crucial concepts (including implosion) were not declassified by the Atomic Energy Commission until 1972.

Small experimental programs were initiated in the early 1960's at the AEC's Lawrence Livermore Laboratory and later at other laboratories throughout the world. The crucial experiments could not be conducted, however, because not enough laser energy could be delivered in a short pulse. Then laser technology improved rapidly, and in 1968 N. G. Basov and his colleagues at the Lebedev Physics Institute in the U.S.S.R. reported the first observations of neutrons emitted by a laser-heated plasma.

Today our laboratory at Livermore, the Los Alamos Scientific Laboratory and groups in the U.S.S.R. are planning to develop lasers that can deliver about 10,000 joules of energy in pulses lasting a nanosecond (a billionth of a second) or less. With such lasers it will be possible to conduct the crucial laser-fusion experiments. If efficient burning of thermonuclear fuel can be demonstrated, advanced laser systems may be able to generate power on a commercial scale. Our current studies indicate that a one-gigawatt (million-kilowatt) power plant would call for 100 fusion microexplosions per second.

The central problem in releasing thermonuclear energy is to confine an intensely hot plasma of light nuclei, such as isotopes of hydrogen, long enough for a reaction to take place. Inside stars plasmas of ordinary hydrogen and other light elements are raised to ignition temperatures and confined by gravitational pressure. In thermonuclear explosives the heavy isotopes of hydrogen (deuterium and tritium) are heated by a fission explosive. The resulting plasma is confined by its own inertia, that is, by the finite time required to accelerate and move material a significant distance. The original approach to the controlled release of thermonuclear energy requires magnetic bottles to confine plasmas of deuterium and tritium (DT). Laser fusion reintroduces the inertial-confinement technique by exploiting lasers to heat plasmas with short pulses of light at extremely high power densities. Over the past four years the AEC has increased funding of the laser-fusion program more than tenfold, so that the laser-fusion effort is now comparable in scale to the magnetic-confinement one.

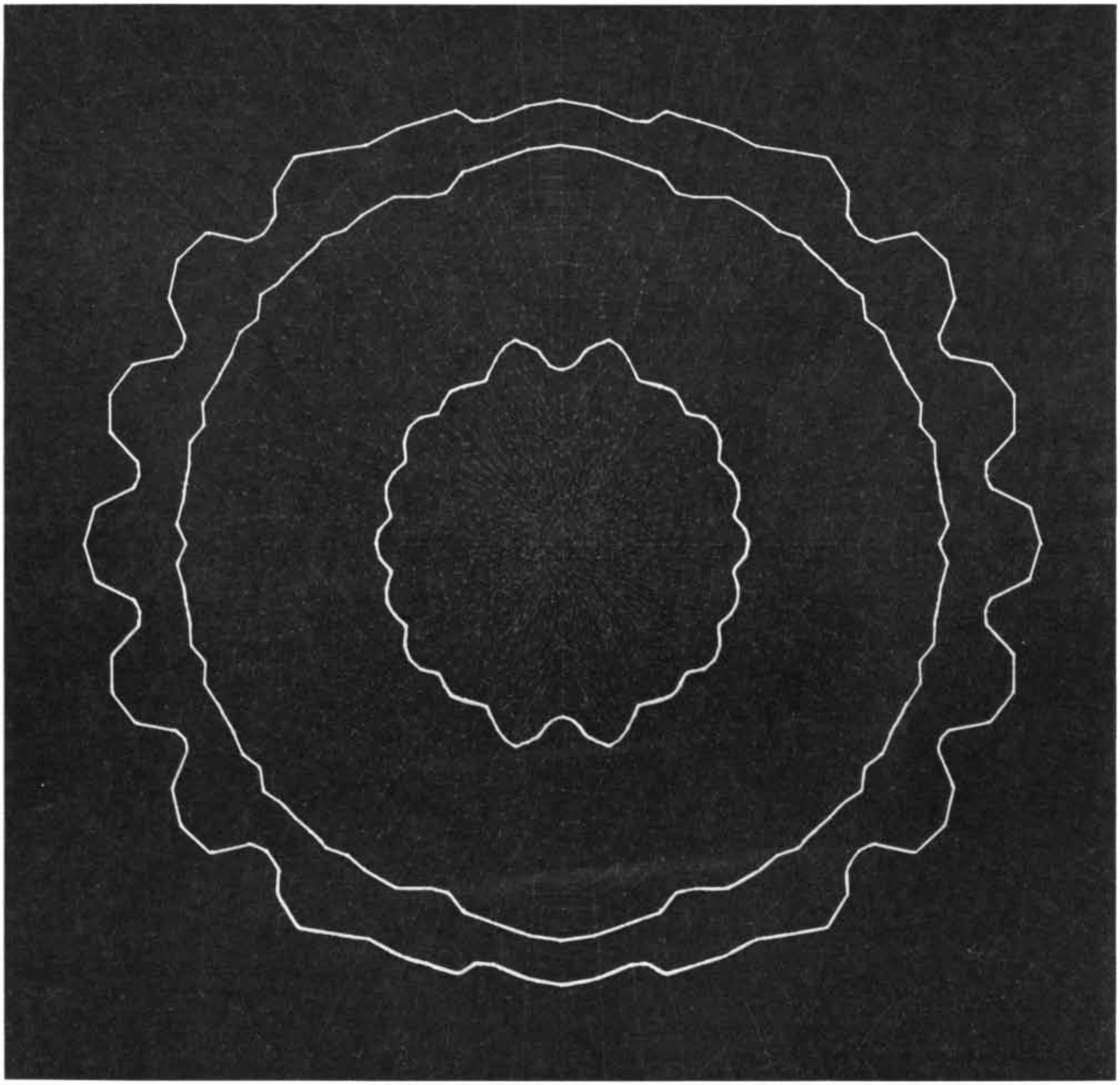
For a given thermonuclear fuel to burn efficiently certain conditions must be met regardless of the confinement scheme. First, the thermonuclear ignition temperature must be reached, which is about 100 million (10^8) degrees Kelvin for DT. Such temperatures are needed to impart sufficient thermal velocity to the nuclei to overcome their mutual electrical repulsion during a collision. At such temperatures the nuclear reaction proceeds fast enough for the energy generated to exceed the energy lost by radiation from the plasma. Second, the plasma must be contained long enough and at a high enough density for a sizable fraction of the nuclei to react. This second requirement is expressed in the Lawson number (devised by the British physicist J. D. Lawson), which is obtained by multiplying the density of the interacting particles by the confinement time. The Lawson number, expressed in seconds per cubic centimeter, specifies the break-even condition on the assumptions that no more than a third of the energy released must be fed back to sustain the reaction and that the plasma temperature is 10^8 degrees K.

The Lawson number is approximately the same for both laser fusion and magnetic-confinement fusion: about 10^{14} seconds per cubic centimeter for a DT fuel mixture. In magnetic confinement the objective is to design magnetic bottles that can confine the fuels for comparatively long times; the maximum density is then determined by the highest magnetic field that can be sustained by available materials. Plausible numbers are a confinement time of about a second and a density of about 10^{14} particles per cubic centimeter. In laser fusion the objective is to implode fuels to extreme densities (about 10^{26} particles per cubic centimeter); the effective confinement time is then determined by the inertia of matter (slightly more than 10^{-12} second). In the first reactors designed around either approach the energy-carrying neutrons released by thermonuclear reactions will probably be absorbed in a lithium blanket, raising its temperature. The heat removed from the blanket will then be used to generate steam and produce electricity in the usual way.

The principles of inertial-confinement fusion and the function of implosion can be easily illustrated with a simple example. Suppose we have a laser capable of producing a million joules of optical energy. This amount of energy is just sufficient to heat one milligram of the fastest-burning thermonuclear fuel (DT) to the ignition temperature of 10^8 degrees K. At normal liquid density one milligram of DT forms a sphere about one millimeter in radius. The efficiency with which the fuel can be burned in inertial confinement is fixed by the competition between the confinement time and the fusion reaction time.

For a one-millimeter pellet at the ignition temperature the inertial confinement time (which is proportional to the pellet radius divided by the thermal velocity of the nuclei) is about 2×10^{-10} second. The fusion reaction time (which is inversely proportional to the pellet density) is about 1,000 times longer, or about 2×10^{-7} second for fuel at liquid density. Consequently only a thousandth, or .1 percent, of the pellet would burn, yielding only a third of the laser's million-joule input.

If, however, the same pellet were imploded tenfold in radius (1,000 times in volume), the confinement time would likewise be reduced tenfold, but more important the burn time would be reduced by a factor of 1,000. As a result the burn efficiency would increase by a



LASER-DRIVEN IMPLOSION of a 60-microgram pellet of the hydrogen isotopes deuterium and tritium is simulated in a two-dimensional calculation at the Lawrence Livermore Laboratory of the University of California. This illustration is one frame from a computer motion picture based on the calculation. The original radius of the pellet before it is heated and imploded by 11 laser beams is 400 micrometers (.4 millimeter). When the pellet is compressed to a radius of about 15 micrometers, its center finally reaches ignition temperature: 10^8 degrees Kelvin. In the frame shown here, about seven picoseconds (7×10^{-12} second) after ignition, the pellet has reexpanded to a radius of about 17 micrometers, or approximately to the position of the second-largest of the three white rings, representing isotherms of ion temperature. The innermost ring is a thermonuclear detonation front at 5×10^8 degrees K. propagating through the pellet behind two earlier fronts, one at 2×10^8 degrees (*middle ring*) and one at 10^8 degrees (*outer ring*). The blue network shows hydrodynamic motions. The motion picture was made by George Zimmerman and Albert Thiessen.

factor of 100 to 10 percent, yielding 30 times the laser-input energy. It is clear that any laser-fusion power-generation scheme based on heating of a pellet of thermonuclear fuel at normal density is doomed to failure. To achieve an energy release significantly greater than the laser input, convergent compression of the fuel pellet is essential.

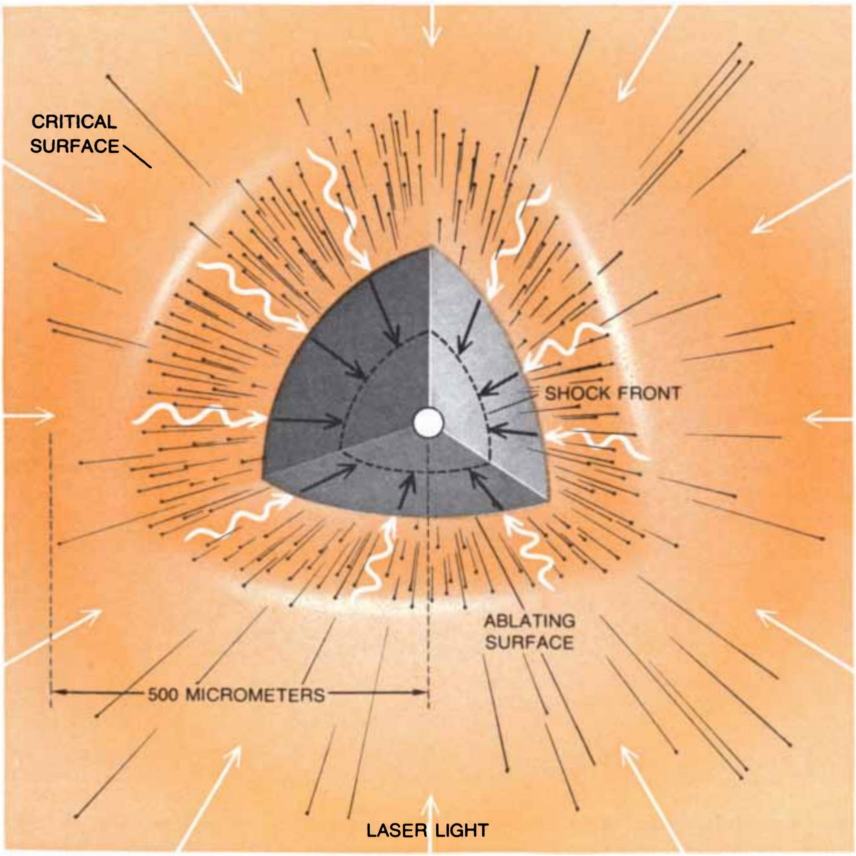
This is not quite the whole story. It turns out that if a pellet is suitably compressed, it is no longer necessary to heat the entire mass to the ignition temperature, and thus the laser energy required is reduced. The energy released by the fusion of DT is carried chiefly by neutrons of 14 million electron volts (MeV) and alpha particles (helium nuclei) of 3.5 MeV. If these particles are released in a plasma of 10^8 degrees K. whose density is that of liquid DT, the particles would have a range more than 10 times greater than the one-millimeter pellet radius. Hence nearly all the energy would escape without heating the rest of the pellet. The ranges of the neutrons and alpha particles, however, are inversely proportional to the density of the medium. Accordingly if they are released in a plasma whose density is 1,000 times greater than liquid density, the neutron range would be only a few times greater than the .1-millimeter radius of the compressed pellet, and the alpha-particle range would be less than a sixth the radius. Consequently the alpha particles would deposit energy within the pellet and propagate a thermonuclear burn front. As a result the effective ignition energy is reduced by a factor of nearly 100. This is extremely fortunate; otherwise laser-fusion power generation would be impossible because the energy released by the burning of DT is not enough larger than the ignition energy to make up for inefficiencies of the laser, the implosion and the reactor [*see illustrations on opposite page and on page 28*].

Given the fundamental necessity for imploding the thermonuclear fuel to high densities in laser fusion based on inertial confinement, one must ask: Are the required densities even conceptually possible? The feasibility is suggested by the surprising fact that the energy needed to compress hydrogen to 10,000 times liquid density is only 1 percent of the energy required to heat DT to the ignition temperature. Energetically, therefore, compression is virtually free compared with ignition. In order to achieve a 10,000-fold compression, however, truly astronomical pressures must be generated: at least 10^{12} atmospheres. Comparable pressures exist in the core of stars, where they are maintained gravitationally by an overlying mass of 10^{33} grams or more.

Can such pressures be generated on the earth? Our analyses indicate that laser-driven spherical implosions should be capable of meeting or exceeding the required values. In the laser-driven spherical implosion, light is focused by a lens (or mirror) onto a low-density atmosphere of material as it evaporates from the surface of a tiny spherical pellet [*see illustration on this page*]. The intense focused laser light is absorbed in the low-density plasma atmosphere by electron-ion collisions or by plasma instabilities, creating "hot" electrons with energies of a few thousand volts. Nearly all this absorption occurs in what is called the critical-density region of the plasma, where the frequency of plasma oscillations approaches that of the laser-light frequency.

The distance that laser light must travel before it is absorbed by collisions in the critical-density region is proportional to the square of the wavelength of the radiation. Thus in a plasma at 10^7 degrees K. the collisional absorption wavelength is one millimeter for radiation whose wavelength is one micrometer and 100 millimeters for radiation whose wavelength is 10 times longer. Since the pellet dimensions are about one millimeter, 10-micrometer radiation cannot be absorbed efficiently by the collisional process. Fortunately, however, in the laser implosions of interest the extremely high intensity of 10-micrometer radiation excites plasma instabilities that lead to efficient absorption.

The hot electrons produced by either process diffuse in through the atmosphere along the radial thermal gradient to heat the surface of the pellet, which is being continuously cooled by ablation: the same process that cools the heat shield of a space vehicle reentering the atmosphere. As the ablated material speeds outward it generates an equal and opposite force (according to Newton's Third Law) that drives the implo-



PELLET OF DEUTERIUM-TRITIUM (DT) FUEL IMPLODES when heated symmetrically by focused laser beams. Maximum absorption of light takes place at the critical surface, a narrow region in the low-density atmosphere surrounding the pellet. "Hot" electrons transport energy inward through the rest of the atmosphere to heat and ablate pellet's surface. Ablated material explodes outward as the reaction force accelerates pellet inward.

sion. In effect the system is a laser-powered spherical rocket whose payload is the rapidly contracting fuel pellet. The energy efficiency of the rocket implosion is low—less than 10 percent—because the exhaust velocity (of the ablated material) is much higher than the vehicle velocity (the imploding pellet).

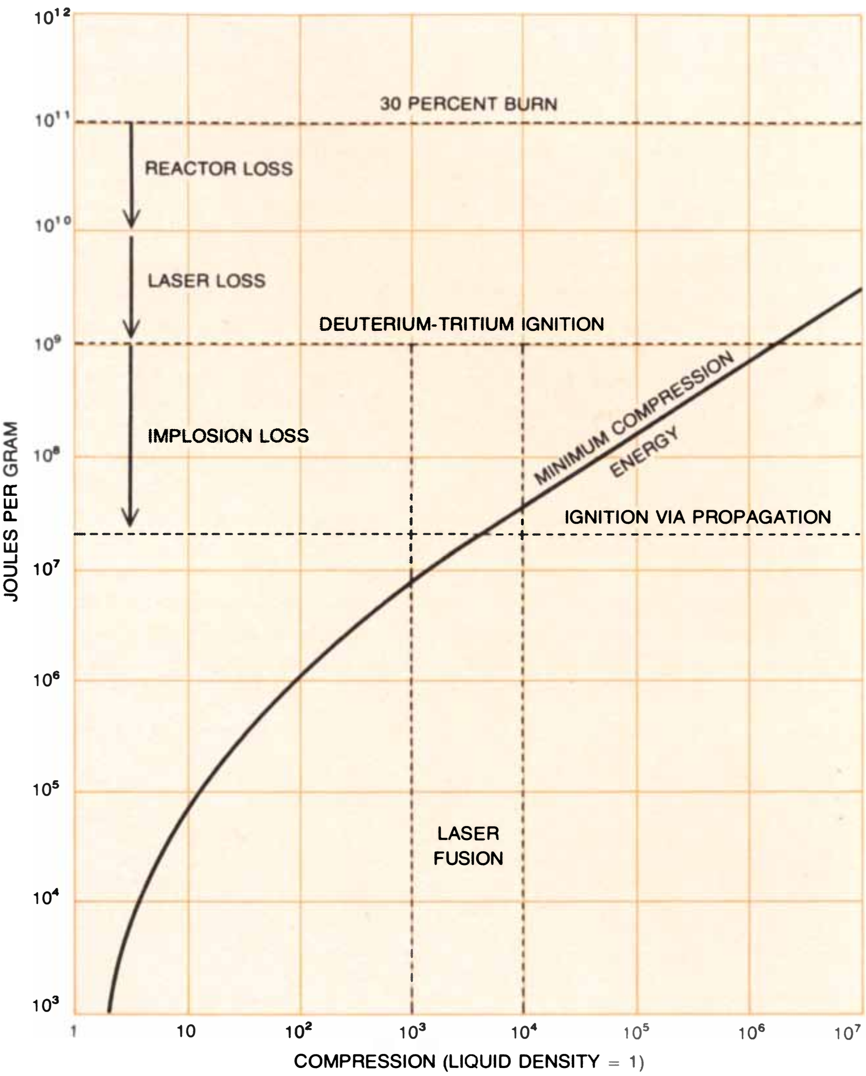
The imploding matter is accelerated inward to a velocity nearly 50 times higher than earth-escape velocity and consequently collapses inward until the pressure generated in the compressed matter finally brakes the implosion. In combination these processes multiply the power per unit area by 14 orders of magnitude from 10^5 watts per square centimeter in the laser-pumping system to 10^{19} watts per square centimeter in the implosion [*see illustrations on page 29*]. At the same time energy is transferred from the laser photons to the deuterium and tritium ions that begin to react.

Unless the pellet is compressed carefully, pressures substantially higher than 10^{12} atmospheres will be required. Thus it is necessary to carefully tailor the pressure-v.-time history of the implosion by tailoring the time history of the laser pulse. For a solid DT pellet the optimum driving pressure increases from 10^6 to 10^{11} atmospheres in the space of 10^{-8} second. Under these conditions the implosion velocity is always comparable to the local speed of sound in the pellet. As a result most of the pellet is compressed without significant heating, although the central region is deliberately driven to ignition temperature.

It is well known that investigators pursuing the magnetic-confinement approach to fusion control have had to contend with various kinds of plasma instability that cause magnetic bottles to leak. One may ask: Is the laser approach beset by similar problems? The answer is yes, but the difficulties appear manageable. Plasma instabilities can indeed be excited when very-high-power lasers are focused onto the atmosphere surrounding a pellet.

The oscillating electric field of the laser light accelerates the plasma electrons and ions in opposite directions so that they acquire a relative oscillating velocity. If the velocity is high enough, this motion leads to unstable plasma waves that grow with time. The effect can be compared to the way high winds whip up waves on the ocean. Although this intense plasma turbulence absorbs laser light efficiently, it may also generate extremely energetic (suprathermal) electrons. These electrons may penetrate deep into the pellet, heating it prematurely and thus making the compression more difficult. The phenomenon is referred to as preheat. The suprathermal electrons may also have such a long range that collisions with "colder" electrons become less frequent, so that the rate of heat transfer between the atmosphere and the pellet drops and the pellet fails to implode properly. This is called decoupling.

In order to avoid preheat and decoupling the first laser implosion experiments are being designed to operate with laser intensities that do not exceed the plasma instability thresholds. The thresholds can be increased by seeding the pellet with traces of materials of high atomic number (to increase the collision frequency) and by using lasers with shorter wavelength and wider bandwidth. Another expedient will be to use hollow pellets, since they allow a lower ablation pressure to act for a longer time and over a larger area and volume than solid pellets. Consequently less driving pressure, and hence less laser intensity, is needed to implode hollow pellets to the required high density. (It should be



FUEL COMPRESSION NEEDED FOR LASER FUSION is determined by anticipated efficiencies of the laser (10 percent), the implosion (5 percent) and the electric-generating system (effectively 10 percent, since not more than a third of the electric energy produced should be used to pump the laser). About 10^9 joules per gram is needed to ignite DT and about 10^{11} joules per gram is released if 30 percent of the fuel actually fuses. The energy gain of 100 is not large enough to accommodate the various efficiencies listed ($.10 \times .05 \times .10$ equals .0005, or .05 percent net efficiency). Fortunately in a highly compressed fuel pellet the fusion energy released by a small portion of the pellet serves to ignite the rest by radial propagation, thereby reducing the laser input required by a factor of about 100. Thus in the compression range between 10^3 and 10^4 the span between the minimum laser-input energy (about 2×10^7 joules per gram for the compression of relatively cold DT) and the energy output at 30 percent burn efficiency is large enough to make fusion power feasible.

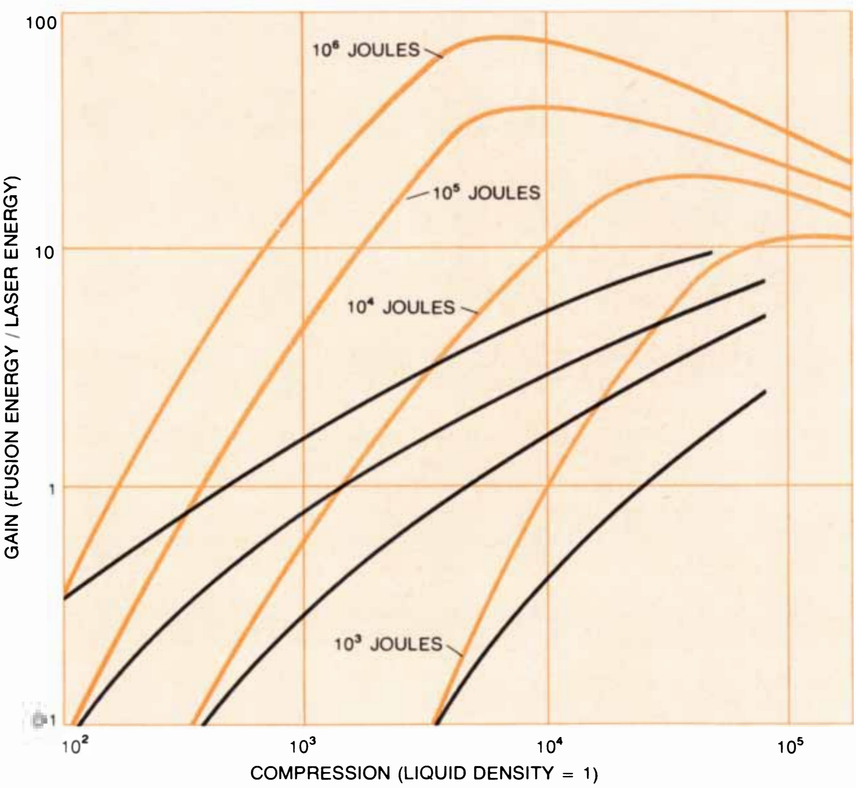
possible to make hollow liquid shells of DT by a process analogous to blowing bubbles.)

If one wishes to compress fusion fuels by a factor of 10,000, how spherically symmetrical must the implosion be? In the process the radius of the pellet is reduced by a factor of 20. Any point-to-point variation in the implosion velocity will produce a proportional error in the distance the imploding material travels. Thus an error of one part in 20 in velocity will produce an error as large as the compressed radius, with the result that the imploded pellet is no longer a sphere. In order to achieve a high degree of spherical symmetry in a twentyfold reduction in radius, temporal and spatial errors in the applied implosion pressure must be less than 1 percent.

The required implosion symmetry is achieved by irradiating the pellet as uniformly and synchronously as possible with multiple laser beams. The atmosphere that rapidly develops around the irradiated pellet does the rest: electron scattering in the atmosphere effectively filters out the remaining nonuniformities in the laser input. Overall the effect of the atmosphere on the implosion symmetry is comparable to having an earth with a hazy atmosphere 4,000 miles thick, heated by 12 symmetrically placed suns, each with an apparent diameter many times larger than that of our actual sun. Under these conditions the illumination would be so uniform that shadows would be virtually eliminated. Two-dimensional computer calculations show that the atmosphere around the pellet can reduce irradiation errors by as much as a factor of 100. As a result it should be possible to achieve sufficient sphericity to guarantee ignition and propagation of a fusion burn front in a compressed pellet.

Highly complex computer programs have been developed to calculate laser-driven implosions and thermonuclear microexplosions. The programs are used to design fusion pellets, to provide guidelines for the laser designers and to calculate laser-plasma experiments. The largest and most complex of the programs, named LASNEX, has been developed at the Lawrence Livermore Laboratory. The programs provide for the calculation of the transport and interactions of laser photons, electrons, ions, X rays and fusion reaction products, together with the magnetic and electric fields and the hydrodynamic behavior of the pellet [*see illustrations on pages 25 and 30*].

For a given laser energy the gain (the ratio of fusion energy released to laser energy input) is determined by four factors: the burn efficiency, the ignition energy (modified for propagation), the compressional energy and the implosion efficiency. Our mathematical model of a laser-fusion reactor indicates that a gain of 75 will be required if laser efficiencies of 10 percent can be achieved. To get this gain the laser input energy must reach 300,000 joules in a carefully shaped pulse lasting approximately a nanosecond. At this input energy the gain peaks when the compression reaches 10,000 times liquid density [*see illustration at left*]. Given a 300,000-joule laser operating at 10 percent efficiency, one may hope eventually to construct a 1,000-megawatt (one gigawatt) power plant in which laser pulses initiate 100 fusion microexplosions per second.



LASER ENERGY NEEDED FOR FUSION REACTOR is sharply reduced by propagation of the fusion reaction that can be realized in a highly compressed pellet. Curves show the ratio of thermonuclear energy released to laser energy invested for lasers of various energies, as influenced by pellet compression. Curves in color assume propagation; black curves assume uniform ignition (no propagation). The gain decreases at very high compressions because the energy required for compression becomes excessive. Otherwise the gain increases with compression because the DT fuel burns with increasing efficiency and less of the pellet need be ignited to sustain radial propagation. For successful power production a laser of 10 percent efficiency must yield 3×10^5 joules to achieve a desirable gain of about 75 : 1. About 95 percent of the 3×10^5 joules is consumed in ablating the fuel pellet down to a few tenths of a milligram. This leaves about 1.5×10^4 joules (or a few times 10^7 joules per gram) in the pellet, having supplied the energy needed for compression and for triggering ignition. The various curves are computed for optimum pellet weight and laser pulse shape.

The lasers needed for this job do not exist today. We believe, however, that the technology is available to prove the scientific feasibility of laser fusion by demonstrating a gain of 1 or greater. Calculations indicate that laser energies of the order of 5,000 joules will be needed. At least four different types of laser are capable of generating pulses exceeding one joule in one nanosecond or less. In order of historical development their active elements are ruby, neodymium glass, carbon dioxide and iodine. The first two are solid-state systems; the other two are gaseous ones.

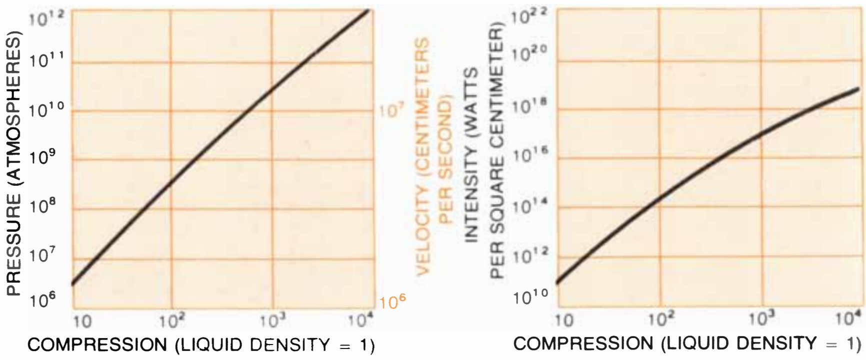
The ruby laser must be ruled out for fusion purposes because its overall efficiency when it is operated in short pulses is barely .001 percent. Moreover, large pieces of high-optical-quality synthetic ruby are nearly impossible to obtain and are quite expensive.

Glass doped with neodymium and "pumped" by xenon flash lamps emits infrared radiation with a wavelength of 1.06 micrometers. The operating efficiencies of short pulses are in the range of .02 to .1 percent, which puts the neodymium-glass laser well within the realm of feasibility for laboratory experiments. Large pieces of neodymium glass of high optical quality are costly but available. One of the attractive features of this laser is that its output wavelength can be halved to .53 micrometer by exploiting the effect known as second-harmonic generation. At high energies workers at the French CEA Limeil facility have achieved efficiencies of better than 50 percent in converting from a wavelength of 1.06 micrometers to .53 micrometer.

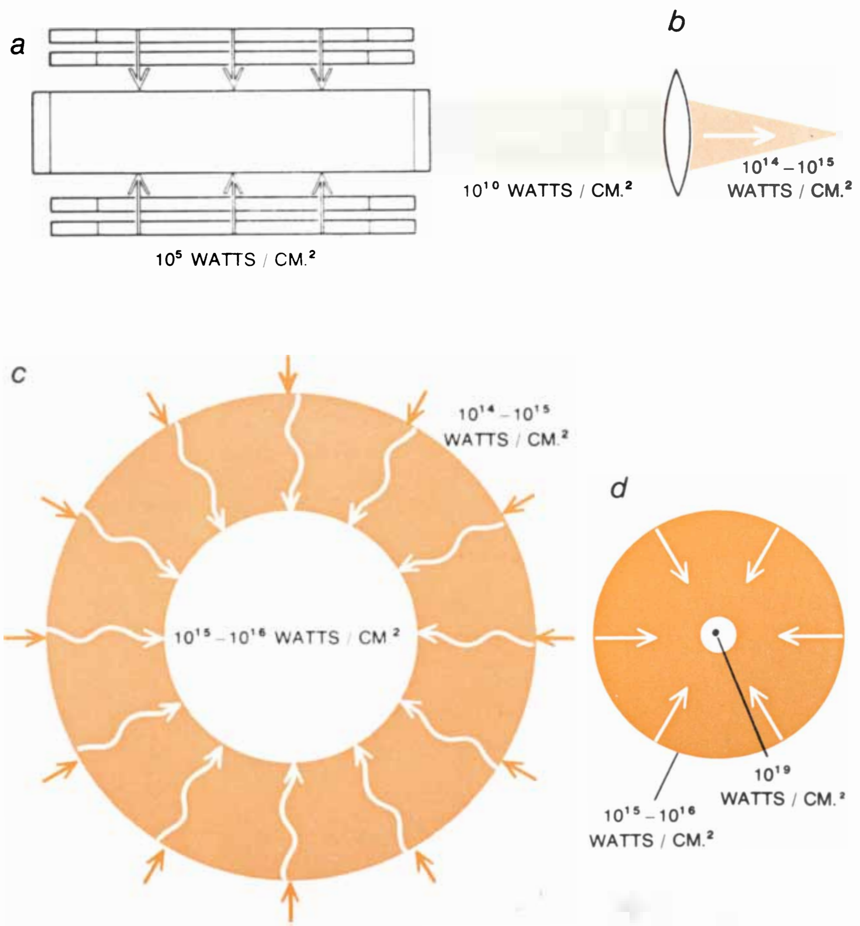
Fourth-harmonic generation, giving rise to ultraviolet radiation at a wavelength of .265 micrometer (2,650 angstroms), at overall efficiencies approaching 25 percent should be possible. Thus a single laser system can be used to explore the effects of different wavelengths on the plasma-heating process. In various laboratories throughout the world neodymium-glass laser systems have achieved impulse energies in the 100-to-1,000-joule range with pulse durations of between .1 nanosecond and two nanoseconds. During the short pulses the unfocused peak power output reaches .1 terawatt to three terawatts (10^{11} to 3×10^{12} watts).

During the past two years the technology of carbon dioxide lasers has been advancing rapidly. The laser operates at 10.6 micrometers in the infrared. Pulses of 100 to 200 joules have been generated with pulse duration of one nanosecond to two nanoseconds (at a peak power level of 100 gigawatts). Short-pulse efficiencies of 1 to 2 percent have been achieved, and values of 5 to 10 percent seem attainable. Pulse durations of less than one nanosecond have not yet been achieved. The major effort with carbon dioxide lasers in the AEC laser-fusion program is being conducted by our colleagues at the Los Alamos Scientific Laboratory. They are currently undertaking the development of a 10,000-joule, one-nanosecond carbon dioxide laser.

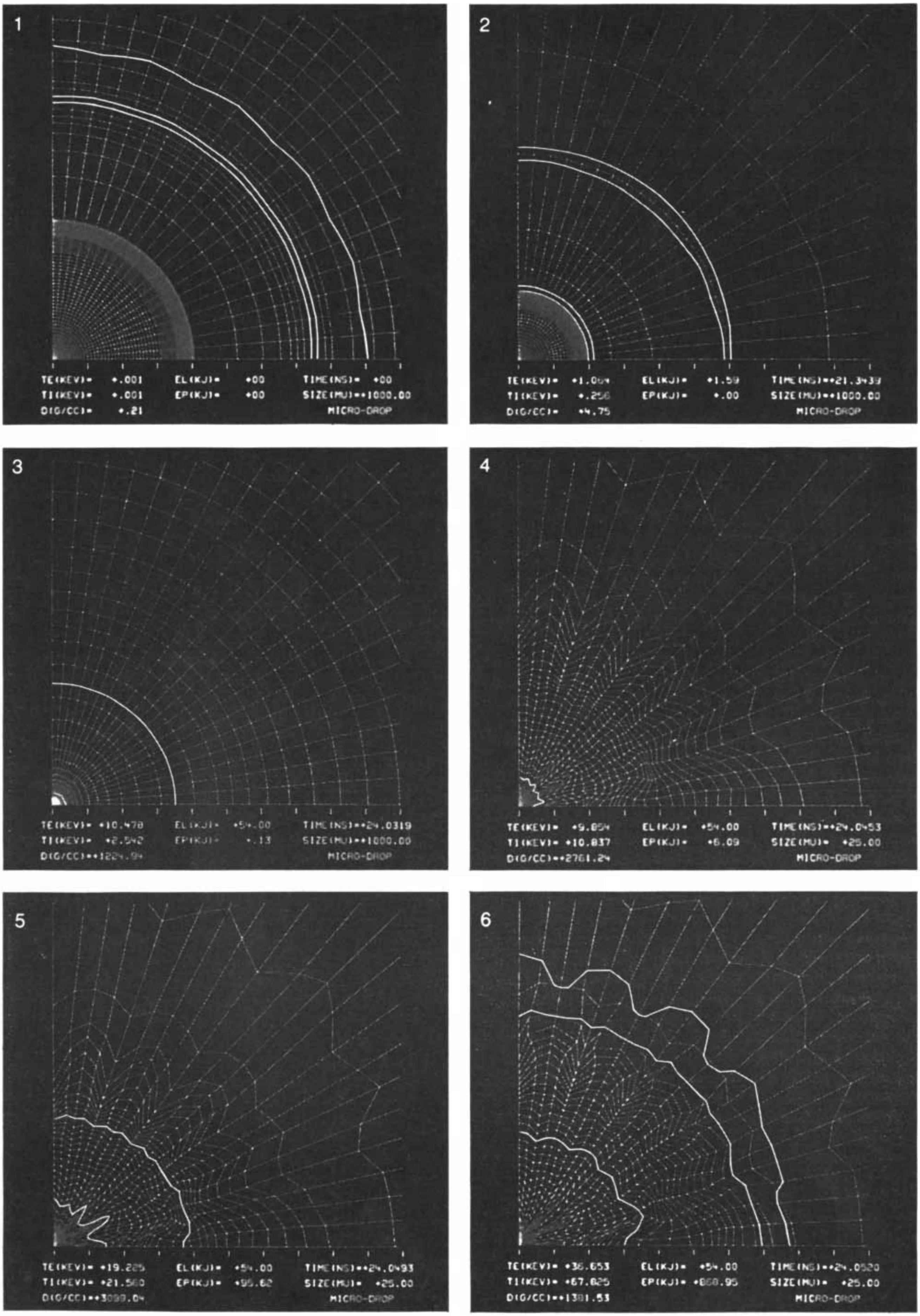
The iodine laser operates at a wavelength of 1.34 micrometers. Atomic iodine is prepared in the proper excited state by the photodissociation of gaseous compounds such as iodotrifluoromethane (CF_3I) with xenon flash lamps. The system is being intensively studied at the Max Planck Institute for Plasma Physics at Garching in West Germany. As in the case of carbon dioxide, the laser medium is cheap and therefore attractive. Iodine lasers have demonstrated an efficiency of about .5 percent. They are hard to control, however, because the iodine system has a high gain coefficient.



COMPRESSION OF FUEL IN LASER FUSION requires the imposition of pressures exceeding those in the center of the sun: approximately 10^{12} (one trillion) atmospheres. At that pressure liquid isotopes of hydrogen can be compressed by a factor of 10^4 (*curve at left*). In imploding a fuel pellet to that density radiant energy of the laser is converted to kinetic form when the matter is accelerated to high velocity inward and then is transferred to internal form when the matter is quickly brought to rest at the center. The minimum implosion velocity required to reach a compression of 10^4 is about 3×10^7 centimeters per second, or a thousandth the speed of light (*also curve at left*). The product of pressure and velocity gives an intensity (power per area) for any compression. To achieve a compression of 10^4 the required intensity is 10^{19} watts per square centimeter (*curve at right*).



ESCALATION OF POWER INTENSITY to 10^{19} watts per square centimeter can be achieved in four stages. The laser amplifier system (*a*) concentrates energy in both space and time from 10^5 to 10^{10} watts per square centimeter. Focusing of the laser beam (*b*) provides another factor of 10^4 to 10^5. The atmosphere around the target pellet acts as a thermal lens (*c*) by transporting the input energy to the smaller surface of the ablating pellet, thus increasing the intensity to 10^{15} to 10^{16} watts per square centimeter. The implosion (*d*) acts as both a hydrodynamic lens and a switch. Material motion concentrates the kinetic energy into a shrinking area. Finally the kinetic energy is converted to internal energy in a much shorter time than the kinetic energy is generated by the applied implosion pressures. The intensity of 10^{19} watts per square centimeter needed for fuel compression is thus achieved.


1
TE(KEV)= +.001
TI(KEV)= +.001
D(G/CC)= +.21
EL(KJ)= +00
EP(KJ)= +00
TIME(NS)= +00
SIZE(MU)=+1000.00
MICRO-DROP
2
TE(KEV)= +1.064
TI(KEV)= +.256
D(G/CC)= +4.75
EL(KJ)= +1.59
EP(KJ)= +.00
TIME(NS)=+21.3439
SIZE(MU)=+1000.00
MICRO-DROP
3
TE(KEV)= +10.478
TI(KEV)= +2.542
D(G/CC)=+1224.94
EL(KJ)= +54.00
EP(KJ)= +.13
TIME(NS)=+24.0319
SIZE(MU)=+1000.00
MICRO-DROP
4
TE(KEV)= +9.854
TI(KEV)= +10.837
D(G/CC)=+2761.24
EL(KJ)= +54.00
EP(KJ)= +6.09
TIME(NS)=+24.0453
SIZE(MU)= +25.00
MICRO-DROP
5
TE(KEV)= +19.225
TI(KEV)= +21.560
D(G/CC)=+3099.04
EL(KJ)= +54.00
EP(KJ)= +95.62
TIME(NS)=+24.0493
SIZE(MU)= +25.00
MICRO-DROP
6
TE(KEV)= +36.653
TI(KEV)= +67.825
D(G/CC)=+1381.53
EL(KJ)= +54.00
EP(KJ)= +868.95
TIME(NS)=+24.0520
SIZE(MU)= +25.00
MICRO-DROP

A high gain coefficient means that the atoms that have been pumped to an excited state are very easily stimulated to emit their characteristic radiation and drop to a lower energy state. Since short-pulse laser systems must store large amounts of energy prior to pulse amplification, high gain coefficients in large-aperture amplifiers present two difficult problems. The first is termed superfluorescence. This is simply the normal fluorescence emitted spontaneously by the excited laser material, amplified by the gain of the material itself. If the specific gain coefficient is high, the superfluorescence loss rate could exceed the maximum available pumping rate. Thus energy cannot be efficiently stored in the laser amplifier.

The second major problem results from feedback in the laser amplifier of the superfluorescent radiation. The feedback generates parasitic laser oscillations within the laser amplifier. The stimulated emission rate from such parasitic oscillations could far exceed any pumping rate and therefore also limits the maximum achievable stored energy. Although it seems almost paradoxical, for large laser-amplifier systems one needs laser transitions that have low gain coefficients.

At this writing neodymium-glass laser systems have been chosen by all but one of the major laser-fusion laboratories in various countries as the prime vehicle they will use to prove the scientific feasibility of laser fusion. At the Lawrence Livermore Laboratory we are developing several neodymium-glass laser systems for short-run fusion experiments. The systems, which are all quite similar, consist of a low-energy laser oscillator and a cascade of laser amplifiers [*see upper illustration on next two pages*]. The oscillator produces a train of five to 30 pulses whose duration is controllable over a range of 20 to 1,000 picoseconds (trillionths of a second). A fast electro-optical shutter is used to switch out a single pulse for amplification through the system. The pulse, which contains only about 10^{-3} joule of laser energy, is shaped spatially (and also temporally, if desired) prior to amplification. Since total single-beam energies of 100 to 1,000 joules are desired, the cascade of amplifiers must produce an optical gain of 100,000 to a million.

The maximum gain at each stage in the cascade is fixed by nonlinear optical effects, which limit the maximum intensity of the laser radiation that can be propagated through matter. The nonlinearity arises because the intense electric field of the light wave produces a small but significant increase in the index of refraction of the optical material. The destruction of spatial coherence and the self-focusing of the laser beam resulting from such nonlinear effects limit the optical power density in glass to between five and 20 gigawatts per square centimeter. If the system is to stay below this limit, the amplifiers must be increased in aperture (beam area) from stage to stage. The first factor of 1,000 in gain can be provided by neodymium-glass rod amplifiers between two and four centimeters in diameter. The last factors of 100 to 1,000 in gain are achieved by amplifiers with diameters ranging from 10 to 30 centimeters. The large-area units are usually constructed from a series of neodymium-glass disks rather than from a single fat rod. Loss of light due to surface reflection from the disks is eliminated by inclining them at a special angle known as the Brewster angle and by properly polarizing the laser beam. The arrangement makes it possible for large amplifiers to be pumped uniformly across a large aperture. The disks are totally surrounded by xenon flash lamps, which pour their energy into the angled faces of the disks.

At Livermore we have successfully tested amplifiers up to 20 centimeters in aperture. A two-beam, one-terawatt (10^{12}-watt) system is currently being used in initial studies of laser-target interaction. The final aperture of this system is nine centimeters. A five-to-10-terawatt system consisting of six to 12 laser beams in a spherically symmetrical array will be in operation in 1975. This system should generate implosions of medium to high compression and facilitate the development of implosion diagnostic techniques.

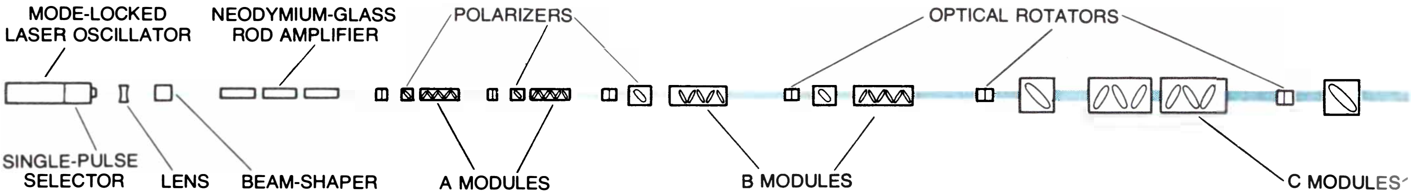
The next step will be the design and construction of a spherically symmetrical facility whose combined beams before focusing will carry between 50 and 100 terawatts of radiant energy. When the beams are focused, they will be capable of delivering 10,000 joules of energy to a pellet of fusion fuel within the span of 100 to 500 picoseconds. With some modification it should be possible to deliver 50,000 joules in a pulse of a few nanoseconds. The facility will cost some $20 million.

The big laser will consist of 12 or 20 parallel amplifier systems driven by a single oscillator to ensure nearly perfect synchronization [*see lower illustration on next two pages*]. The choice of 12 or 20 beams to achieve the 10,000 joules of energy on the target will be dictated by considerations of implosion symmetry and by the cost of finishing large optical elements. The project is currently under way, with the development of one of the amplifier chains being nearly complete. With this facility, scheduled for completion early in 1977, we believe it will be possible to prove the scientific feasibility of laser fusion by approaching or exceeding the condition where thermonuclear yield equals input of optical energy.

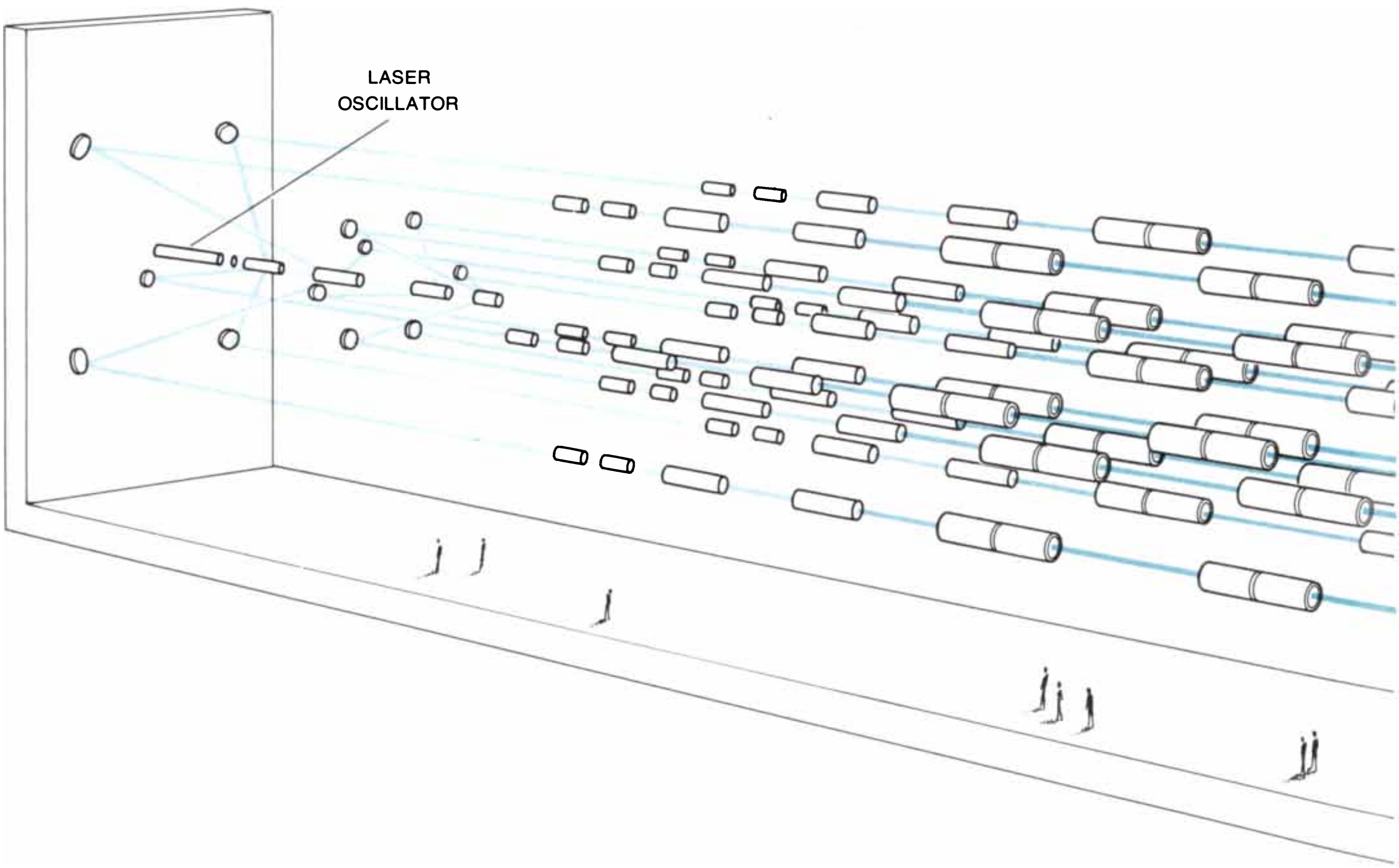
EXPLOSION OF DT PELLET is modeled by a Livermore computer program called LASNEX. Six frames from a computer motion picture of a typical calculation are shown on the opposite page. (The last frame shows the same stage of the explosion as the full-circle representation on page 25.) Counters under each frame indicate maximum electron temperature (*TE*) and maximum ion temperature (*TI*) in kilovolts (KeV). One KeV equals about 10^7 degrees K. *D* is maximum density in grams per cubic centimeter. *EL* is cumulative input of laser energy and *EP* is output of fusion energy, both in kilojoules. Time is given in nanoseconds and the width of the frame is given in micrometers. (Note fortyfold change in scale between frames No. 3 and No. 4.) White lines in first three frames are isotherms of electron temperature; those in next three frames are isotherms of ion temperature. In Frame No. 1, at zero time, a 400-micrometer pellet of normal liquid density is surrounded by a low-density atmosphere extending beyond 1,000 micrometers. Laser light is being absorbed between the two closely spaced isotherms, at a radius of 750 micrometers and a temperature of 3×10^6 degrees. The isotherm at 900 micrometers is at 10^6 degrees. Hot electrons have not yet penetrated to the pellet surface. In Frame No. 2 the pellet has imploded to 200 micrometers and laser light is being absorbed between a pair of 10^7-degree isotherms at 600 micrometers. In Frame No. 3 the full 54 kilojoules of laser energy has been delivered and the pellet is compressed to a radius of less than 15 micrometers. In Frame No. 4 the onset of thermonuclear ignition is indicated by an ion-temperature isotherm at 10^8 degrees near the center. In Frame No. 5, four picoseconds later, the 10^8-degree burn front has propagated across most of the pellet and a ragged isotherm twice as hot has started moving outward from center. In Frame No. 6 an isotherm at 5×10^8 degrees has traveled 7.5 micrometers from the center and half of the fusion energy has been released. The remainder will be generated in the next 12 picoseconds (.012 nanosecond). Such a computer calculation requires several hours on the world's most powerful computer, the CDC 7600. Sixty variables are computed at 2,000 points in space and 10,000 points in time to produce more than a billion numbers.

Although a neodymium-glass laser system should be adequate to demonstrate the feasibility of laser fusion, we know that it will not serve for a practical fusion reactor. A power plant of one-gigawatt electrical capacity burning deuterium-tritium fuel will require a laser of approximately 300,000 joules operating at a repetition rate of 100 pulses per second (or several lasers operating at a lower repetition rate) with an efficiency of 10 percent. If the fuel is in the form of solid pellets, the laser must deliver its 300,000 joules at a peak power level of 1,000 terawatts; if the fuel is in the form of hollow pellets, the power level can be reduced to 100 terawatts or less.

If the laser system has an efficiency of 10 percent, waste heat will be generated in the laser medium at an average rate of 300 megawatts. The laser medium must therefore be a fluid, so that the waste heat can be removed by high-speed flow rather than slow thermal conduction. The large nonlinear optical coefficient associated with most liquids argues strongly for a gaseous laser medium since the laser must operate at very high optical power densities. Gases, on the other hand, are subject to optical breakdown from intense beams. A simple theory of the process indicates that the threshold for optical breakdown varies inversely with the square of the optical wavelength. Thus shorter wavelengths are preferable to longer ones. For a wavelength of 10.6 micrometers the measured breakdown threshold near atmospheric pressure is on the order of 10^9



ONE CHAIN OF LIVERMORE LASER SYSTEM will be 165 feet long and will consist of a small laser, or oscillator, to provide the original pulse, followed by 11 stages of amplification. The first three stages are solid rods of neodymium glass. Subsequent stages incorporate neodymium glass in the form of disks of increasing diameter. Disk amplifiers designated *A*, *B*, *C* and *D* have respective

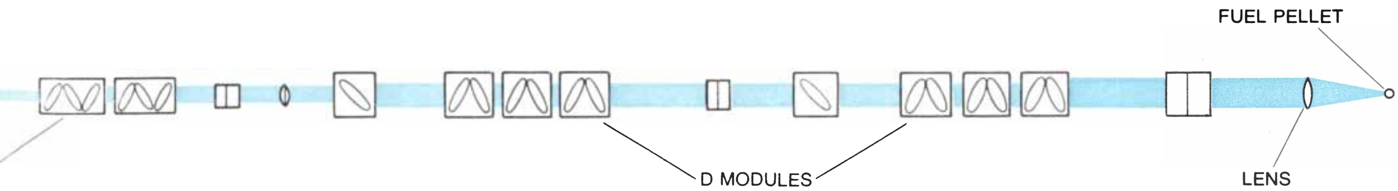


MULTIBEAM LASER FACILITY, scheduled for completion at the Lawrence Livermore Laboratory in 1977, is designed to prove the feasibility of initiating thermonuclear microexplosions by implosion of pellets of hydrogen isotopes to ultrahigh densities. It will contain 12 amplifier chains like the one shown in the illustration above this one or, depending on cost analyses, 20 somewhat

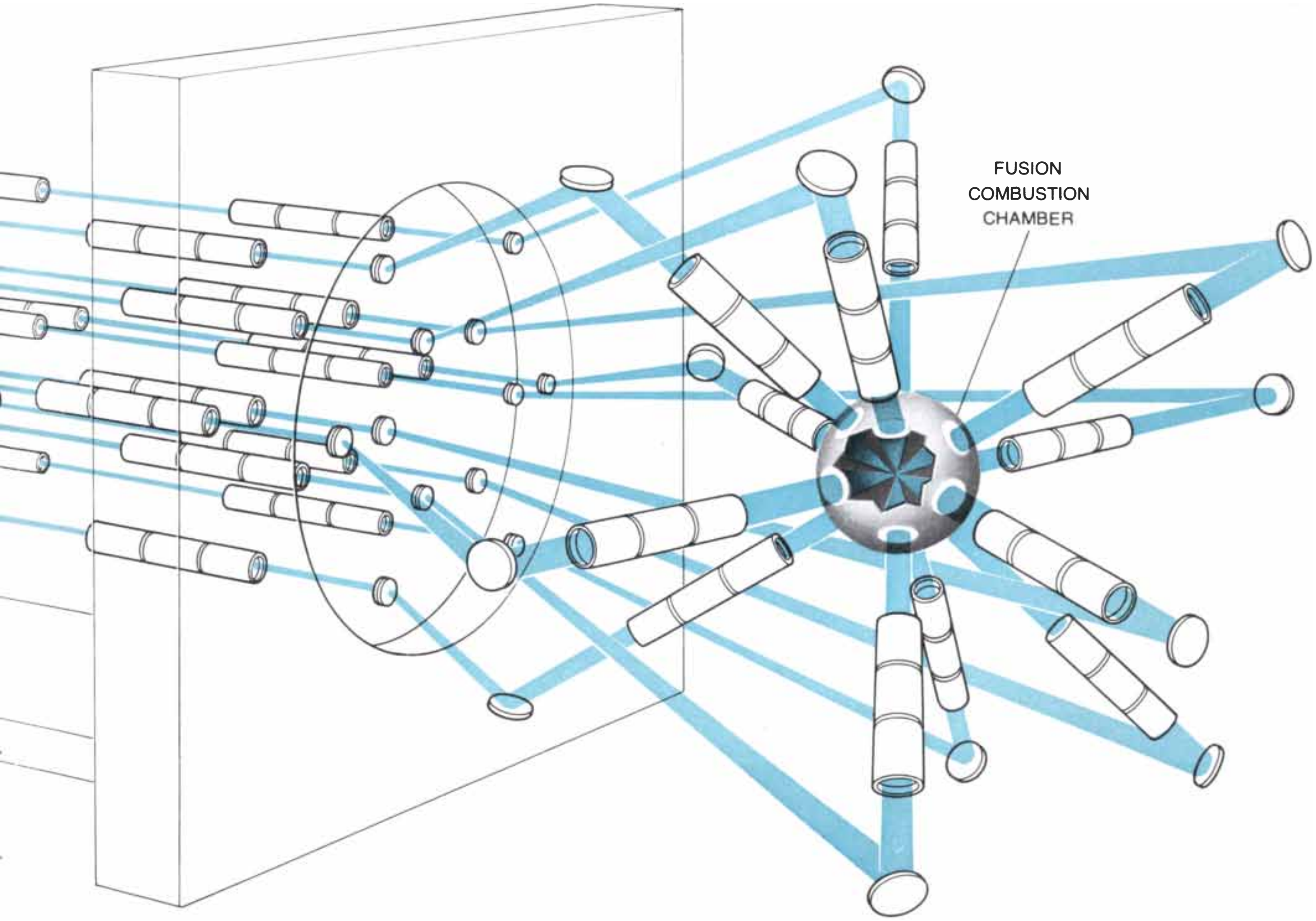
watts per square centimeter; for a wavelength of 1.06 micrometers the threshold rises by a factor of 100. Thus optical breakdown of the laser medium and plasma instabilities in the target both argue for radiation of short wavelength. If, however, the wavelength is so short that two photons can be absorbed simultaneously by an atom or molecule of the laser medium, with consequent photoionization, the laser beam will be strongly absorbed rather than amplified. The optimum wavelength for a laser fusion power plant therefore lies between approximately .3 and .8 micrometer.

If the pressure of the gas medium in our hypothetical laser system does not exceed a few atmospheres, the system should be capable of a power level of perhaps 5×10^{10} watts per square centimeter. Hence a total power of 10^{14} watts (for a hollow-pellet target) or 10^{15} watts (for a solid target) calls for a laser system with a final aperture of 2,000 or 20,000 square centimeters respectively. If we use 12 laser beams to achieve the required irradiation symmetry, the individual final amplifiers need only be from 15 to 45 centimeters in diameter. Such sizes are well within the present state of optical technology.

To summarize, the laser system for a practical fusion power plant must meet the following criteria. The operating wavelength should lie between 3,000 and 8,000 angstroms, the medium must be a gas at pressures below a few atmospheres and, to limit superfluorescence and parasitic oscillations, the gas must



apertures of five, nine, 20 and 30 centimeters. Each disk amplifier includes a polarizer and a Faraday optical rotator for rotating the polarization 45 degrees. The polarizer-rotator combination ensures that light reflected from the target will not propagate back into high-gain amplifier system. The chain will amplify the millijoule (10^{-3} joule) output of the oscillator by a factor of 10^5 or 10^6.



less powerful chains. The combined beams will be capable of irradiating test pellets with 10 kilojoules of optical energy in a period of 100 to 500 picoseconds, equivalent to a peak power output of some 20 to 100 terawatts. The laser power, energy and pulse shape will be highly flexible in order to test a variety of pellet designs. It is estimated the completed facility will cost $20 million.

have a low specific-gain coefficient. Accordingly we must look for a weakly allowed electronic transition in an atom or a molecule. Finally, an efficiency of roughly 10 percent must be attainable.

The search for a new laser medium is currently under way at our laboratory, at the Los Alamos Scientific Laboratory, at the Sandia Laboratory in Albuquerque—all operated for the AEC—and at many other laboratories throughout the world. One class of electronic transitions being studied is typified by the one at .5577 micrometer in atomic oxygen, the spectral line responsible for the green color of the aurora borealis. Other elements that share the same column with oxygen in the periodic table (sulfur, selenium and tellurium) show similar transitions, which are also being examined. We believe that whereas the ideal laser medium for a fusion power plant has not yet been found, its discovery and development appear to be a straightforward, albeit time-consuming, endeavor.

In addition to a substantial advance in laser technology, a laser-fusion power plant will require the solution of many other technological problems. The high-efficiency detonation of fusion-fuel pellets for practical electricity generation will occur on a time scale of 10^{-11} second or less. Since the energy released will be at least 10^{7} joules, the peak rate of fusion-power production will be at least 10^{18} watts (10^{7} divided by 10^{-11}). This rate (which, to be sure, is intermittent) is a million times greater than the power of all man-made machinery put together and is about 10 times greater than the total radiant power of sunlight falling on the entire earth. The technological challenge of laser fusion is how to wrap a power plant around fusion microexplosions of these astronomically large peak powers that can endure their effects for dozens to hundreds of times every second for many years. Perhaps surprisingly, it appears possible to do so.

The energy output of a deuterium-tritium microexplosion is carried by neutrons, X rays and charged particles. Each of these radiations presents a special threat to the survival of the "first wall," or innermost surface, of the combustion chamber. The individual particles of radiation emerge with a spectrum of energies and a wide range of velocities, up to and including the velocity of light. The spread in velocities is fortunate because it means that the particles do not all hit the first wall at the same time. This greatly lowers the peak rate at which energy is deposited and makes possible the long-term survival of the wall.

THERMONUCLEAR POWER PLANT GENERATION	FUSION FUELS		FUSION PRODUCTS
FIRST	DEUTERIUM + TRITIUM	→	HELIUM 4 + NEUTRON
SECOND	DEUTERIUM + DEUTERIUM	↗	HELIUM 3 + NEUTRON
		↘	TRITIUM + HYDROGEN
THIRD	BORON 11 + HYDROGEN	→	HELIUM 4 (THREE ATOMS)

FUELS FOR FUSION POWER PLANTS may become cheaper and cleaner-burning if the development of laser technology makes it possible to achieve densities still higher than those needed for fusion of deuterium and tritium. Since the fuels for the more advanced plants burn less rapidly, they require for efficient burning that the product of pellet radius and density be higher by a factor of 10 to 100 than is needed for DT. System performance is

The threat to the wall presented by neutrons, which carry about three-fourths of the fusion energy released, is perhaps the subtlest. Although high-energy neutrons pass through matter much more readily than any of the other radiations do, they nonetheless leave scars of their passage in the form of dislocated and disintegrated atoms. The dislocations, which tear atoms out of their position in a crystal lattice, result from collisions like those of billiard balls as the neutrons career through the first wall. The neutrons also disrupt occasional atomic nuclei in the wall by knocking out a proton, a neutron, a deuteron (deuterium nucleus), a triton (tritium nucleus) or a helium nucleus. All the particles knocked out (with the exception of secondary neutrons) form gas atoms within the wall as they slow down from their birth events. The atoms eventually aggregate into tiny gas bubbles whose pressure can rise to thousands of pounds per square inch before they finally rupture the surface of the wall. Moreover, the nucleus that remains behind after its disruption is usually radioactive and is added to the radioisotope inventory of the power plant, which consists primarily of radioisotopes created when the primary neutrons are eventually captured by the nuclei of the atoms forming the wall.

Neutron-degradation problems place an upper limit on the lifetime of the first wall in all DT-burning fusion power plants, whether they are of the laser or the magnetic-confinement type. Preliminary experiments indicate that the best first-wall materials may survive exposure to 14-MeV neutrons (the kind released by deuterium-tritium fusion) for a few dozen years, provided that the neutron power flux is limited to about one megawatt per square meter. This implies that any one-gigawatt DT-burning fusion reactor may need as much as 1,000 square meters of first-wall area, either in one chamber or in several chambers. In the case of magnetic confinement, increasing the chamber volume to increase the first-wall area presents a particular handicap because the magnets must be outside the shielding blanket; as they are moved outward they must be made proportionately larger (and more expensive) to generate the same magnetic field at the center of the chamber.

The threat presented to the first wall by X rays is more straightforward. Since most of the X-ray photons have energies greater than 5,000 volts, they penetrate deep into the wall and deposit their energy quite harmlessly. Although X rays with energies of less than 2,000 volts carry off less than .1 percent of the energy of the microexplosion, they deposit their energy within a micrometer of the wall's surface, very rapidly heating a thin skin of the first wall to a high temperature. The large thermal gradients and mechanical stresses so produced could cause the front surface of the first wall to flake away very slightly with each microexplosion. A possible solution is to build the first wall out of materials such as lithium, beryllium and carbon, which are relatively transparent to soft X rays. The soft X rays would therefore penetrate a considerable distance into such a surface layer, thereby heating a larger amount of mass to a lower temperature and reducing the peak stresses to acceptable levels.

The threats presented to the integrity of the first wall by charged particles streaming from the microexplosion are complex and substantial. One threat is analogous to that presented by soft X rays: the less energetic ions (thermal deuterons and tritons), also absorbed in a thin layer, are capable of producing sharp thermomechanical stresses in a thin skin of the first wall. The more energetic ions

FUEL DENSITY × RADIUS (GRAMS PER / CM.^{2})	FUEL DENSITY (GRAMS PER / CM.^{3})	REQUIRED LASER PULSE (JOULES/NANOSECONDS)	LASER EFFICIENCY (PERCENT)	PERCENT OF ENERGY RELEASED IN:		
				X RAYS	CHARGED PARTICLES	NEUTRONS
2–5	$10^3 - 4 \times 10^3$	10^5 / 3	~10	2–10	25–30	65–75
10–20	~10^4	10^6 / 1	≥10	20–50	30–60	20–45
~200	~3×10^5	10^8 / .3	≥30	30–70	30–70	<.1

optimized by raising both the fuel density and the pellet radius to obtain the desired values. This in turn requires both shorter and more energetic laser pulses, produced with higher laser efficiency. The fuels proposed for second-generation and third-generation power plants yield increasingly less of their energy in the form of neutrons. This is desirable because the amount of radioactivity produced in the walls of the combustion chamber is roughly proportional to the emission of neutrons. The reaction of a proton (the nucleus of ordinary hydrogen) and the common isotope of boron, boron 11, looks attractive for third-generation fusion power plants.

are sufficiently penetrating and few in number, so that they should produce relatively little damage.

The second major threat presented by charged particles resembles the gas-bubble problem associated with neutrons: the energetic nuclei that bury themselves in the first few micrometers of the surface quickly acquire electrons and become gas atoms that agglomerate into growing bubbles with high internal pressures. These bubbles are also capable of rupturing the surface of any solid wall in a much shorter time than the desired service life of a fusion power plant.

One possible way to deal with the various assaults on the inner surface of the reactor's first wall is to coat the surface with a regenerating layer of liquid lithium metal a few tenths of a millimeter thick [*see illustration on next page*]. The lithium could be held in place by surface tension developed against a grooved backing shell of, say, vanadium or niobium. Such a liquid skin would be essentially immune to thermomechanical stresses. Moreover, the gas atoms produced by the action of neutrons and by the electrical neutralization of low-energy ions would migrate rapidly to the inner surface of the film of liquid lithium and could be pumped away, along with the small amount of lithium vaporized during each heating pulse.

Contrary to what intuition might suggest, the purely mechanical stress imposed on the first wall by the outwardly streaming debris of a fusion microexplosion is extremely small. For example, a microexplosion producing about 10 million joules of energy, comparable to the energy released by two kilograms of chemical high explosive, imposes a momentary pressure on the thermonuclear combustion chamber no greater than what would be produced by a medium-size firecracker. The explanation for this seeming paradox is that the mass of the fusion pellet is only about a milligram, or a factor two million less than two kilograms of chemical explosive. Since the impulse of a blast wave is proportional to the square root of the product of the energy and the mass associated with the wave, the tiny mass of the pellet compared with that of a chemical explosive implies a reduction of 1,400 in the impulse of a fusion blast wave compared with the impulse of a chemical-explosive blast wave.

The heat associated with the X-ray and plasma pulses, which constitute about a fourth of the total fusion energy from DT explosions, can be removed from the back of the first wall by a stream of coolant or the evaporator surface of a heat pipe. The neutrons, which carry about three-fourths of the fusion energy, pass through the first wall and are stopped in a neutron-absorbing blanket containing lithium; the reaction of a neutron and an atom of lithium 6 yields tritium and helium. Some of the tritium is fed back into the pellet factory for new pellet production. The absence of magnetic fields in the laser system allows the tritium-breeding blanket to be cooled by fluids, such as lithium, that are electrically conducting; conducting fluids in magnetic confinement systems present a special set of problems.

Several other considerations are important in designing the shielding blanket around a fusion combustion chamber. The laser beams and fuel pellets must be directed through the blanket on zigzag paths to prevent the escape of X rays and neutrons [*see illustration on page 37*]. The fuel pellets could be electrically charged and steered through a zigzag path electrostatically. The laser beams can be guided through similar zigzag paths by mirrors. The final mirror in each laser path would be curved to focus the light onto the pellet. The front surface of the final mirror will probably require a continuously renewable liquid-metal surface to withstand the nuclear environment. The mirror need not have high optical performance, however; it will be sufficient if most of the focused laser light falls within a two-millimeter spot from a distance of one or two meters.

The first laser-fusion reactors will probably use deuterium-tritium fuel because it calls for the smallest lasers. If tritium were not needed, however, lithium blankets for breeding tritium would not be needed. For example, with 10^6-joule lasers "straight" deuterium (DD) could be used as a fuel. The products of deuterium fusion are helium 3 and a neutron in half of the events and tritium and a proton in the other half [*see illustration on these two pages*]. Most of this tritium is burned by the time the compressed pellet has blown itself apart, but the remaining tritium ends up in the debris of the microexplosion. That tritium can then be recovered from the vacuum system that keeps the combustion chamber pumped down and is sent to the pellet "factory" to be combined with deuterium. Eventually a steady state is reached in which a small, constant amount of tritium is incorporated in each deuterium pellet, thereby facilitating its ignition.

Not having to breed tritium in the shielding blanket would simplify the blanket's construction and operation. One can visualize, for example, a neutron-shielding blanket consisting of graphite blocks perforated with cooling channels whose sole function is to transform neutron energy into heat at, say, 2,000 degrees K., which could be carried away by helium under high pressure. The system would be a fusion analogue of the promising High Temperature Gas Cooled Reactor now attaining commercial feasibility for the generation of fission power. The higher the temperature of the fluid used to drive turbogenerators, the higher the thermal efficiency of the power plant and the lower the discharge of waste heat to the environment. Since a fusion reactor would not have to contain large amounts of volatile radioactive and toxic materials, it might be operated at temperatures even higher than those to be used in the helium-cooled fission reactors.

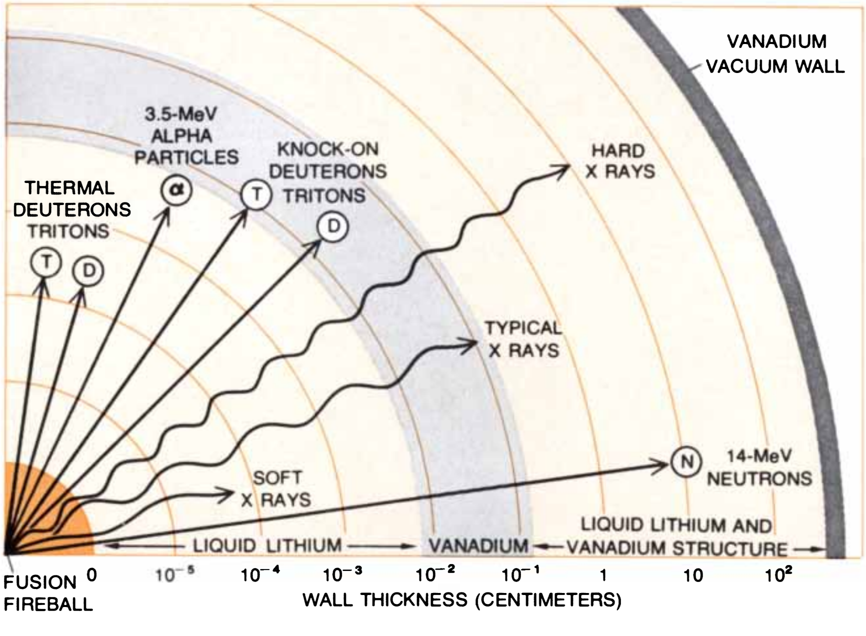
The fusion neutrons can also be used to convert nonfissionable isotopes of uranium or thorium to fissionable isotopes, just as is done in a fission reactor of the breeder type. The neutrons released by DT or DD microexplosions are sufficient to release 20 to 50 times as much energy in the form of fissionable uranium or plutonium isotopes as was released in the original fusion reaction. Such hybrid fusion-fission reactors may make their appearance before pure fusion systems become economically competitive.

Looking beyond pure deuterium fusion, what other nuclear reactions might be considered? There is one fascinating possibility. As the density-radius product of the imploded pellet is raised beyond a few hundred (which implies pellet densities greater than 100,000 times liquid density) the X rays emitted by the burning fuel can no longer readily escape and are effectively trapped. If the ultrahigh densities needed to achieve this effect can be attained, certain fusion fuels that do not ignite in the characteristic fashion may nevertheless be burned with adequate efficiency. For these applications, however, lasers with an output of 10^8 joules and an efficiency of 50 percent must be developed. Perhaps the most interesting candidate is the reaction between ordinary hydrogen and ordinary boron: boron 11. The reaction could properly be called a thermonuclear fission reaction since more particles are produced than are consumed. Its great virtue is that it converts nonradioactive reactants to nonradioactive products. The reaction between a proton and the nucleus of boron 11 takes place in three steps. In the first step the proton simply joins the nucleus of boron 11, forming carbon 12 in an excited state. The excited carbon immediately fissions into helium 4 and beryllium 8, and the beryllium 8 then promptly fissions into two more nuclei of helium 4. Thus the final products of the reaction are three energetic atoms of helium. Relatively rare side reactions, however, produce a low-energy neutron or a weakly radioactive nucleus (carbon 14) in about .1 percent of all the reactions. Even so, a power plant based on the reaction of boron 11 and a proton should produce 1,000 times less radioactive debris than one employing the DT reaction.

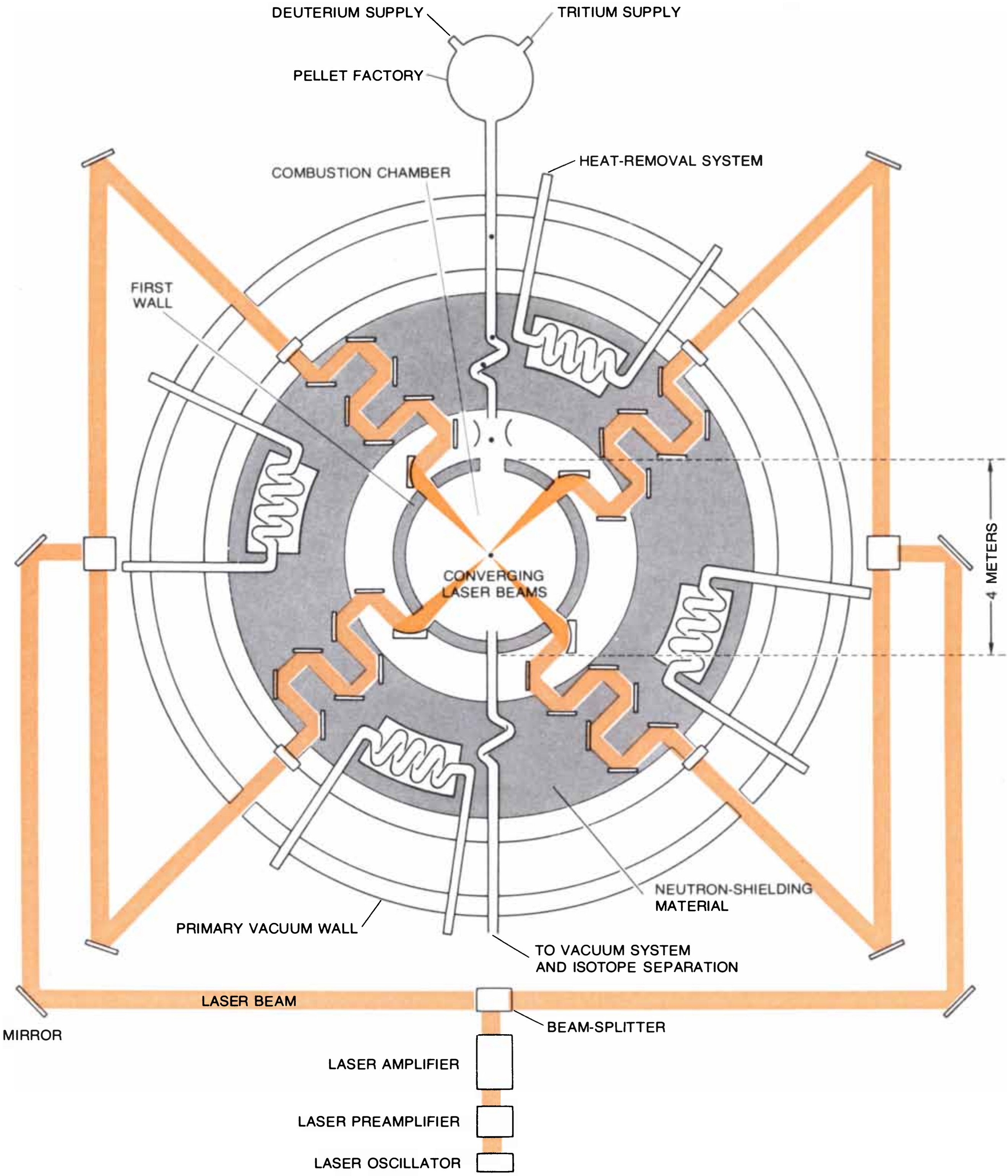
So far we have described only conventional heat-transfer methods for generating electricity in a laser-fusion power plant. Might it not be possible to capture the energy released in the microexplosion more directly? After all, we start with electric charges streaming at high velocity outward from the point of explosion. We want to end up with electric charges streaming through transmission lines in power grids. Is it really necessary to transmute and degrade the energy along the way by passing it through a steam boiler?

Perhaps not. In deuterium-burning microexplosions most of the fusion energy is released as charged particles and much of the neutron energy is deposited in the highly compressed plasma. It is well known that if a plasma is allowed to expand against a surrounding magnetic field, it will compress the field and push it outward. The compression of the magnetic field is available for direct conversion into electricity by the electromagnetic-induction principle discovered by Faraday. One simply arranges for the moving lines of magnetic force to cut through the loops of an induction coil. In experiments with laser-heated plasmas surrounded by a magnetic field it has been shown that somewhat more than 70 percent of all the energy originally put into the plasma by the laser is transferred into the magnetic field in the form of compressed lines of magnetic force. It has been estimated that under conditions more similar to those existing in deuterium-burning microexplosions 80 to 90 percent of the initial fireball expansion energy would be absorbed in compressing the magnetic field, from which it could be extracted by induction for direct conversion into electricity.

For 20 years the goal of harnessing the fusion process to power our civilization has been pursued along the avenue of magnetic confinement. Now laser fusion, an entirely independent approach, has been conceived to attain this goal. In any problem of major importance it is good to have a diversity of approaches, since such diversity substantially increases the probability of success.



FUSION RADIATIONS STRIKING FIRST WALL of a thermonuclear combustion chamber will penetrate to different depths depending on their energy. The wall must be designed accordingly to minimize damage to its structure. Roughly 70 percent of the energy released by the fusion of deuterium and tritium is carried by 14-MeV (million-electron-volt) neutrons, about 25 percent by charged particles (deuterons, tritons and alpha particles) and a few percent by X rays. Deuterons and tritons that carry only the normal fireball energy, equivalent to 10^9 degrees K., are called thermal. "Knock on" deuterons and tritons have been accelerated to energies about 100 times higher through collisions with neutrons in the fireball. The film of liquid lithium bathing the surface of the first wall will stop the softest X rays as well as all charged particles capable of aggregating into bubbles. Vanadium will probably be used as the structural material because it produces least radioactivity under neutron bombardment of any high-temperature material compatible with liquid lithium.



LASER-FUSION POWER PLANT is conceptually simple. Liquid pellets of deuterium and tritium, about a millimeter in diameter, are guided electrostatically on a zigzag path through a neutron-shielding wall until they can fall freely to the center of the combustion chamber. There the pellets are symmetrically irradiated and imploded by converging laser beams. The zigzag paths are necessary to prevent the escape of X rays and neutrons from the explosion. In first-generation fusion power plants heat will be removed by conventional heat-exchange systems, such as a circulating flow of liquid lithium, and used to make steam for driving turbogenerators. In more advanced plants it may be possible to convert a large fraction of the microexplosion energy directly into electricity. A plant capable of triggering 100 microexplosions per second, perhaps in several combustion chambers, could generate between 100 and 1,000 megawatts of electricity. For economical power generation the cost of each fusion pellet should not exceed one cent.

Population Cycles in Rodents

Field populations of many small rodents rise and fall in a cycle lasting three to four years. Associated with these fluctuations is a periodic change in the genetic constitution of the population

by Judith H. Myers and Charles J. Krebs

Legend has it that the lemmings of Scandinavia periodically overpopulate their range, then reduce their numbers by migrating into the sea. The story is not true—lemmings make no "suicide marches" into the sea—but it is not without insight. Populations of lemmings and other small rodents such as voles and field mice do fluctuate from year to year, and a population that has grown steadily for a few seasons may decline abruptly, and may even seem to disappear entirely. Moreover, it has recently been determined that migrations, although they do not end in the sea, are an essential part of the process that regulates these populations.

In many species the fluctuations are cyclical, with a period between peak populations of from three to four years. Oscillations of great amplitude are common: an acre harboring 100 rodents during the period of maximum population may have only one or two at the minimum; in some cases none at all are found. This characteristic three-to-four-year rhythm has been described in numerous species throughout temperate and arctic North America, Europe and Asia. (Corresponding patterns have not been observed, however, in tropical rodents or in Southern Hemisphere species.) The cycles are relatively regular when measured over a period of decades; what is more remarkable, the cycles of widely separated communities of rodents are often synchronous.

A number of these observations have not been satisfactorily explained. It is not known, for example, why the length of the cycle is almost invariably three to four years, or why populations apparently isolated from one another fluctuate synchronously. Only the most fundamental question about rodent population cycles has yet been investigated: Why do the populations fluctuate at all? A model of the mechanism that powers these cycles is now emerging.

Aristotle believed heavy rains caused mice to disappear; early Norwegians suggested to the contrary that lemmings fell from the sky during storms; the notion that they plunge into the sea is a later development. Later still natural historians proposed that abrupt reductions in rodent populations were caused by epidemic disease or by the activities of predatory birds and mammals. Although some of these theories were more plausible than others (and some more whimsical), none was supported by any reliable observational or experimental evidence.

The first systematic investigations of population dynamics in rodents were made by Charles S. Elton of the University of Oxford. He was one of the first to recognize the periodic nature of the fluctuations. Little direct evidence of population cycles was available when Elton began his work in the 1920's. A survey of the literature revealed that the population of the Norwegian lemming (*Lemmus lemmus*) in southern Norway had been high in 1862, 1866, 1868, 1871, 1875, 1879, 1883, 1887, 1890, 1894, 1897, 1902, 1906 and 1909. The average interval between peak populations was 3.8 years. These cycles were documented, however, only by observations of the hordes of lemmings easily seen in years of peak density.

Further evidence could be obtained indirectly. On the assumption that the size of rodent populations would be reflected in the numbers of their predators, many of which are valuable to the fur trade, Elton examined statistics compiled by the Hudson's Bay Company and other companies trading in furs from North America. He found that in northern Labrador the number of furs of the red fox, which feeds on voles, varied with a period of from three to four years; in the area surrounding Hudson Bay a similar cycle was established for the arctic fox, which preys chiefly on lemmings [*see top illustration on page 40*].

Elton's data were somewhat inexact and uncertain; fur-trade statistics, for example, could be influenced by fashion, economic considerations and other factors unrelated to the abundance of prey. The pattern of the cycles was sufficiently clear and sufficiently regular, however, for him to be able to conclude that population cycles in rodents are a widespread phenomenon, common to agricultural areas of Europe and to virtually undisturbed arctic communities. He was particularly intrigued by the fact that many independent populations in the Canadian arctic, separated by vast distances, fluctuate in phase with one another. The pulsating rhythm of these populations, he pointed out, could be expected to have important effects on the plant food supply of an area and on predators as well as on the rodents themselves.

In the late 1920's Elton and his co-workers applied more refined techniques to the study of rodent populations. By employing traps designed to catch living specimens they were able to capture the same animal repeatedly; by marking and releasing the trapped animals they were then able to follow the life cycles of identified individuals. At that time epidemic disease was considered the most likely cause of population cycles, and Elton attempted to identify the specific disease responsible for the population cycles of the British field mouse. Although he and his co-workers did discover a new disease in the animals (vole tuberculosis), he had to reject the disease hypothesis. Epidemics were encountered, but they were not always correlated with the cycles; some populations declined with no trace of disease.

The failure of the disease hypothesis in the late 1930's left ecologists without a satisfactory explanation of rodent population cycles. The two most obvious alternatives, predation and starvation, were also contradicted by the evidence. Predators did not seem to be abundant enough to eliminate the large peak populations, and rodents captured during periods of decline showed no apparent undernourishment.

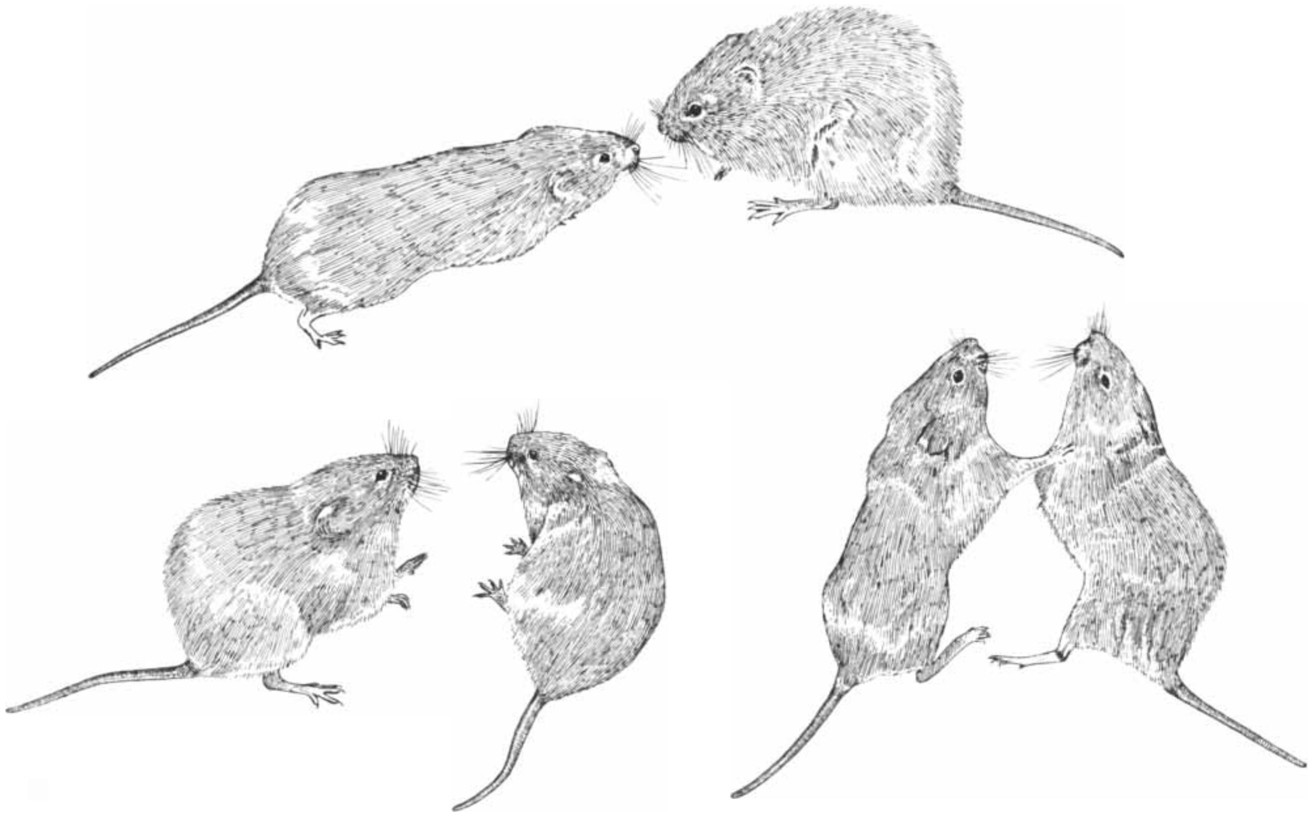
The early data compiled by Elton, such as the fur-trade statistics, consisted essentially of annual surveys. By sampling populations more frequently we have been able to describe rodent population cycles in much greater detail. Through this technique it is possible not only to identify the years of maximum and minimum population but also to determine, week by week, the rate of population growth or decline, to identify any short-term fluctuations superimposed on the cycle and to categorize the composition of the population at any time during the cycle.

All the rodent species involved in these studies are herbivores, feeding on a great variety of green plants. They are short-lived: few individuals survive for more than a year. Their reproductive rate, however, is very high. Voles and lemmings are active all year, burrowing under the snow in the winter, and many species can breed throughout the year.

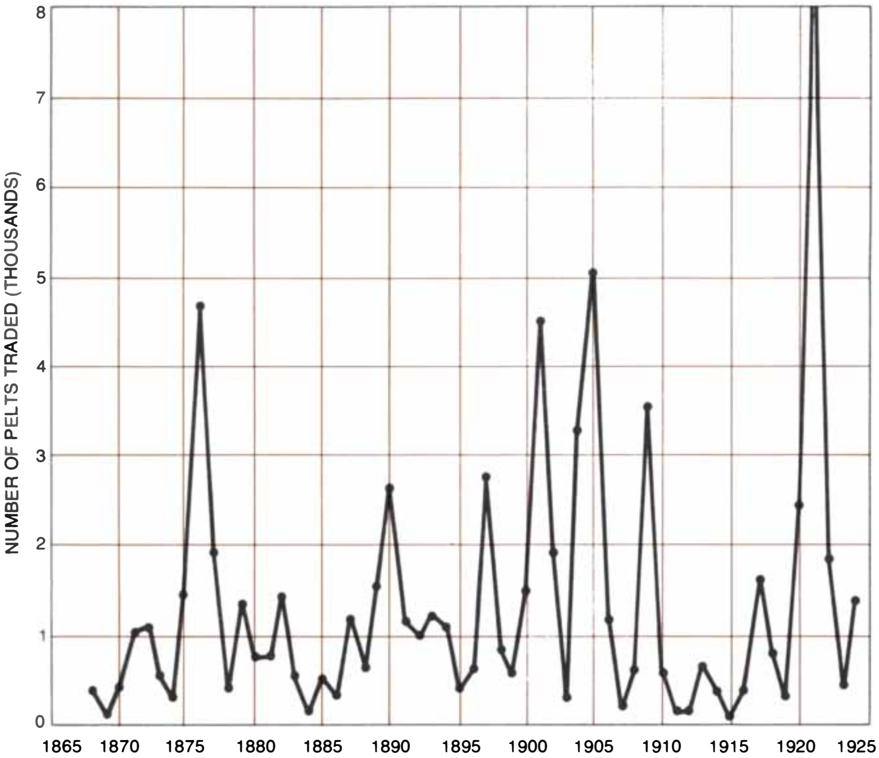
A typical population cycle is that of the meadow mouse *Microtus pennsylvanicus* of the central and eastern U.S., which we studied in fields near Indiana University [*see top illustration on page 41*]. The cycle can be assumed to start in late spring or early summer, during a period of minimum population. Density initially is about one animal per acre; extremely rapid growth brings the population to very high density by the following spring. At times the number of rodents may increase by 10 to 15 percent per week, and it is during this period of increasing population that breeding often continues throughout the winter.

The maximum density attained by the community in the early spring depends on local conditions; it may range from 50 to 300 animals per acre, but it is most commonly about 100. (By laboratory standards this peak density is quite low: 100 animals per acre is equivalent to one animal per 20-by-20-foot room.) During the spring there is often a sudden dip in density, and the population may be diminished by as much as half; this decline is often more pronounced in the number of males than in that of females. During the summer the population begins to recover, and by the fall it has usually returned to approximately the level of the spring peak.

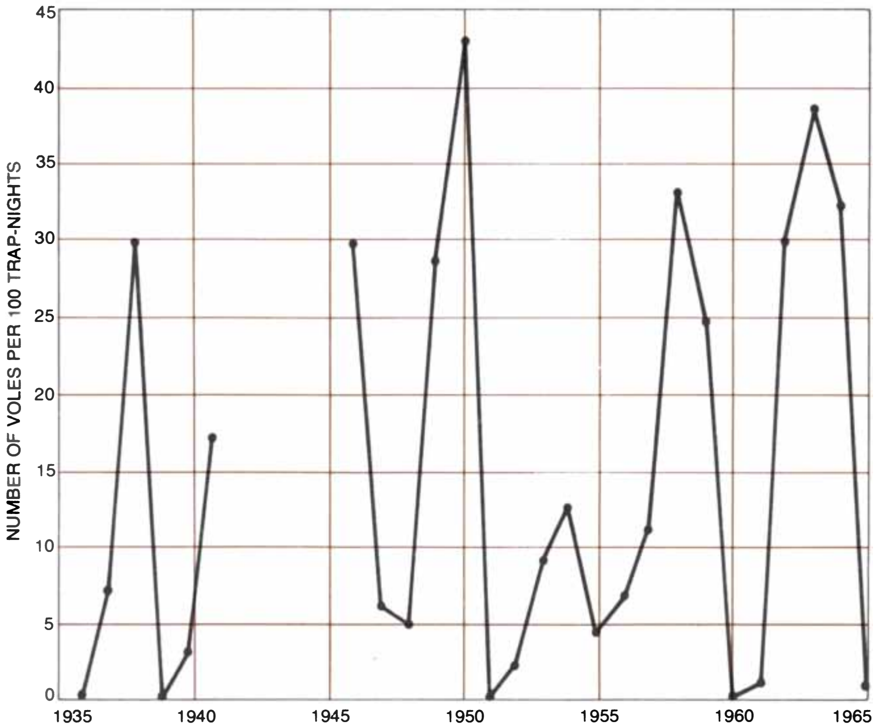
Breeding tapers off during the fall, stops entirely and does not resume until the following spring. Population decreases throughout the winter, but only slowly; normal mortality eliminates older or weaker individuals, and no young are being added to the community. In the spring following this second winter the population drops sharply, beginning at about the time breeding resumes, and may reach low density by early summer. The cycle is now complete, but what happens next is variable and impossible to predict. The population may start immediately to grow again; it may continue to decline and become so sparse that it is difficult to catch a single animal, or it may remain at a relatively low level and



AGGRESSIVENESS IN VOLES was measured in the authors' laboratory by confining two males in a small cage and observing their reaction to each other. Behavior regarded as aggressive consisted of acts labeled "approach" (*top*), "threat" (*left*) and "attack" (*right*). Each vole was tested against at least two opponents. The incidence of aggressive behavior varied during the population cycle of the rodents: an animal from a dense population was more likely to be aggressive than one from a sparse population. Fluctuations in aggressiveness are considered evidence that natural selection acts on genetically determined characteristics expressed in social interactions.



INDIRECT EVIDENCE of a population cycle in rodents was derived from the records of the Hudson's Bay Company for the fur trade in northern Labrador. The data are for the arctic fox, which feeds primarily on lemmings, and show a cycle with a period of three to four years in the number of pelts traded. The statistics were compiled in the 1920's by Charles S. Elton of the University of Oxford. Although factors other than the abundance of prey could influence the volume of the fur trade, the regularity of the fluctuations suggested to Elton a cyclical variation in rodent populations. Number of furs traded in 1921 was 9,797.



ANNUAL SURVEYS of the population of the red-gray vole in the Kola Peninsula of the northern U.S.S.R. reveal a cycle three to four years long. The method of the surveys gives not the total population but the relative density. The data were collected by T. V. Koshkina, then at the Kandalaksha Reserve in Lapland. No surveys were made from 1943 to 1946.

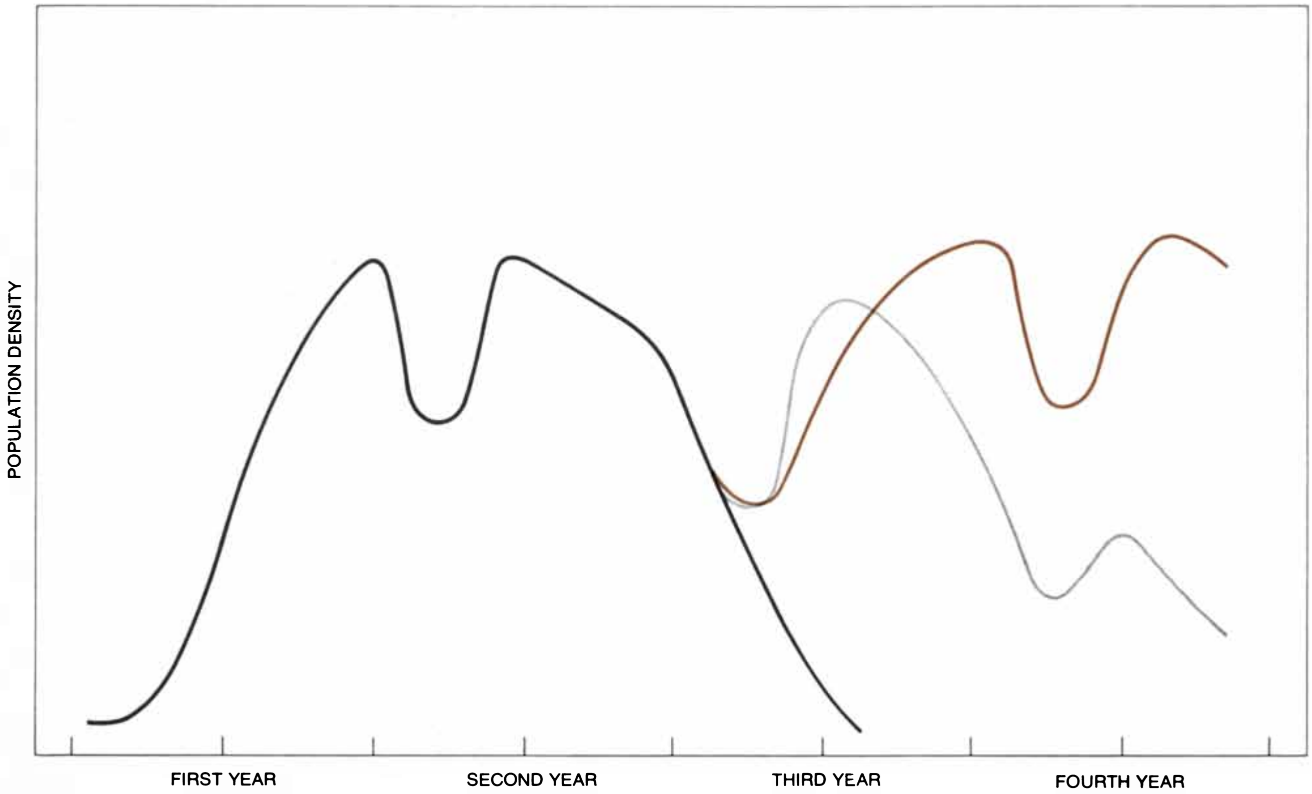
neither increase nor decrease for one or two years.

The changing composition of the population by age and sex and other characteristics during the cycle has been described in detail for several species [*see bottom illustration on opposite page*]. Survival of both adults and juveniles is good during the period of population increase, but juvenile mortality increases during the peak and decline. Adult survival remains high during the peak but drops off sharply during the decline. The reproductive rate varies steadily from a maximum during the increase to a minimum during the decline. Growth rates of individual animals also vary from a maximum during the population increase to a minimum during the decline, with the result that peak populations are composed of animals of larger than average size.

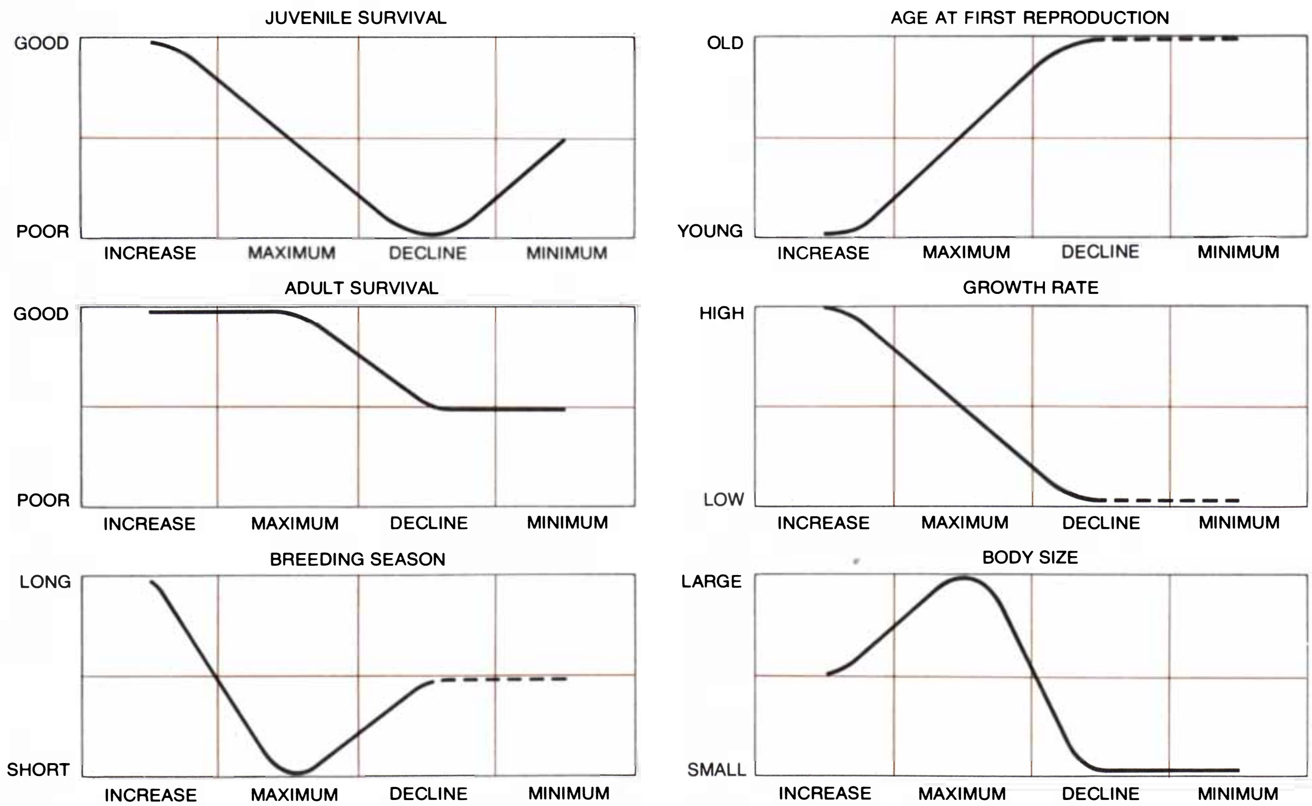
The discovery of complex patterns of changing birth, death and growth rates in population cycles has made traditional explanations of the cycles seem unlikely. Disease, predation or starvation could explain some of these observations but not all of them. In fact, no external agent accounts for the entire pattern; what is indicated instead is an internal control mechanism regulating the mode of interaction of an animal with its neighbors.

A new approach to the study of population cycles was proposed in 1950 by John J. Christian, who is now at the State University of New York at Binghamton. Hans Selye of the University of Montreal had previously shown that stress produced by a variety of physical and chemical agents could lead to disorders of the endocrine system in laboratory animals; the resulting condition is known as shock disease, and it can result in death. Christian suggested that high densities could produce stress and thus precipitate a decline in population from endocrine disturbances.

Some of the effects of stress are in fact found in field populations of small rodents, but there was nevertheless a serious flaw in Christian's argument. Dennis Chitty, working in Elton's group at Oxford, pointed out that the expected effects of high density were found not in the generation of animals exposed to crowding but in their offspring. A hereditary mechanism was obviously implicated, and Chitty speculated that the stress associated with conditions in peak populations might be genetically selective. Social interactions in a crowded community, he suggested, might favor reproduction by individuals possessing



POPULATION CYCLE of the meadow mouse *Microtus pennsylvanicus* begins in a period of extremely rapid growth, during which breeding often continues through the winter. In the second year the population drops abruptly in the spring but recovers by fall, then declines gradually as breeding ceases. In the following spring, about the time breeding resumes, the decline becomes precipitous. It may continue until the population reaches very low density (*black curve*); it may be gradual, with partial recovery in the fall, reaching minimum density the next year (*gray curve*), or it may end and the increase resume in six to eight months (*colored curve*).



SHIFTING CONSTITUTION of the population during a single cycle is reflected in measurements of adult and juvenile survival, length of breeding season, age at first reproduction, growth rate and size of individuals. Two other significant parameters—litter size and the proportion of mature adults actually breeding—are observed to change, but not in a manner determined by the cycle.

certain genetically determined characteristics and discourage the reproduction of others; the changed genetic composition of the population would in turn alter the character of social interactions within the community.

Chitty's hypothesis was initially met with skepticism. Natural selection was believed to operate much too slowly to influence a cycle with a period of only a few years. Nevertheless, the older explanations were clearly inadequate.

In 1965 we began a series of experiments intended to test some of the predictions implicit in Chitty's hypothesis. One of these predictions was that social interactions within the population would change during the cycle; in particular we decided to investigate changes in hostility between individuals, or aggressiveness.

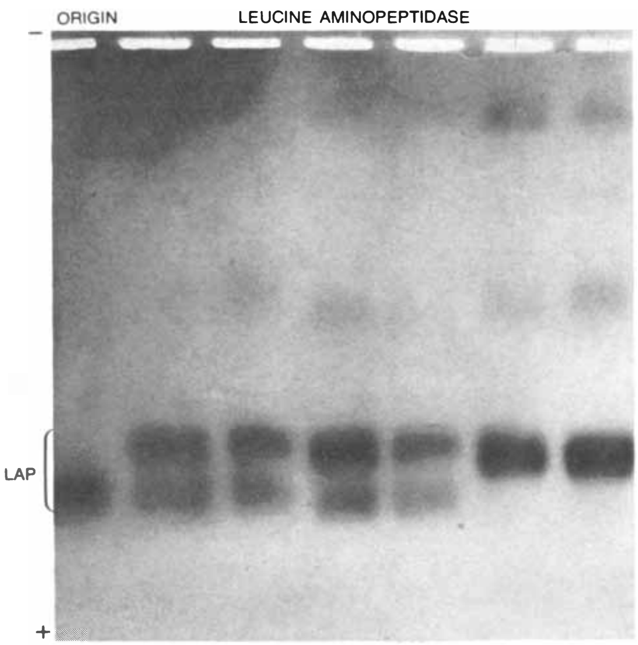
Aggressiveness was evaluated by confining two males in a small enclosure and noting their reaction to each other. The repertory of behavior defined as aggressive consisted of acts we labeled "approach," "threat" and "attack"; submission to any of these, or the failure of both animals in a pair to approach, threaten or attack, was considered evidence of lack of aggressiveness. Each male was tested with at least two opponents so that we could measure its range of reactions [*see illustration on page 39*].

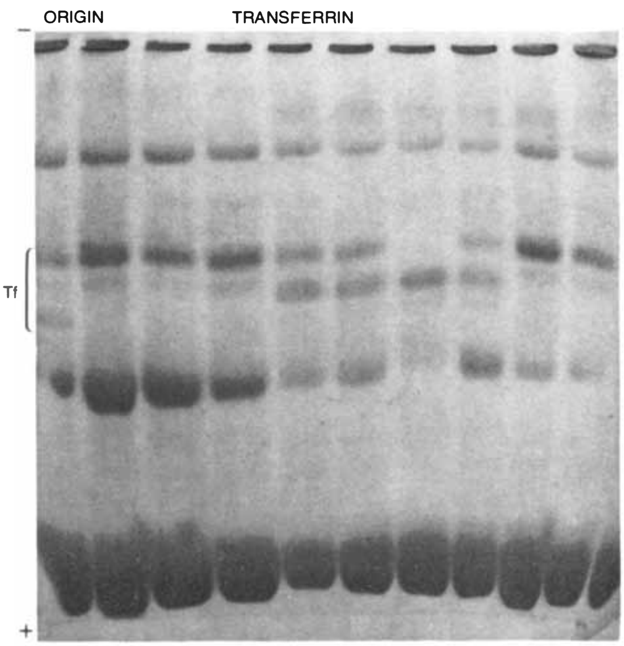
Between 1965 and 1970 we tested 1,140 males of *Microtus pennsylvanicus* and 1,450 males of *Microtus ochrogaster*. Males in peak populations were more aggressive than males from either increasing or decreasing populations; the change in behavior in response to changing population density confirmed Christian's predictions and Chitty's.

An even more crucial test of Chitty's proposal was one intended to determine whether natural selection could operate rapidly enough to account for the observed variations in social behavior. In order to make this test it was necessary to measure genetic variations in the population; where the tests of aggressiveness had been concerned with a phenotypic trait of individuals, it was now necessary to consider the genotype. When dealing with complex genetically transmitted traits, such as animal behavior patterns, making such a determination can be difficult, since the trait is governed not by a single gene (as seed color in peas is, for example) but by several or many genes that can be expressed in an individual in various combinations. It is therefore not possible merely by inspecting the animal to classify it as to genotype.

One sensitive indicator of differences in genotype is suitable for studies such as ours: by the technique of electrophoresis it is possible to detect minor differences in the chemical composition of proteins in the blood. The amino acid composition of certain proteins in some cases varies slightly from individual to individual; these variations presumably have little effect on the functioning of the protein, but they may alter the balance of electric charges on the protein molecule. Electrophoresis takes advantage of small differences in charge; the proteins, obtained from the blood serum, are placed on a gel and an electric potential is applied across the gel. Those molecules that are more strongly charged migrate more rapidly through the gel and form bands separate from those of the more slowly migrating molecules. Because proteins represent a direct expression of genetic information, the bands formed by electrophoresis can serve as reliable indicators of genotype. It is important to note that these minor variations in protein composition do not cause the behavioral changes observed in our experiments; they are merely incidental markers by which we can monitor genetic changes in the population concurrent with variations in behavioral traits [*see illustration below*].

We have worked chiefly with two protein systems: the iron-transporting protein transferrin (Tf) and the enzyme leucine aminopeptidase (LAP), both of





ELECTROPHORESIS discriminates between genotypes by detecting minor differences in the genetically specified composition of protein molecules. In these studies of vole leucine aminopeptidase (*left*) and transferrin (*right*) specimens of blood serum were placed at the origin, near one end of a gel made of boiled starch. An electric potential applied across the gel caused the proteins to migrate toward the positive terminal at rates determined by the charged groups present on the molecule. By treating the gel with stains specific to the proteins under investigation the relevant bands (as well as some extraneous bands) were revealed. Specimens showing only a fast-migrating band or a slow-migrating band are homozygous; those showing both bands are heterozygous. By observing changes in the relative frequency of the protein variants, changes in gene frequency within the population can be measured.

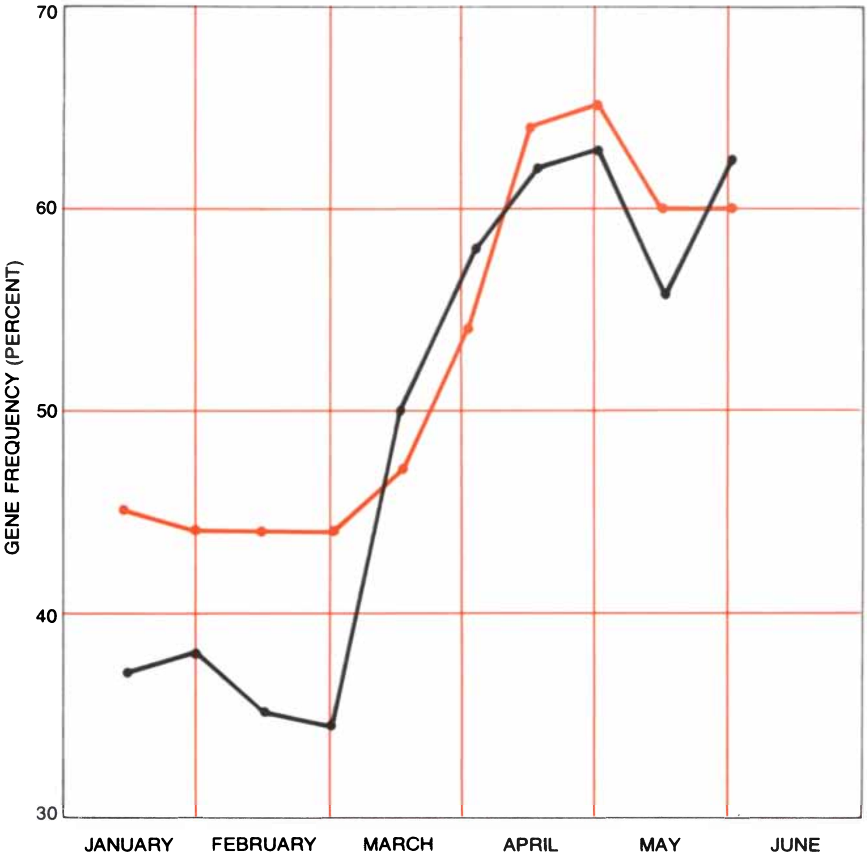
which have two main electrophoretically distinguishable variants in the rodent populations we studied. In both proteins the relative proportion of the variants in the population changed during the cycle, indicating that the frequency of the gene specifying each of the variant forms also changed. Particularly large changes in gene frequency were observed during the decline phase of the cycle, and the direction of this change was the same in several different cycles.

A change in the relative frequency of a particular genotype within a population can be brought about through a change in either the birth rate or the death rate of individuals of that genotype. In the population cycles of rodents we found that both mechanisms are important [*see top illustration at right*]. In *Microtus pennsylvanicus,* for example, study of transferrin revealed that both the birth rates and the death rates of different genotypes change systematically throughout the cycle.

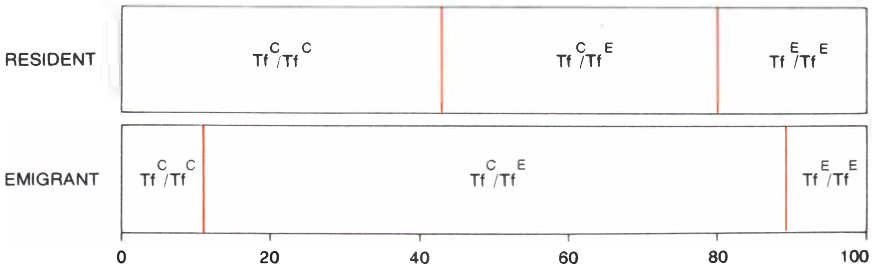
Three genotypes for transferrin are found in *Microtus pennsylvanicus:* the homozygote with genes specifying only the fast-migrating form (designated Tf^C/Tf^C), the homozygote having only the slow-migrating form (Tf^E/Tf^E) and the heterozygote, which has one of each gene and displays both bands in electrophoresis (Tf^C/Tf^E). The homozygote Tf^C/Tf^C was always more likely to survive during population increase; the heterozygote Tf^C/Tf^E always survived best during peak periods, and the other homozygote, Tf^E/Tf^E, seemed to be favored during the decline. These variations in survival were accompanied by disparities in the reproductive capability of segments of the population. Throughout the cycle heterozygous males were more often in breeding condition than males of either homozygous genotype, and females that were either heterozygotes or Tf^C/Tf^C homozygotes were more often capable of reproduction than female Tf^E/Tf^E homozygotes. Finally, the genotypes could be distinguished by the age at which the animals were first able to breed. This age was calculated indirectly by measuring the growth rates of young rodents. Tf^C/Tf^C animals grew faster than the others during the increase and peak phases of the cycle and thus became sexually mature sooner, whereas in the declining populations the Tf^E/Tf^E animals grew more quickly.

Our studies of protein variation and natural selection have led us to the following overview of the population cycle in *Microtus pennsylvanicus:* During the period of increase the survival rate is high and the selective advantage rests with those genotypes that have a high reproductive rate, a high growth rate and early sexual maturity. During the decline mortality is much higher, particularly among juveniles. Reproductive and growth rates are then less important than mere survival, and those individuals are selected that succeed in enduring, even though they produce relatively few offspring.

This model accounts for the observed correlations between genetic variability and population fluctuations, but it does not indicate which phenomenon is the cause and which the effect. Chitty's hypothesis demands that the cycle be the consequence of the changing genetic composition of the population, but from the results of these experiments alone the genetic changes could just as easily be a side effect of a population cycle governed by some external agent. To prove that the genetic mechanism is essential it was necessary to study the



CHANGES IN GENE FREQUENCY in male *Microtus pennsylvanicus* during the spring of 1969 were measured by protein electrophoresis. Relative frequencies of leucine aminopeptidase gene *F* (*black*) and transferrin gene *E* (*color*) are closely related during the period. The genes do not represent specific, known characteristics of the animals but serve merely as convenient indicators of genotype. During this period the population was declining.



RESIDENT AND EMIGRANT individuals among female *Microtus pennsylvanicus* are distinguished by genotype, as determined by electrophoresis studies of transferrin. Among dispersing voles the heterozygous genotype predominated; those that remained in the population had a broader distribution of genotypes, with Tf^C/Tf^C homozygote the commonest.

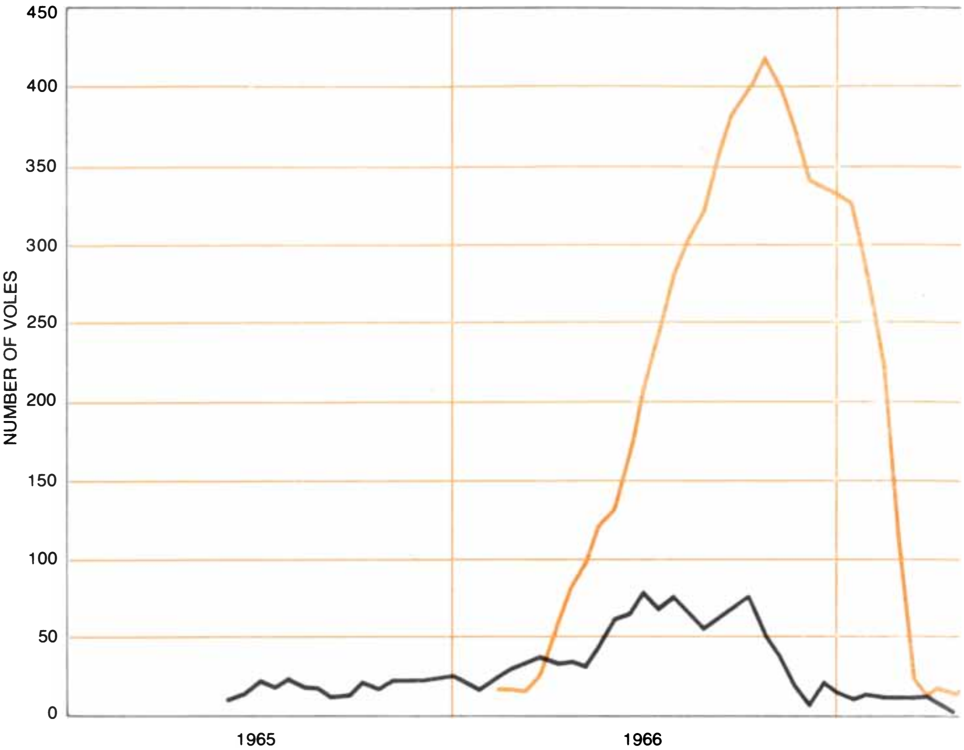
population cycle of the vole in still greater detail.

One important aid to such a study is an experimental procedure that allows one to start or stop a cycle at will. Only one such technique has been found: if a rodent population is confined to an enclosure, its population cycle is halted. This interesting fact was discovered by John Clarke of Oxford in work with *Microtus agrestis* and was independently confirmed by A. van Wijngaarden, who was then working for the Plant Protection Service at Wageningen in the Netherlands. Both Clarke and van Wijngaarden employed relatively small enclosures, about the size of a large room; we decided to investigate the effects of enclosing much larger areas, and in 1965 we established three mouseproof areas, each enclosing about two acres, in a field in southern Indiana. The areas were surrounded by a fence of wire mesh, which was sunk about two feet below the turf so that rodents could not burrow under it and which had a metal flange at the top so that they could not climb over it [*see illustration on page 46*].

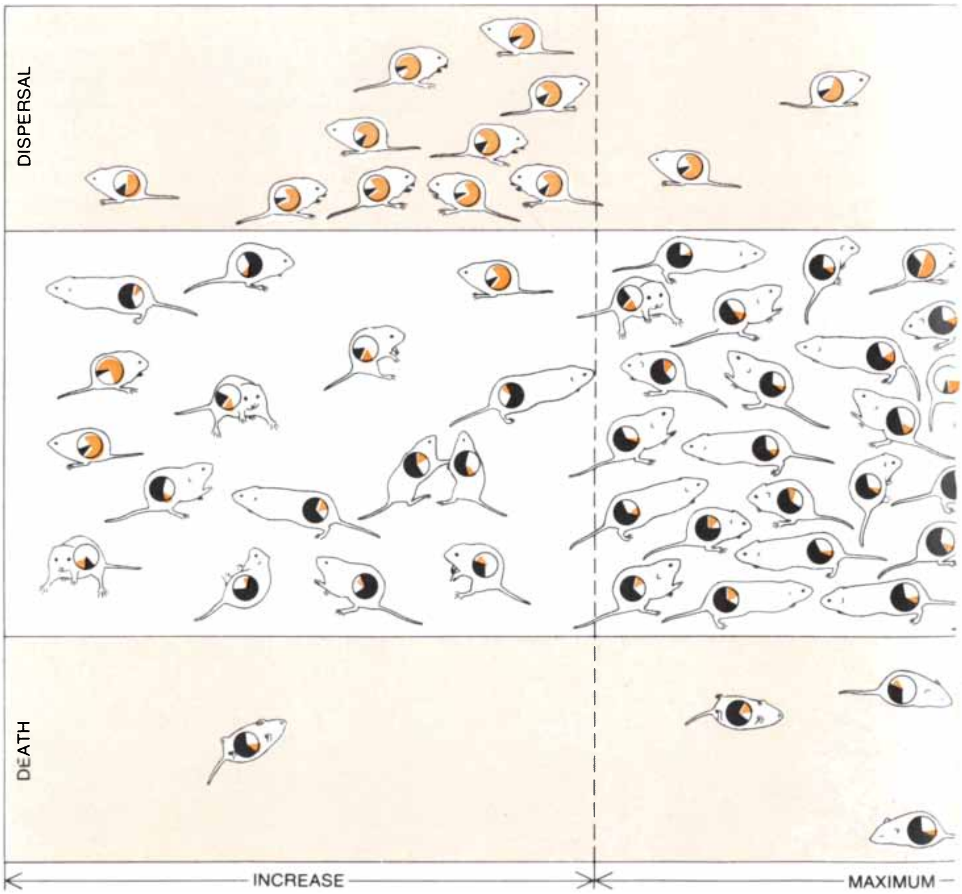
The effect of enclosure on a population of voles was dramatic. Within the fenced area populations reached densities from four to 20 times as high as those of nearby control populations. Although the enclosed animals did considerable damage to their habitat during the winter, the vegetation recovered rapidly each spring. The growth and survival rates of the animals varied in response to the changing condition of the habitat, but each spring the renewed growth of vegetation was accompanied by an immediate increase in the number of voles. Only seasonal fluctuations in population were observed; the characteristic three-to-four-year cycle was suppressed [*see top illustration on these two pages*].

These simple experiments revealed mechanisms of population regulation in voles that could not have been discovered through the observation of unrestrained populations. The results immediately suggested that the stress associated with high densities could not of itself be sufficient to cause the population to decline. Within the enclosures densities reached levels much higher than those observed under natural conditions, and the associated stress was also presumably greater than normal, yet it did not even halt population growth. The experiment also confirmed that at normal densities the deterioration of the habitat could not be a factor limiting population.

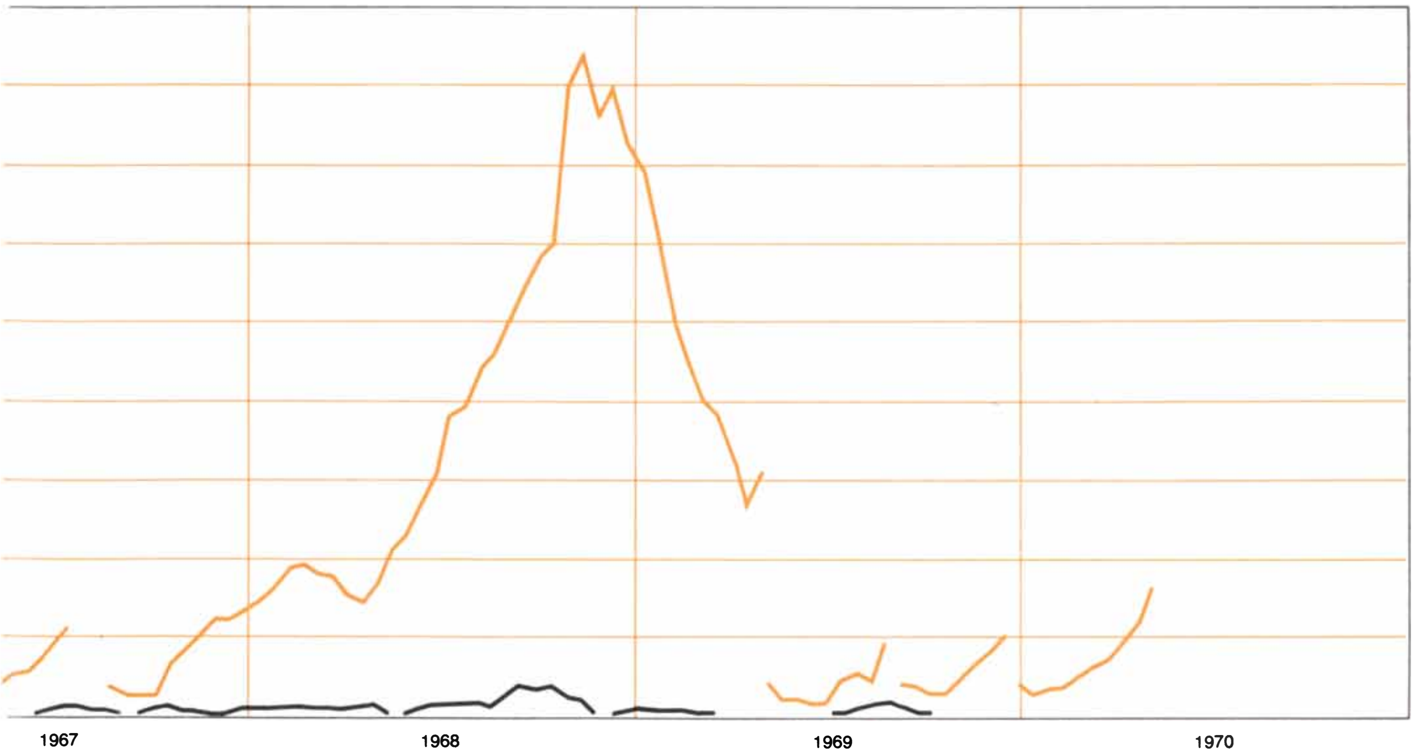
The rampant growth of the enclosed



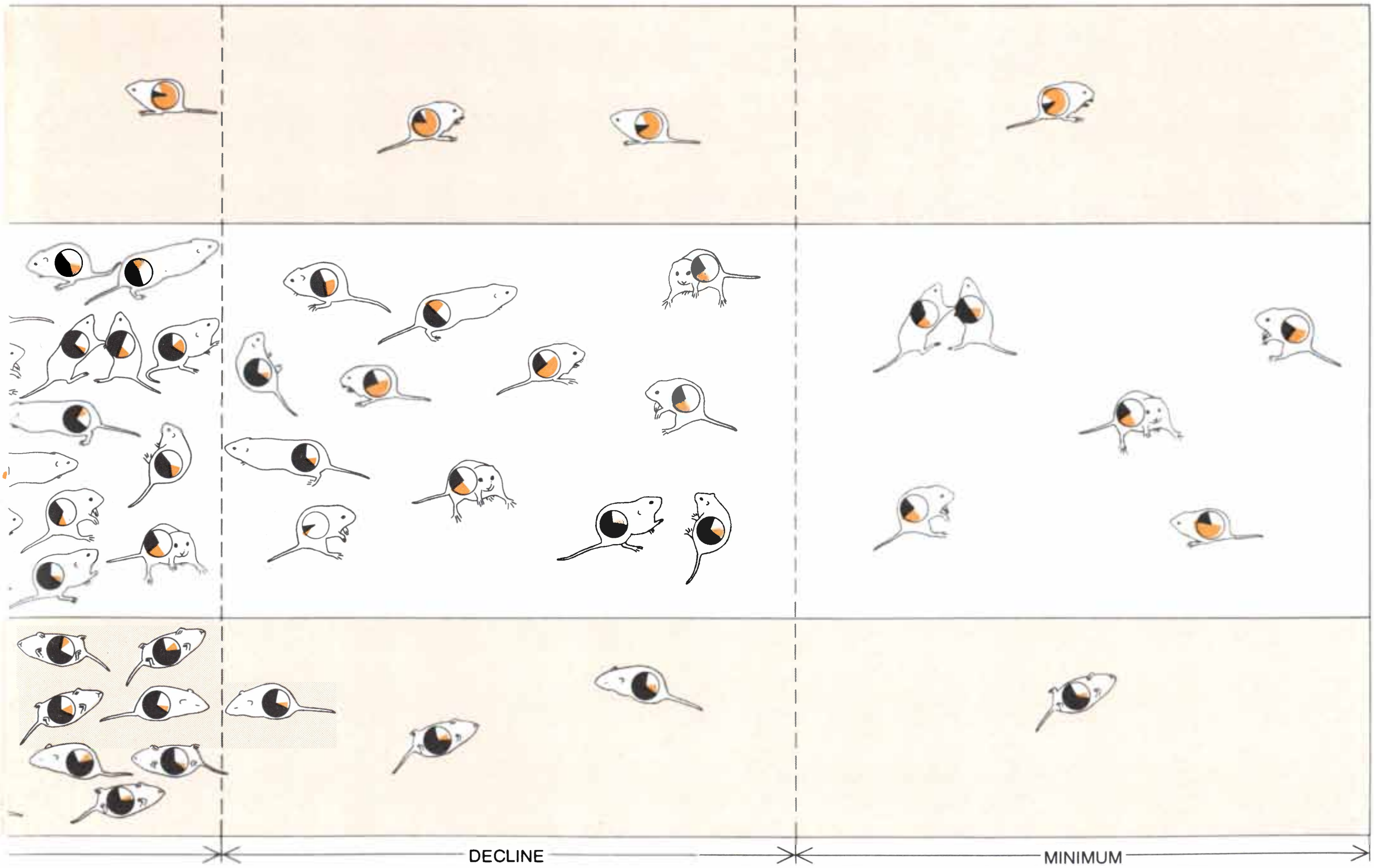
ENCLOSED POPULATIONS of *Microtus ochrogaster* failed to display the characteristic three-to-four-year population cycle; instead they expanded rapidly, reaching densities five times as high as those of nearby control populations. New populations were introduced



MODEL OF THE MECHANISM proposed to explain the population cycle is based on periodic changes observed in the genetic constitution of rodent communities. Voles with an inherited propensity for dispersal, which is associated with a high reproductive rate (*col-*



into the area five times (*colored curves*); the fact that extreme density was attained each time supports the hypothesis that dispersal is an important part of the mechanism of the cycle. Even though peak density far exceeded that of the control population (*black curve*), the enclosed communities thrived, confirming that deterioration of habitat does not limit density of vole populations.



ored areas), emigrate during the period of population increase. In the remaining population genes specifying aggressiveness (*black areas*) predominate. Aggressiveness is probably an adaptive advantage in crowded communities, but aggressive rodents are apparently poorly adapted with regard to other characteristics (*white areas*); when their mortality rate rises, the population declines.

populations suggested an obvious mechanism of population regulation in voles: dispersal. It was this process above all other processes that was disrupted by enclosure.

In order to study the characteristics of dispersing individuals we studied the animals moving into a newly available area. For two years we periodically removed all the voles from two areas of open grassland, allowed immigration into the vacant habitat for two weeks and then removed all the immigrants. The new colonists were subsequently compared with animals living in neighboring control populations.

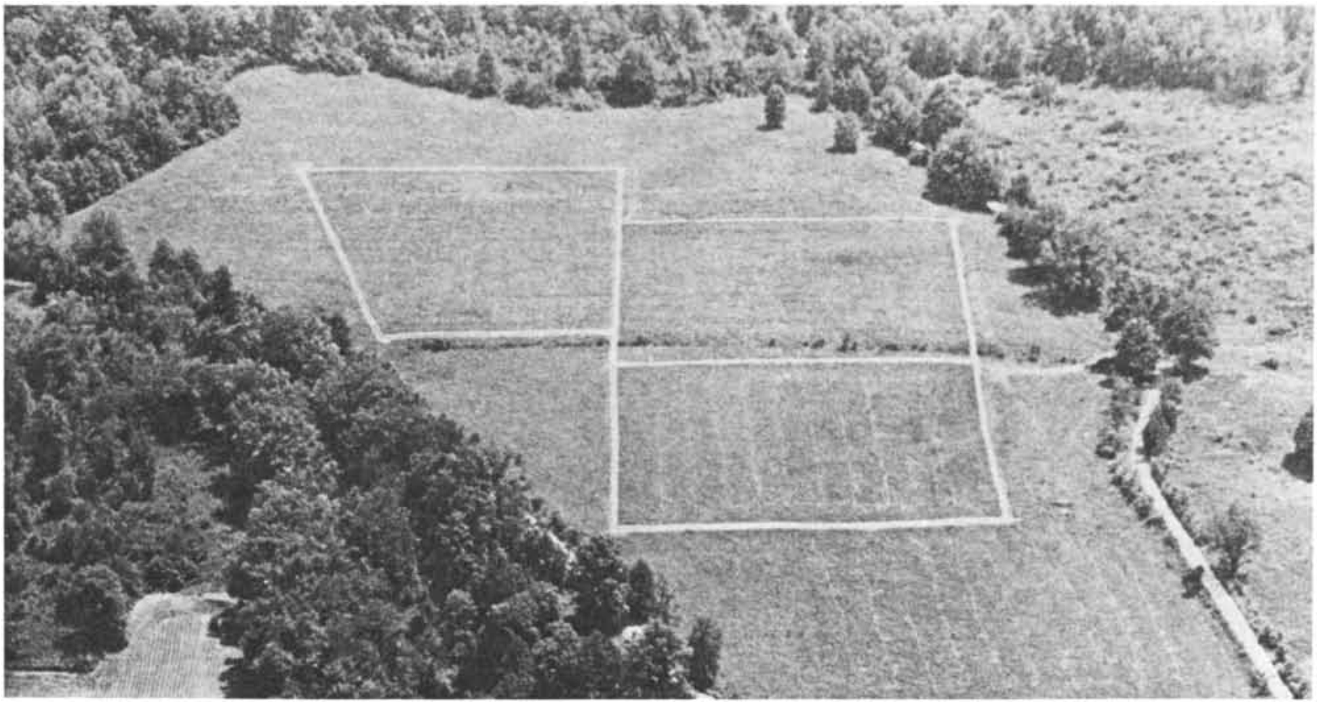
Several characteristics distinguished dispersing voles from residential ones. Young females that had recently become sexually mature were disproportionately represented among the immigrants and therefore apparently had a high tendency to disperse. As a consequence the colonists of a new territory were likely to have a high reproductive potential. During the period of population increase the genotypes of migrant female *Microtus pennsylvanicus* differed from those of residential females, suggesting that the movement of animals from crowded areas to low-density ones could significantly alter the genetic constitution of both populations [*see bottom illustration on page 43*].

The highest incidence of dispersal was observed during periods of population growth. When one of the control populations was expanding, more than half the loss of individuals from it could be accounted for by dispersal into the vacant habitat. During the decline phase, on the other hand, only about 15 percent of the loss was a result of emigration. Dispersal therefore seems to selectively remove individuals that are intolerant of crowding, and it has a significant effect on the population only when density is rising.

A model of the rodent population cycle combining Chitty's hypothesis with our findings on dispersal can now be constructed [*see bottom illustration on preceding two pages*]. Its essential postulate is that rodent populations comprise two broad genotypes, one reproductively superior but intolerant of high density, the other adapted to survival under crowded conditions but having a lower reproductive rate. As the population increases, the density-intolerant individuals leave. The probability of success of these animals is still unmeasured, but those that find a suitable uncrowded habitat probably do well. The animals that are tolerant of high density remain behind; for this segment of the population aggressiveness rather than reproductive potential is the trait of the greatest selective advantage.

The decline in the resident population can be precipitated by any of a number of factors, and the variability of this phase of the cycle suggests that the instigating event may be different in different cases and may be largely determined by chance. The ultimate cause of the decline, however, almost certainly lies in the composition of the population itself: the animals left behind, more aggressive but less prolific, are apparently more susceptible than others to ordinary causes of mortality.

Many of these conclusions are based on observations of a few species in a small geographic area; their validity and generality must now be tested elsewhere. Refugee populations may play an important role in generating population fluctuations, yet it is difficult even to prove that they exist. The attributes of dispersing animals must be studied in other species if two crucial points are to be confirmed: that dispersal is greatest from increasing populations and that young, fecund females are present in disproportionate numbers among the emigrants. The genetic and behavioral changes associated with the population cycle must also be identified in other vole species. With a broader and more thorough understanding of the mechanism of the cycle it may be possible to solve the two outstanding mysteries that remain: why the cycle has a period of from three to four years and how the cycles of isolated populations are synchronized.



RODENTPROOF ENCLOSURES near Indiana University in southern Indiana are surrounded by a wire fence in the middle of strips of land from which the vegetation has been cleared. The fence is sunk about two feet below the turf so that the animals cannot burrow under it and is capped with a metal flange so that they cannot climb over it. Each of the enclosures is about two acres.

"Kodak, I need your product. How do I get my hands on it?"

Much Kodak product finds its use in the ultimate interests of better medicine, though someone other than a physician may actually be calling for the product. Suggestions to help paramedics, administrators, biological researchers—and even physicians themselves—lay hands quicker on the proper yellow boxes:

Kodak's most specific current major service to medicine is in **radiography, cineradiography,** and **nuclear medicine.** If you are neither a radiologist nor a dentist and therefore unacquainted with the local suppliers of films, film-processing equipment, and processing chemicals used therein, and if you get no help from the Yellow Pages under "X-ray," a friendly radiologist can lead you to such suppliers.

If the need is photographic, not radiographic, and a notch beyond the requirements of most folks who shop the food markets and drug stores down the street or at the plaza, the Yellow Pages or just a store sign can guide you to a photo shop. The smaller ones may offer convenience, proximity, a more personal kind of counseling; the larger ones, scope. Some Kodak **film and paper** products may not be carried by retailers oriented toward the photo-hobbyist. A professional photographer would be pleased to direct you to a photographic stockhouse, particularly if you take the opportunity to ask about an appointment for personal portraiture. (Make it and keep it!)

It may well be that your best buy from a photo shop is literature before anything else. Some dealers are very alert to its value in building customer satisfaction with merchandise subsequently purchased. Last year, for example, there appeared the **Kodak book "Medical Photography."** Covering clinical, ultraviolet, and infrared photographic techniques, it's in stock or can be ordered for you. For the free *1974 Index to Kodak Information,* drop a card to Dept. 412-L, Kodak, Rochester, N.Y. 14650 and ask for "L-5."

A favorite Kodak publication with biomedical people who teach or present papers and talks is *Slides With a Purpose.* A note to A. T. Brown, Kodak, Dept. 640W, Rochester, N.Y. 14650 can get you a free copy and the name of a nearby organization to keep you abreast of the burgeoning new form of inexpensive **visual and aural presentation** by super 8 film.

In microfiche, Kodak offers the biomedical community a modern alternative to costly traditional forms of publication for specialized material requiring good illustration. For this and such other applications of **microfilm** as patient records, find us in the Yellow Pages under "Microfilming," or write to C. S. Bishara, Kodak, Rochester, N.Y. 14650.

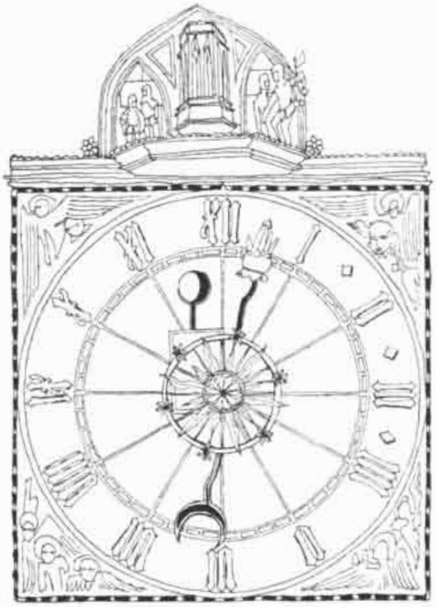
Light-sensitive recording material for instruments? Available from outfits mentioned that handle Kodak x-ray films. Or write P. J. Coder, Kodak, Dept. 659, Rochester, N.Y. 14650.

Supplies for **liquid scintillation counting, thin-layer chromatography, electrophoresis, protein sequenation and synthesis? Stains? Analytical reagents? Building blocks for synthesis?** Lists and product information from G. S. Grau, Kodak, Rochester, N.Y. 14650. You can order the Eastman organic chemicals themselves from leading lab supply houses (not photo dealers).

Materials for **electron micrography** and **photomicrography?** For **autoradiography?** For **special photographic recording** techniques? **Light filters** and **attenuators?** Inquire of W. A. Sykes, Kodak, Rochester, N.Y. 14650.

Workers in biology and medicine who **still** think we are playing hard-to-get should complain to Richard Fuller, Kodak, Rochester, N.Y. 14650. Advice given here applies only in the U.S.A. For other countries, Fuller can forward inquiries.

Kodak



Fuel and Food

Most discussions of increasing the food supply of poor countries assume that the key ingredients are high-yielding grains, generous use of fertilizer and adequate irrigation. It is assumed, in short, that the U.S. food system provides a model that can be copied almost anywhere. When the U.S. food-supply system is closely examined, however, it turns out to be highly dependent on the input of energy, and hence less exportable than one might have thought.

The amount of energy contained in the food that reaches American tables adds up to some 250×10^{12} food calories, or 10^{15} British thermal units, per year, which is equivalent to about 1.4 percent of the total annual U.S. energy consumption. For every calorie that reaches the table, however, another nine calories must be expended to put it there in the form of fertilizer, farm machinery, irrigation, fuel for farm vehicles, energy for food processing and packaging, fuel for delivery trucks and finally energy for refrigeration and cooking. The ratio of nine calories of energy expended per calorie of food consumed means that about 13 percent of the total U.S. energy production is devoted to keeping Americans fed. The estimate, recently published in *Science*, was made by John S. Steinhart of the University of Wisconsin and Carol E. Steinhart. They observe: "To feed the people of India at the U.S. level of about 3,000 food calories per day (instead of their present 2,000) would require more energy than India now uses for all purposes."

The Steinharts note that their estimate omits a number of important activities that could well raise the energy investment per food calorie from nine calories to as many as 12. Not included, for example, is a share of the cost of the highway network traveled by food delivery trucks. (Nearly half of all truck traffic is devoted to moving food and agricultural items.) Also omitted is the energy in fuel consumed by private automobiles on food-shopping trips, which may exceed the energy in the food by half.

Between 1920 and 1970 the energy input to the U.S. food system increased roughly tenfold from about 200×10^{12} kilocalories per year to $2,170 \times 10^{12}$ kilocalories. (A food "calorie" is actually a kilocalorie.) In the same period the index of farm output increased by a factor of only two. When output is plotted against energy input, the result is an S-shaped curve, similar to the typical growth curve in which a brief period of exponential growth is followed by a marked slowdown. This suggests to the Steinharts that "we are near the end of an era.... It is likely that further increases in food production from increasing energy inputs will be harder and harder to come by."

The Steinharts suggest that the energy required for agriculture might be reduced in a number of ways. One would be to use natural manure from animal feed lots to replace at least some of the commercial fertilizer. Greater reliance on crop rotation by interplanting legumes would provide additional "green manure." Such measures have only limited potential, however, because three-fourths of the energy investment in the U.S. food-supply system is made off the farms in processing and distribution and in food-preparation in the home.

More about Repressors

Most genes, like machines in an efficient factory, are turned on only when there is a demand for their products. The rest of the time they are turned off, and one of the devices for keeping them turned off is a repressor: a protein molecule that blocks the transcription of a gene or a set of related genes into RNA for subsequent translation into protein. It does so by binding to a site (the operator) on the DNA and preventing the enzyme RNA polymerase from transcribing the gene. When there is a need for the product of a repressed gene, the shape of the repressor is modified by an "inducer" so that the repressor lets go of the operator and RNA polymerase gets on with its job. The concept of repression was put forward on the basis of genetic data by François Jacob and Jacques Monod of the Pasteur Institute in 1961. In 1966 the first molecular evidence was developed: Walter Gilbert and Benno Müller-Hill of Harvard University isolated the *lac* repressor, which controls the synthesis of three enzymes that metabolize lactose in the bacterium *Escherichia coli.* The next month Mark Ptashne of Harvard isolated the repressor that keeps the genes of a bacterial virus, phage lambda, turned off so that the phage lies dormant within an infected *E. coli* cell instead of multiplying. Soon the binding of specific repressor to its operator had been demonstrated for lambda and then *lac* [see "Genetic Repressors," by Mark Ptashne and Walter Gilbert; SCIENTIFIC AMERICAN, June, 1970]. Since then the *lac* and lambda repressors and operators have been under intensive study.

Gilbert isolated the short region of DNA that constitutes the *lac* operator by letting pieces of DNA that contained the region interact with the repressor, which bound to the operator. With an enzyme he digested away the nonoperator DNA, which was not protected by the repressor, and then he used an inducer to make the repressor let go. He was left with a bit of DNA: the operator. Gilbert and Allan Maxam worked out the operator's sequence of bases, the subunits whose sequence along the inside of a DNA double helix encodes genetic information. Many of the 21 bases turned out to be arranged with twofold symmetry, they reported in *Proceedings of the National Academy of Sciences,* that is, the sequence of complementary base pairs is read in reverse order as one moves from the center to the ends of the operator. This seems to conform with what has been learned about the *lac* repressor. It is a tetramer: a molecule composed of four identical subunits. Konrad Beyreuther and his colleagues in Müller-Hill's laboratory at the University of Cologne have published the complete sequence of its 347 amino acids in *Proceedings,* and more recently Thomas A. Steitz, Timothy J. Richmond, David Wise and Donald Engelman of Yale University reported on electron-microscope studies indicating that the four subunits fit together to form two grooves, in either of which the DNA double helix may lie "like a hotdog in a hotdog bun." Each half of the groove may be read by half of the symmetrical operator sequence. Finally, Nancy Maizels of Gilbert's laboratory transcribed *lac* DNA into RNA and found that transcription begins in a region that binds repressor, showing that the repressor works by blocking the binding of polymerase directly.

Meanwhile Ptashne and his colleagues have been publishing new information about repressor control in lambda. They had found there were two operators in lambda, one on each side of the gene that codes for the repressor protein; each of them controls transcription of a different gene for one of the "early" products the virus synthesizes. Ptashne and Tom Maniatis found that each of the operators actually consists of six binding sites, each of which has a somewhat different sequence of bases and therefore a different affinity for repressor. As increasing amounts of repressor are added to a laboratory preparation of lambda DNA, a double molecule (dimer) of repressor first binds to one site, which is about 35 base pairs long. Then single repressor molecules bind to successive, shorter sites. The multiple binding may provide a kind of modulation of repressor activity. The investigators have also found that the sites to which RNA polymerase binds, the promoters, are right in the operators: lambda repressor thus blocks the binding of RNA polymerase directly. Maniatis and Ptashne, with B. G. Barrell and John Donelson of the Medical Research Council Laboratory of Molecular Biology in Cambridge, are now publishing in *Nature* the base sequence of a region including a single repressor-binding site. It turns out to have three axes of interdigitating two-fold symmetry. They suggest that each of the overlapping symmetries may be recognized by a different protein—not only by the repressor itself and the RNA polymerase but also by the products of one or more of the very genes whose transcription is blocked by the lambda repressor, since these products seem also to be regulatory proteins.

Consortium

American libraries have been hit hard in recent years by the rising volume of publications, the rising cost of operation and the falling level of Federal assistance. Four of the major research libraries in the U.S. (the research libraries of the New York Public Library, Harvard University, Yale University and Columbia University) have moved to meet these problems by establishing as a separate entity the Research Libraries Group, which will work to increase cooperation among the four members of the group and any other research libraries in North America that may choose to join it. The first move is the establishment at Yale of a bibliographic information center that will collect and distribute information on the holdings of the four libraries.

Later the group intends to institute a program of cooperative purchasing, aimed at making it unnecessary for each library to buy an expensive work that is unlikely to be much in demand. Instead one library will buy it, and steps will be taken to facilitate the exchange of material among the libraries by photocopying and the electronic transmission of published material. These prospects have caused the Association of American Publishers, which represents book publishing companies, to protest to the U.S. Senate's subcommittee on copyright that the plan "ignores the rights of copyright owners."

Homing Cells

It has long been apparent that the major neural pathways and circuits of the brain are genetically determined. In the developing embryo, nerve cells in one part of the brain grow fibers that seek out specific targets in other parts of the brain. A major question is: How does a fiber "recognize" its target?

The leading hypothesis is that each nerve fiber is endowed with a specific chemical component and that the target has a complementary component. When the growing tip of a fiber encounters a cell surface for which it has a chemical affinity, it adheres to that surface. Of the experiments supporting this hypothesis perhaps the best-known are those conducted by R. W. Sperry of the California Institute of Technology [see "The Eye and the Brain," by R. W. Sperry; SCIENTIFIC AMERICAN, May, 1956].

Further support is provided by an elegant demonstration of adhesive selectivity in nerve cells recently achieved by Stephen Roth, Anthony J. Barbera and Richard B. Marchase of Johns Hopkins University. Earlier investigators had established that fibers from the dorsal retina of the chick form spatially specific connections with cells on the surface of the ventral tectum in the midbrain, and that fibers from the ventral retina similarly innervate the dorsal tectum. Roth and his colleagues dissected nerve cells from the dorsal and ventral regions of the retina in chick embryos and labeled them with the radioactive isotope phosphorus 32. Then the labeled cells were incubated in various combinations with tissue from the dorsal and ventral regions of the tectum. The experimenters found that the cells from the dorsal region of the retina adhered to the tissue from the ventral region of the tectum, and vice versa. "A simple explanation of these data," write Roth, Barbera and Marchase in *Proceedings of the National Academy of Sciences,* "is that moieties on the cell surface participate in a process of recognition and specific adhesion, and that these molecules play a major role in an early selectivity that ultimately determines the retinotectal projection."

Breathless

If one takes a deep breath and holds it, a feeling of distress radiates from the chest to the head and the limbs; after about a minute the sense of unease becomes urgent; eventually the compulsion to resume breathing must be obeyed. The warning provided by these sensations is of obvious value to the individual, but the physiological mechanism that gives rise to them is not well understood. A recent investigation suggests that the response to breath-holding may be controlled in part by the carotid bodies, two small structures in the neck that adjoin the carotid arteries and are richly supplied with nerve endings.

It has been known for some time that the carotid bodies are involved in the regulation of respiration, and resection, or partial removal, of the bodies has been employed in the treatment of certain respiratory conditions, such as asthma and emphysema. Five patients who had undergone this surgical procedure and six people with normal carotid bodies were studied by Joseph T. Davidson and his co-workers. Writing in *The New England Journal of Medicine,* they describe large differences between the two groups in the time required to reach the "breaking point" of breath-holding. The report is written by Davidson, of the Harbor General Hospital in Torrance, Calif., and Brian J. Whipp, Karlman Wasserman, Sankar N. Koyal and Robert Lugliani.

Before holding their breath the subjects inhaled mixtures of gases with oxygen concentrations of 100 percent, 50 percent, 21 percent (approximately the oxygen content of air) and 12 percent. When breathing pure oxygen, all the subjects were able to hold their breath for about the same length of time: an average of slightly more than two minutes. As the oxygen content was decreased, however, the results for the two groups diverged. For the subjects whose carotid bodies had been resected the breath-holding time declined only slightly; at an oxygen level of 12 percent it was 115 seconds. For the control group the length of time to the breaking point declined steeply; at the level of 12 percent it was only 55 seconds.

This large difference in the duration of breath-holding, Davidson and his coworkers conclude, suggests that the sensation of breathlessness arises at least in part from the carotid bodies. A clue to how they might generate this sensation is provided by analysis of the exhaled gases: the subjects whose carotid bodies had been resected were able to tolerate a lower partial pressure of oxygen and a higher partial pressure of carbon dioxide than the control subjects. Davidson infers that the carotid bodies may be chemoreceptors capable of monitoring the concentration in the blood of both gases, and that they may be influenced by other stimuli as well.

The Day the Dam Broke

"Channeled Scablands" are what geologists call some 15,000 square miles of eastern Washington that include Grand Coulee, a canyon 50 miles long and 900 feet deep. The name was coined in the 1920's by J. Harlen Bretz of the University of Chicago to describe the heavily eroded area, which is roughly bounded by the Snake, Spokane and Columbia rivers. Bretz suggested that the soil in the scablands had been stripped down to bedrock thousands of years ago by a tremendous flood. Since Bretz's day detailed field studies have confirmed his hypothesis, and the U.S. Geological Survey has now published a reconstruction of the 18,000-year-old catastrophe.

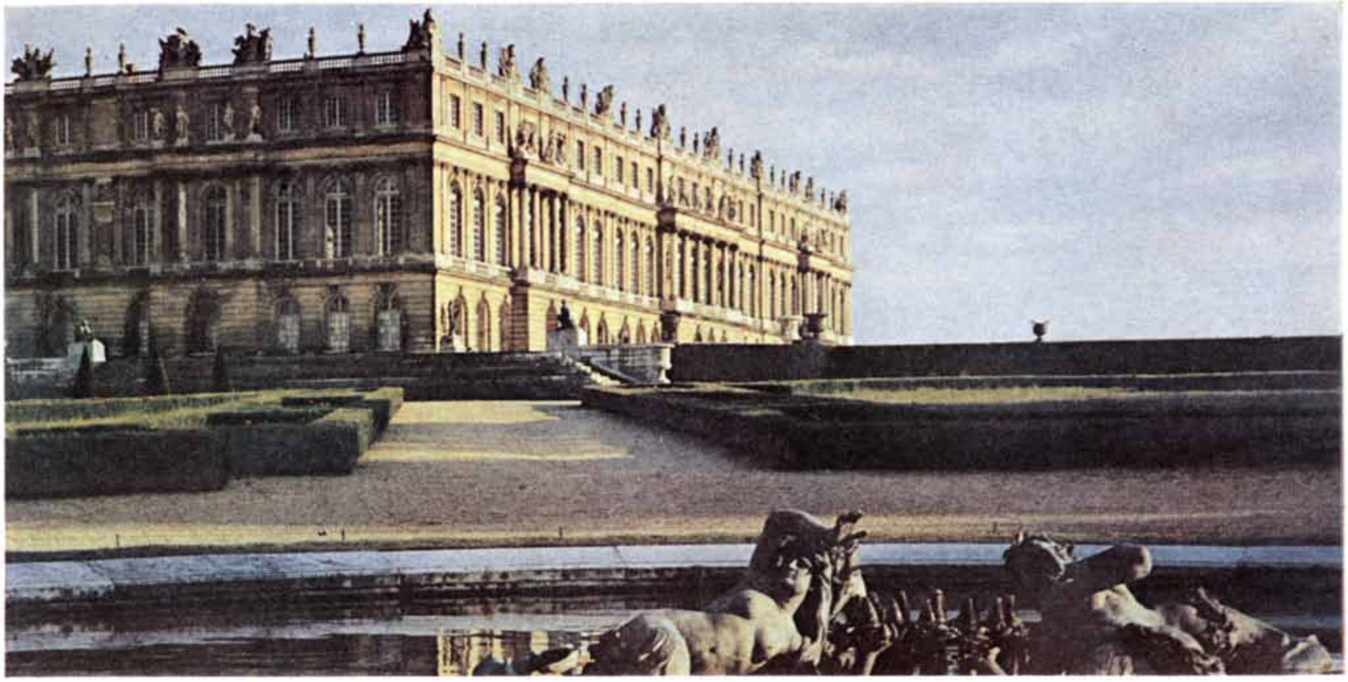
The Pleistocene ice sheet that then covered northern Washington, Idaho and Montana extended a tongue from time to time in the vicinity of today's Pend Oreille Lake in Idaho; each extension formed a natural ice dam that created a temporary glacial lake stretching eastward toward the Continental Divide. The lake that rose behind the dam 18,000 years ago eventually reached deep into Montana, until it contained about half as much water as Lake Michigan does today. When the lake level at last rose above the top of the dam, the ice gave way, releasing an accumulated 500 cubic miles of water in a flash flood that eventually reached the Pacific 350 miles to the west.

The rate of flow as the lake drained away is estimated to have been slightly less than 400 million cubic feet per second, or about 10 times the combined flow of all the rivers of the world. The water had only one way to go: down the Spokane valley to the north rim of an ancient tilted lava field that was then deeply buried under thousands of years' accumulation of loess. The slope of the field, 25 to 35 feet per mile, helped the floodwaters to maintain a velocity high enough to sweep away the loess and pluck great blocks out of the basalt bedrock. At about the point where the city of Spokane now stands the flood split into three streams. The easternmost, in some places 20 miles wide and 600 feet deep, excavated a channel to the southwest, known today as the Cheney-Palouse Tract. The middle stream carved out what is now Crab Creek Channel. The westernmost stream, which may have carried the greatest volume of water, formed Grand Coulee.

The three streams rejoined in the Pasco Basin to the south, near the Oregon border, and spilled through Wallula Gap west of modern Walla Walla into the Columbia River Gorge. The floodwaters followed the channel of the Columbia through the Cascade Range and finally built a delta on the lower Columbia. Within 30 days, according to Geological Survey estimates, the streams of the channeled scablands had returned to their normal rate of flow. The terrain of the scablands, however, had been permanently changed.

The Laughing Conformist

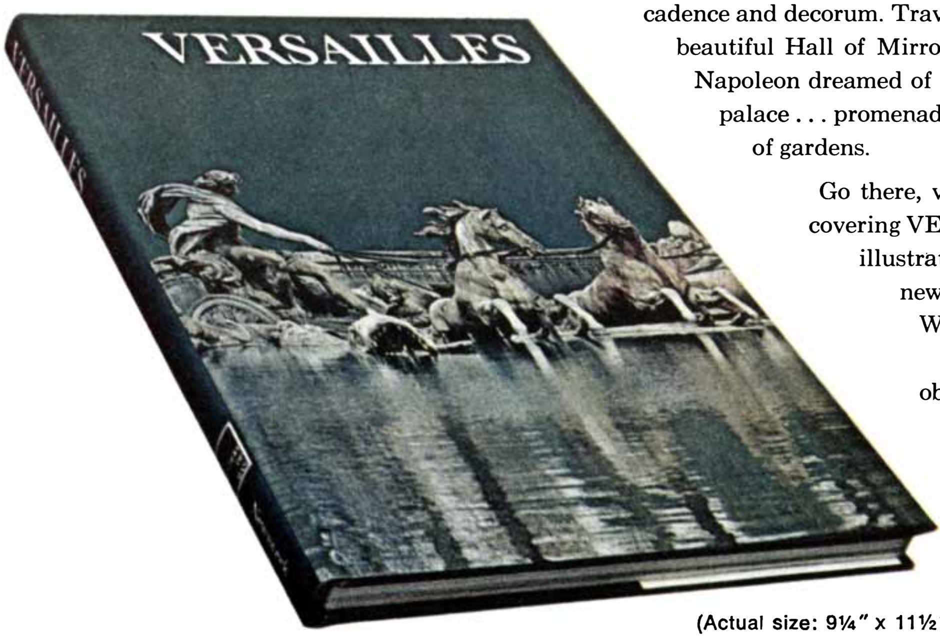
Social conformity is not easy to measure, but Solomon E. Asch managed to do it in a classic experiment of some 20 years ago. What Asch did was to assemble a group of confederates, tell them that an experimental subject would be joining the group shortly and arrange to have them give deliberately false responses to questions so that the effect on the unwitting subject's responses could be observed. For example, Asch would ask if a line on a card was longer or shorter than a line on another card, and his confederates would give a unanimously wrong answer. Seventy-five percent of the subjects went along in varying degrees with the synthetic majority, against the evidence of their own senses [see "Opinions and Social



a magnificent obsession

It had been his father's hunting lodge isolated in a marsh far from Paris . . . but for Louis XIV it soon became a magnificent obsession called Versailles —a grand palace for his elegant whims . . . where leopards and lions would roam in courtyards below as hundreds of guests dined at silver tables in week-long banquets and celebration . . . where he, the Sun King of France, reigned supremely from his world of mirrors and chandeliers, from his acres of gardens stocked with hummingbirds and flamingoes to the affairs of state and the affairs of passion.

Versailles—the life-long creation of Louis XIV, who at the age of 30 was Europe's most formidable ruler—is the story of France in her most opulent era . . . of court intrigue where rival mistresses used poison as well as passion to gain favor . . . of the famous and powerful Madame de Pompadour and the haughty, lovely Marie Antoinette . . . of revolution and artistry, decadence and decorum. Travel to Versailles and walk the coldly beautiful Hall of Mirrors . . . stand at the window where Napoleon dreamed of future glories in an echoing, empty palace . . . promenade through a vast and incredible city of gardens.



Go there, visit Versailles as our guest by discovering VERSAILLES—a unique and lavishly illustrated volume that's part of an exciting, new series from Newsweek Books called Wonders of Man. You can sample this volume—and the series—without obligating yourself to buy anything. The order card (or the coupon) will bring you VERSAILLES to examine in your own home for 10 days.

(Actual size: 9¼" x 11½")

Picture credits: Adam Woolfitt; Versailles (Musées Nationaux) and Louvre.

A series called Wonders of Man

VERSAILLES is just one of the volumes in Newsweek's Wonders of Man series. Each book in the series uses a particular monument as a focal point and examines the important aspects of a nation's history, culture, government and entire civilization. Each volume is 9¼" x 11½" and has 176 pages. Each lucid text is combined with art and literature in an attractive, interesting-to-read package that includes:

- a 40,000-word narrative history—lively and lucid—written by a world-renowned historian and richly illustrated in color and black and white. There are about 100 illustrations in each volume—more than half in full color.
- a section on literature—with excerpts from writings about the monument. Also included in this section is a chronology of the monument and the civilizations with which it is associated plus a guide to other monuments in the city or country, a supplementary reading list, and an index.

From VERSAILLES to THE COLOSSEUM, from the STATUE OF LIBERTY to THE PYRAMIDS AND SPHINX, from NOTRE DAME De PARIS to THE KREMLIN—you'll enjoy a chronological sweep of history from ancient Egypt and the Roman Empire to 19th-century America.

FREE GIFT! **A beautiful addition to the Wonders of Man tour through history.**

Musical England

A special selection of melodies to delight the ear

A musical tour that gives you the feeling of Merrie Olde England. A collection of aires, ballads, folk tunes and classical music chosen to please the ear. This special 12" stereo record contains selections performed by such artists as organist E. Power Biggs and guitarist John Williams, and melodies sung by the Norman Luboff Choir and Joel Grey.

This one-of-a-kind record is yours—FREE, no cost, no obligation—for agreeing to simply examine VERSAILLES!

Please note: Additional copies of MUSICAL ENGLAND are available, on request, at $5.95 per copy, plus a small charge for postage and handling.

Yours FREE for 10 Days

Don't miss this no-risk opportunity to begin your collection of these distinctive volumes in such a convenient and inexpensive way. Our plan, you'll find, is indeed simple:

1. The reply form will bring your copy of VERSAILLES (176 pages; 9¼" x 11½" overall) to examine at your leisure for 10 days—without any cost or obligation.
2. If you're not thoroughly delighted with VERSAILLES, return it in 10 days—and owe nothing.
3. On the other hand, should you decide to keep VERSAILLES, just send $6.95 (plus the small postage and handling charge). You will then receive future volumes of Wonders of Man at regular two-month intervals—every one as lavish and dynamic as the first. Every volume with about 100 illustrations—half of them in color.
4. Each volume will be offered on the same 10-day free-examination basis—at the same low price of only $6.95, plus postage and handling.
5. At no time are you under any obligation to buy. You accept as few or as many books as you wish—there's no minimum. Build your library at your own pace, according to your own preferences. And you're free to cancel whenever you decide.

Deluxe Edition Available

You may prefer VERSAILLES and future volumes in the Deluxe Edition. The fine-grain binding is designed to outlast and outshine ordinary leather. Handsome gold-tooling on the spine and a full-color laminated dust jacket add extra luxury. These volumes have the elegance of fine books costing a good deal more. Yet your cost is only $2 additional per book. Just check the appropriate box on the reply form.

Send for your copy of Versailles today!

No need to send money now. No obligation ever!

To: Newsweek Books/ WONDERS OF MAN CWX34
The Newsweek Building, Box 408, Livingston, N.J. 07039

Please send me **VERSAILLES** to examine **FREE for 10 Days** and enter my trial subscription to Wonders of Man. If, after examining this book, I do not choose to keep it, I may return it within 10 days and I will be under no obligation. If I keep **VERSAILLES**, I will pay only $6.95 plus shipping and handling. You may then send me future volumes in the series, shipped a volume at a time at two month intervals at the same price of $6.95 plus shipping and handling.
Each book is to be sent to me on approval. I may return any book in this series within 10 days after receiving it, and I can cancel my subscription at any time simply by notifying you.

My initials:............................

☐ I would prefer having **VERSAILLES**—and any future volumes I may take—in the Deluxe Edition at only $2 more per book.

Include my FREE GIFT (mine to keep whatever I decide about the book!)

Mr.
Mrs. ____________________
Miss (please print)

Address ____________________

City State Zip

NOTE: Available only in U.S. and Canada.
Residents of Canada: Add $1.00 more per volume.

Pressure," by Solomon E. Asch; SCIENTIFIC AMERICAN, November, 1955].

Asch's experiment has been repeated in many forms, and the results have done much to demonstrate the pervasiveness of social conformity and to illuminate how it is achieved. The technique has always had the drawback, however, of being painful to the subject who is dismayed to discover how easily he succumbed to group pressure. Now T. A. Nosanchuk and Jack Lightstone of Carleton University report a new technique that appears to overcome this drawback. In their experiment the social pressure is exerted not by deliberately false responses but by canned laughter.

Nosanchuk and Lightstone, who describe their work in *Journal of Personality and Social Psychology,* tell their subjects they are participating in a study of humor and present them with recorded jokes, some funny, some extremely funny and some unfunny. The subject, who sits in a booth with a set of headphones and a microphone, is under the impression that four other subjects are listening to the jokes at the same time and that he is hearing their responses. Actually what he hears is recorded laughter (or the lack of it). His own laughter, however, can be recorded over the microphone. He is asked to give a written rating of the joke, and his response is scored from 0 (not funny at all) to 4 (extremely funny). His audible response is scored in the same way, from 0 (no perceptible sound) to 4 (loud laughter).

Nosanchuk and Lightstone found that playing laughter to their subjects immediately after the jokes influenced the subjects' written opinion of the jokes only slightly. The subjects' audible response, however, was profoundly influenced; in general jokes followed by laughter elicited laughter twice as often as jokes not followed by laughter.

After the experiment the subjects were told that they had been involved in an experiment on conformity rather than one on humor. They showed no sign of stress, and most of them expressed much interest in the work. Nosanchuk and Lightstone write that their experiment "appears to distinguish between public and private conformity, and seems to be quite painless for the subjects."

Corrugahorns

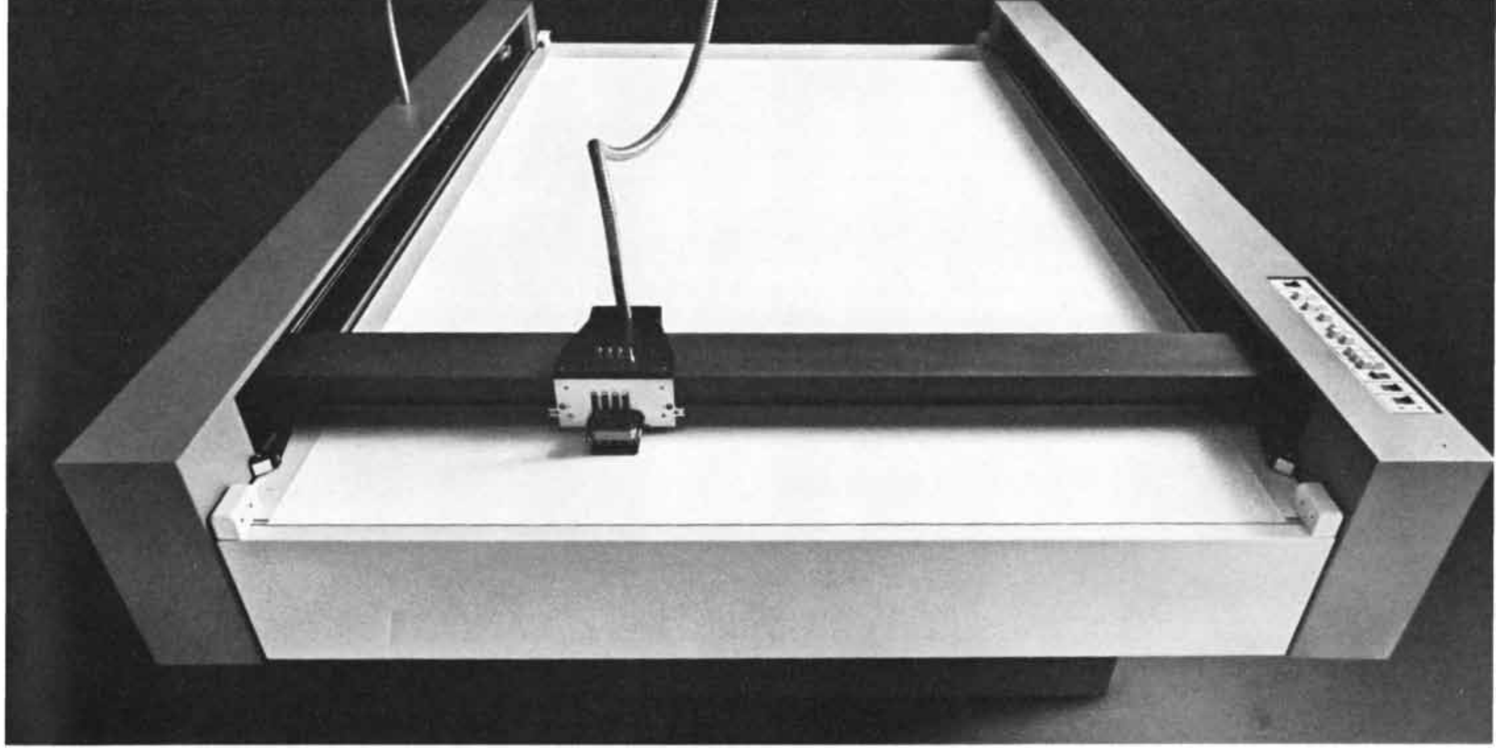
A few years ago an acoustical toy known variously as the Hummer, the Freeka and the Whirl-A-Sound was widely sold in the U.S. It was simply a corrugated flexible plastic tube about an inch in diameter and three feet long. When the tube was held at one end and whirled around the head, it emitted a loud, pure tone. The faster the tube was whirled, the higher the tone was.

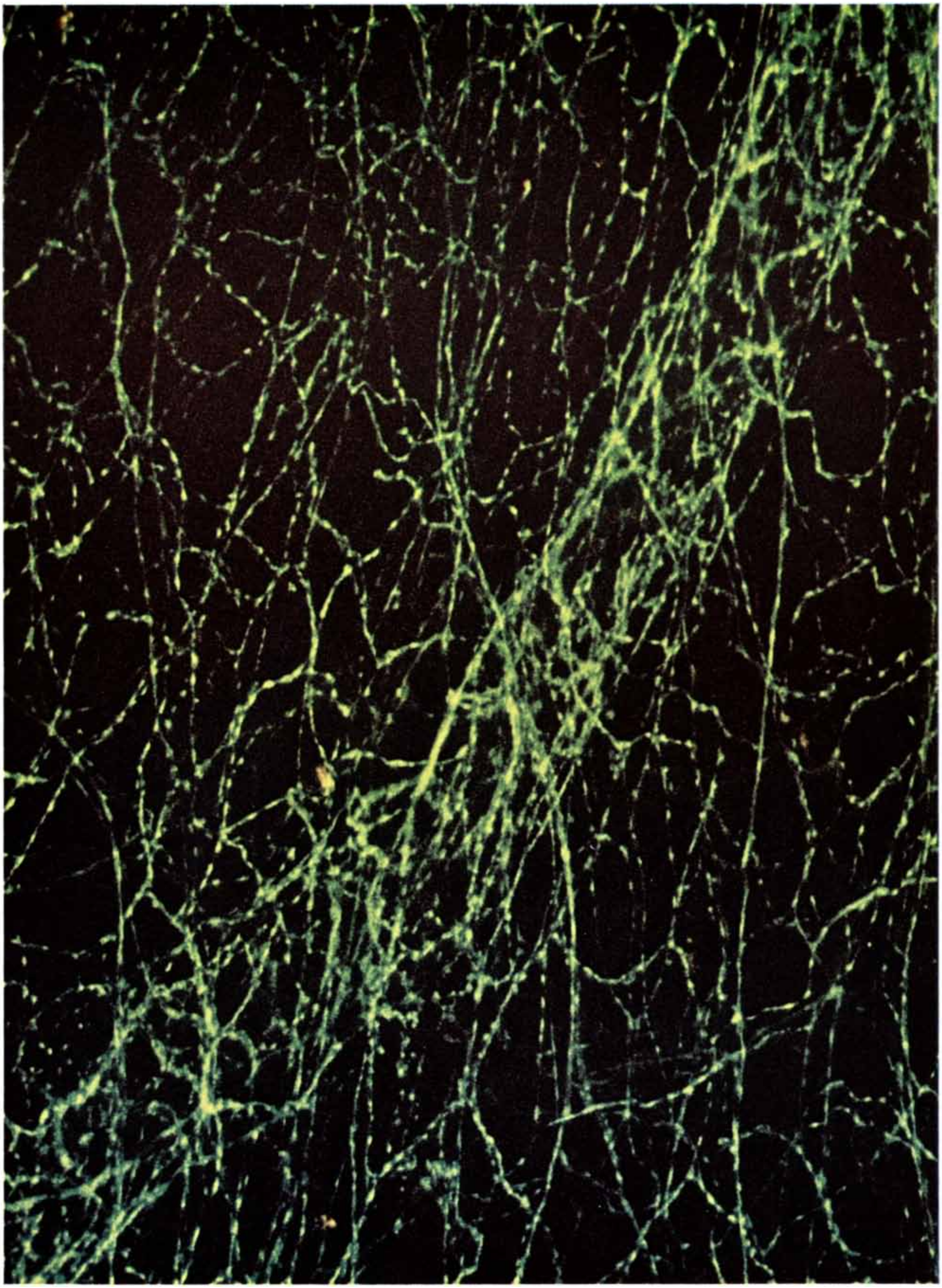
The Hummer intrigued Frank S. Crawford of the University of California at Berkeley. He writes in the *American Journal of Physics:* "I have been playing with this toy off and on for about a year and have learned something about how it works."

Crawford found that a smooth plastic tube will not sing; the corrugations are essential for the Hummer to work. Each note is one of the natural harmonics of the tube, and it is produced when the velocity of the air flowing through the tube is such that the "bump frequency" (the frequency at which the air bumps into the corrugations) is equal to the frequency of the harmonic. One of Crawford's experiments entailed holding the tube outside the window of his van truck with one end pointing into the wind. "The tube starts to sing . . . at about 15 miles per hour," he observes. "By about 35 miles per hour I get the fifth harmonic. . . . I get the 11th harmonic at about 80 miles per hour."

In order to attain even higher harmonics under controlled conditions Crawford got a plastic wastebasket 15 inches in diameter, cut a one-inch hole in its bottom and inserted one end of a Hummer. Then he inverted the basket in a large tub of water. "By pushing the basket down or pulling it up I could use the water as a piston to force air through the Hummer," he reports. "I can easily get up to the 11th harmonic . . . , whereas I can only reach the seventh harmonic, with great effort, by whirling. With some practice I can play typical bugle songs."

On the basis of these results and the reactions of neighborhood children, Crawford was encouraged to develop several new musical instruments. The first was the inverted-wastebasket water piston, which he named the Water Pipe. A more sophisticated family of instruments he based on the much narrower corrugated flexible metal tubes used in gas plumbing, which can be played by mouth. This family he calls the Corrugahorns; it includes the E-Flat and E-Natural Gas-Pipe Corrugahorn Bugles (Corrugabugles) and the E-Natural Gas-Pipe Blues Corrugahorn. The two E-Natural Corrugahorns are in an easy guitar key for stringed accompaniment, and the Blues Corrugahorn "is a very pleasant instrument on which to play 12-bar blues in E-Natural."





SYMPATHETIC-NERVE TERMINALS from the iris of a rat's eye emit a green glow after treatment with formaldehyde, showing that they contain noradrenaline, one of the neurotransmitters. The terminals, studded with varicosities where the noradrenaline is stored, are enlarged 2,400 diameters in this fluorescence micrograph made by David Jacobowitz of the National Institute of Mental Health.

NEUROTRANSMITTERS

These chemicals released from nerve-fiber endings are the messengers by means of which nerve cells communicate. Neurotransmitters mediate functions ranging from muscle contraction to the control of behavior

by Julius Axelrod

In 1901 the noted English physiologist J. N. Langley observed that the injection of an extract of the adrenal gland into an animal stimulated tissues innervated by the sympathetic nerves: the nerves of the autonomic nervous system that increase the heart rate, raise the blood pressure and cause smooth muscles to contract. Just three years before that John J. Abel of Johns Hopkins University had isolated the hormone adrenaline from the adrenal gland, and so Langley's observation prompted T. R. Elliott, his student at the University of Cambridge, to inject adrenaline into experimental animals. Elliott saw that the hormone, like the crude extract, produced a response in a number of organs that was similar to the response evoked by the electrical stimulation of sympathetic nerves. He thereupon made the brilliant and germinal suggestion that adrenaline might be released from sympathetic nerves and then cause a response in muscle cells with which the nerves form junctions. Elliott thus first enunciated the concept of neural communication by means of chemical transmitters. A neurotransmitter is a chemical that is discharged from a nerve-fiber ending. It reaches and is recognized by a receptor on the surface of a postsynaptic nerve cell or other excitable postjunctional cell and either stimulates or inhibits the second cell. Today it is clear that many different neurotransmitters influence a variety of tissues and physiological processes. Neurotransmitters make the heart beat faster or slower and make muscles contract or relax. They cause glands to synthesize hormone-producing enzymes or to secrete hormones. And they are the agents through which the brain regulates movement and changes mood and behavior.

Elliott's concept of chemical neurotransmission was accepted slowly. Langley, who disliked theories of any kind, discouraged further speculation by Elliott until more facts were available. That took time. The first definite evidence for neurochemical transmission was obtained in 1921 by Otto Loewi, who was then working at the University of Graz in Austria, through an elegant and crucial experiment. Loewi put the heart of a frog in a bath in which the heart could be kept beating. The fluid bathing the heart was allowed to perfuse a second heart. When Loewi stimulated the first heart's vagus nerve (a nerve of the parasympathetic system that reduces the heart rate), the beat of the second heart was slowed, showing that some substance was liberated from the stimulated vagus nerve, was transported by the fluid and influenced the perfused heart. The substance was later identified by Sir Henry Dale as acetylcholine, one of the first neurotransmitters to be recognized. In a similar experiment the stimulation of the accelerans nerve (the sympathetic nerve that increases the heart rate) of a frog heart speeded up the beat of an unstimulated perfused heart. In 1946 the Swedish physiologist Ulf von Euler isolated the neurotransmitter of the sympathetic system and identified it as noradrenaline.

The Transmitters

To be classed as a neurotransmitter a chemical should fulfill a certain set of criteria. Nerves should have the enzymes required to produce the chemical; when nerves are stimulated, they should liberate the chemical, which should then react with a specific receptor on the postjunctional cell and produce a biological response; mechanisms should be available to terminate the actions of the chemical rapidly. On the basis of these criteria two compounds are now established as neurotransmitters: acetylcholine and noradrenaline. Nerves that contain them are respectively called cholinergic and noradrenergic nerves. There are a number of other nerve chemicals that meet many of the listed criteria but have not yet been shown to meet them all. These "putative" transmitters are dopamine, adrenaline, serotonin, octopamine, histamine, gamma aminobutyric acid, glutamic acid, aspartic acid and glycine.

This article will deal mainly with one class of neurotransmitters, the catecholamines, since more is known about these compounds than about some other transmitters and since many of the principles governing their disposition appear to govern those of transmitters in general. The catecholamines include noradrenaline (also known as norepinephrine), dopamine and adrenaline (or epinephrine). They have in common a chemical structure that consists of a benzene ring on which there are two adjacent hydroxyl groups and an ethylamine side chain. Noradrenaline is present in peripheral nerves, the brain and the spinal cord and in the medulla, or inner core, of the adrenal gland. In peripheral tissues and in the brain noradrenaline acts as a neurotransmitter, that is, it exerts most of its effect locally on postjunctional cells. In the adrenal medulla it functions as a hormone, that is, it is released into the bloodstream and acts on distant target organs. Dopamine, once thought to be simply an intermediate in the synthesis of noradrenaline and adrenaline, is also a neurotransmitter in its own right in the brain, where it functions in nerves that influence movement and behavior. The third catecholamine, adrenaline, is largely concentrated in the adrenal medulla. It is discharged into the bloodstream in fear, anger or other stress and acts as a hormone on a number of organs, including the heart, the liver and the intestines. Just in the past year it has developed that adrenaline is probably also a neurotransmitter, since it is found in nerves in the brain.

Techniques developed a decade ago in Sweden made it possible to visualize catecholamines in neurons directly, by the fluorescent glow they emit after treatment with formaldehyde vapor. Fluorescence photomicrography, electron microscopy and radioautography have revealed the structure and functioning of the sympathetic nerve cell in great detail [*see illustration below*]. The neuron has a cell body and a long axon, or main fiber, that branches into a large number of terminals. Each nerve ending is studded with varicosities, or swellings, that look like beads on a string, so that a single sympathetic neuron can innervate thousands of other cells: "effector" cells.

In 1960 Georg Hertting, Gordon Whitby and I were able to show that radioactive noradrenaline (noradrenaline in which tritium, the radioactive isotope of hydrogen, has been substituted for some of the hydrogen atoms) is taken up selectively and retained in sympathetic nerves. In my laboratory at the National Institute of Mental Health we went on to find out where the neurotransmitter is stored within the nerve cell. Electron micrographs of sympathetic-neuron varicosities reveal large numbers of vesicles with dark, granular cores. When photographic film was exposed to tissues from rats injected with labeled noradrenaline, the silver grains developed by the radiation from the radioactive hydrogen atoms were strikingly localized over the granulated vesicles [*see illustrations on opposite page*]. This indicated that it is in those vesicles that noradrenaline is stored within the nerve.

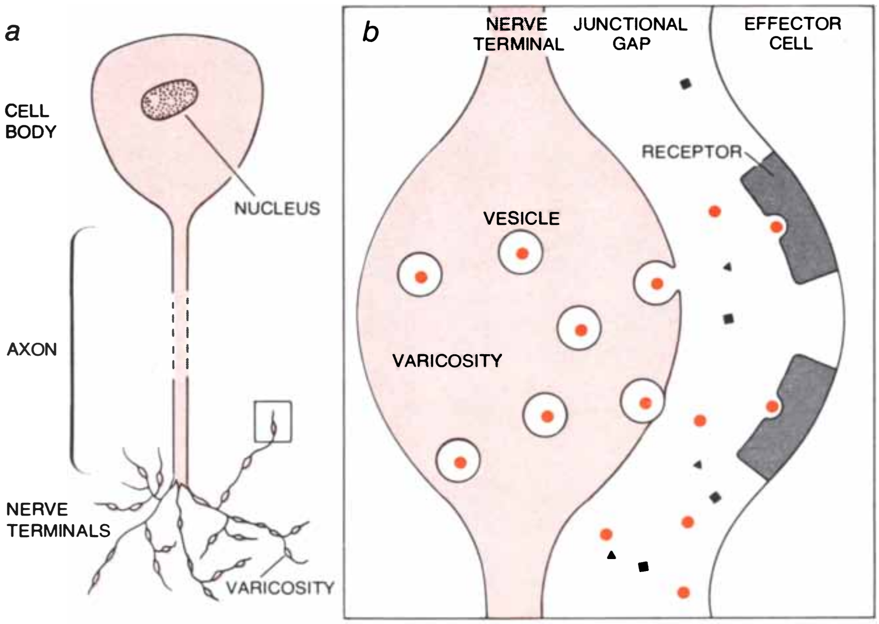
Synthesis and Release

The process leading to the synthesis of catecholamine transmitters begins in the cell body, which has the machinery for making the four enzymes needed for their formation: tyrosine hydroxylase, dopa decarboxylase, dopamine beta-hydroxylase (DBH) and phenylethanolamine N-methyltransferase (PNMT). The enzymes synthesized in the cell body are carried down the axon by a natural flowing process to the nerve endings, where the synthesis of the catecholamines is achieved.

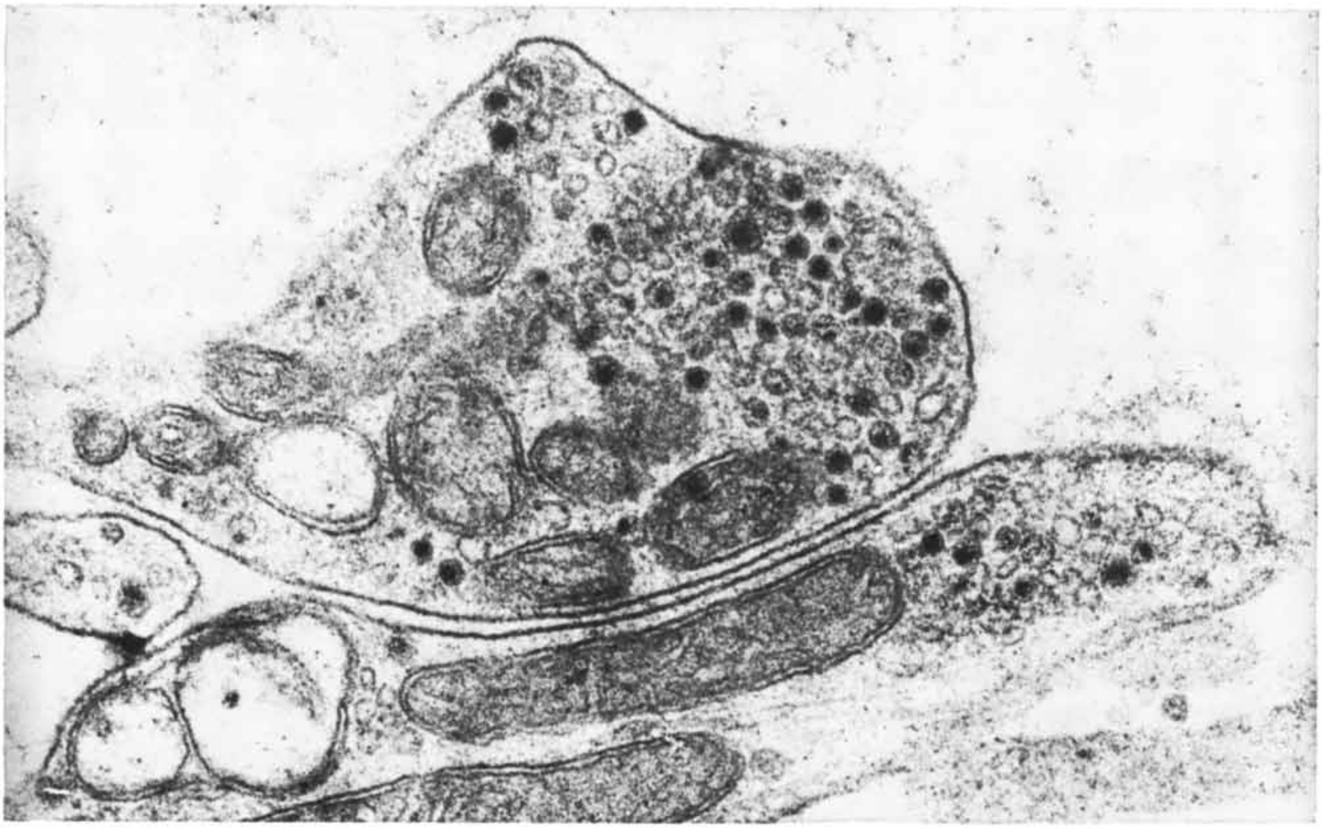
The discharge of neurotransmitter from the nerve endings caps a complex series of events. When a nerve is stimulated, its membrane is depolarized, with sodium moving into the nerve as potassium comes out; the nerve signal propagates as a wave of depolarization that moves along the nerve axon to the endings. As Bernhard Katz of University College London first showed for acetylcholine, the depolarization causes a quantum—a packet or spurt, as it were—of the transmitter to be discharged from the nerve ending into the synaptic cleft.

Biochemical evidence recently obtained in our laboratory and others shows that noradrenaline is released from nerves in much the same way. The vesicles in the endings contain not only noradrenaline but also the enzyme, DBH, that converts dopamine into noradrenaline. When the sympathetic nerve is stimulated electrically, noradrenaline and the enzyme are released in about the same proportions in which they are present in the vesicles. The only way that could happen would be through the fusion of the vesicle with the outer membrane of the nerve, followed by the formation of an opening large enough to allow molecules of noradrenaline to be extruded along with the much larger molecules of the enzyme. Such a release mechanism is called exocytosis. The detailed events whereby the vesicle fuses with the neural membrane and makes an opening to discharge its soluble contents are uncertain, as is the subsequent fate of the vesicle. We do know that certain conditions prevent the release of noradrenaline and DBH. One is the presence of vinblastin, a compound that breaks down the protein structures in nerve cells called neurotubules. Another is the presence of cytocholasin-beta, a substance that disrupts the function of the contractile filament system in cells. A third is the absence of calcium. These findings suggest that the long, tubelike protein structures may orient the vesicles to a site on the neuronal membrane from which the release occurs. It is well known that microfilaments in cells other than nerves, such as muscle cells, can be activated by calcium so that they contract. It is therefore possible that depolarization causes calcium to activate a contractile filament on the neural membrane, which thereupon contracts to make an opening large enough so that the soluble contents of the vesicle can be discharged.

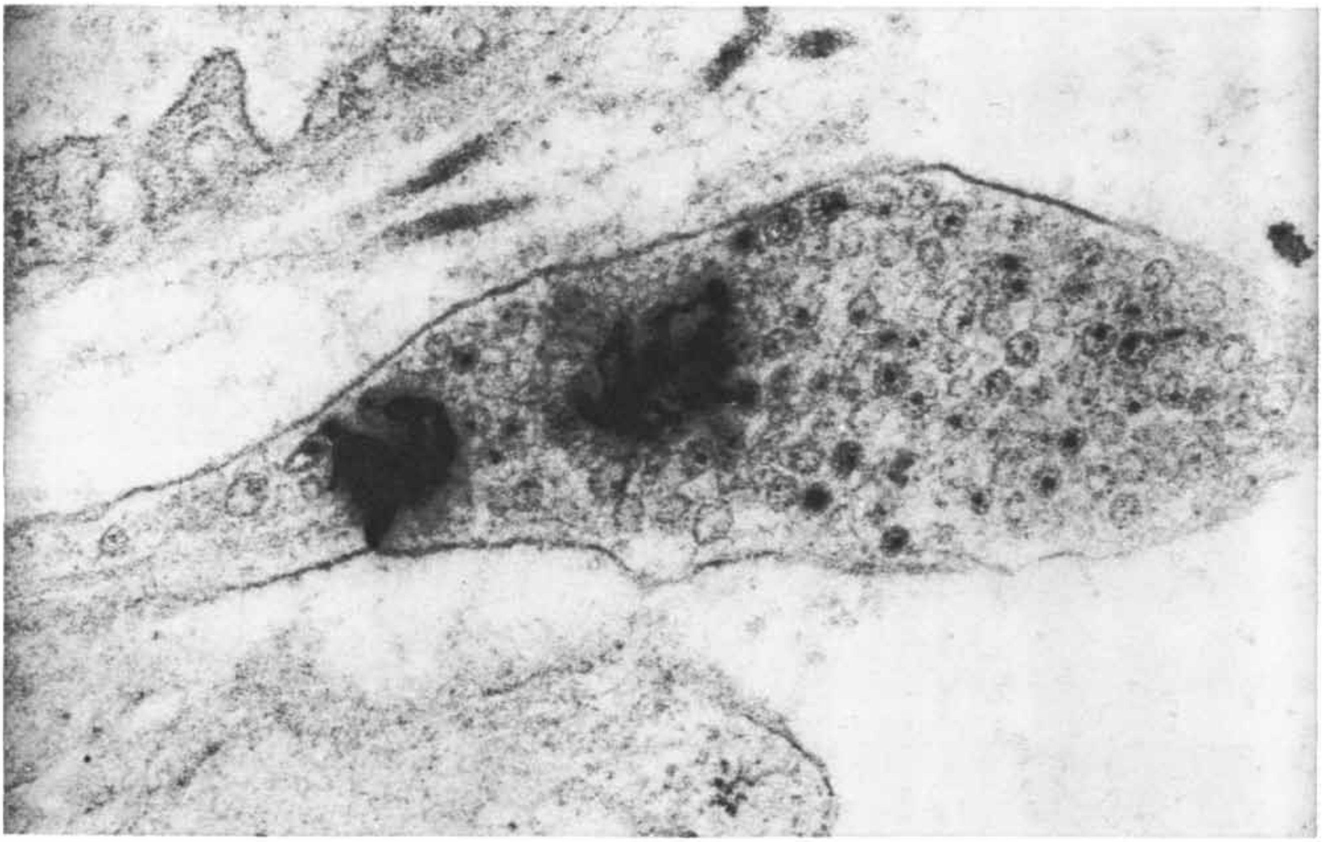
The observation that DBH is released from nerves suggested to Richard Weinshilboum, a research associate in my laboratory, that the enzyme might find its way into the bloodstream. We devised a sensitive assay for the enzyme and found it is indeed present in the blood, and we and others went on to measure the amount of the enzyme (which is found specifically in sympathetic nerves) in a



SYMPATHETIC NEURON, or nerve cell (*a*), consists of a cell body, a long axon and numerous nerve terminals studded with varicosities. Enzymes involved in the synthesis of the transmitters are made in the nucleus and transported down the axon to the varicosities, where the transmitters are manufactured and stored. A varicosity and its junction with another cell are enlarged (*b*). The transmitter (*colored dots*) is stored in vesicles. When the nerve is excited and depolarized, vesicles move to the cell membrane and fuse with it, releasing transmitter into the junctional gap. The transmitter reaches receptors on the effector cell that recognize only it, not any other chemicals (*black shapes*) that are present in the gap.



VARICOSITIES on noradrenaline-containing nerve endings from a rat pineal gland are enlarged 90,000 diameters in an electron micrograph made by Floyd Bloom of the National Institute of Mental Health. The varicosities contain vesicles, many with dense cores.



RADIOACTIVE NORADRENALINE is shown by radioautography to be localized in the vesicles. A pineal-tissue sample was taken from rats injected with labeled noradrenaline. The radioactivity developed silver grains (*black blots*) in a photographic film laid on the sample. The developed grains were strikingly localized over the vesicles, which were thus identified as sites of transmitter storage.

variety of disease states. It is low in the hereditary disorder of the autonomic system called familial dysautonomia and in Down's syndrome (mongolism), and it is high in torsion dystonia (a neurological disease involving muscle spasticity), in neuroblastoma (a cancer of nervous tissue) and in certain forms of hypertension. The findings suggest that in each of these diseases there are abnormalities in the functioning of the sympathetic nervous system.

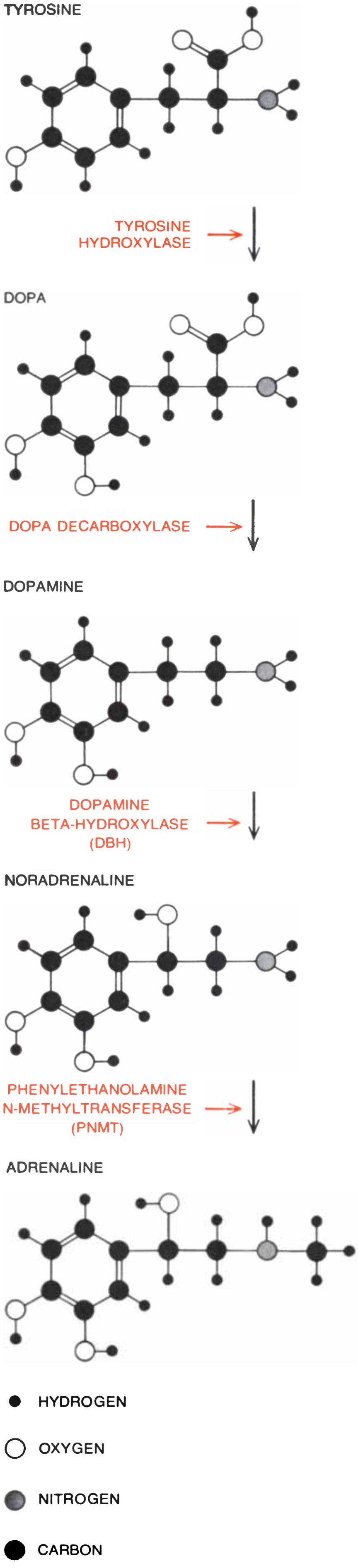
Action and Inactivation

Once the neurortansmitter is liberated it diffuses across the cleft between the nerve terminal and adjacent cells. The capacity of a neighboring effector cell to respond to the transmitter then depends on the ability of a receptor on the postjunctional cell's surface to selectively recognize and combine with the neurotransmitter. When the receptor and transmitter interact, a series of events is triggered that causes the effector cell to carry out its special function. Some of these responses occur rapidly (in a fraction of a second), as in the propagation of nerve transmission across a synapse; others occur slowly, in minutes or sometimes hours, as in the synthesis of intracellular enzymes. There are two receptors that recognize noradrenaline, alpha and beta adrenergic receptors, and there is one for dopamine. These receptors can be distinguished from one another by the specific response each elicits and by the ability of specific drugs to block those responses.

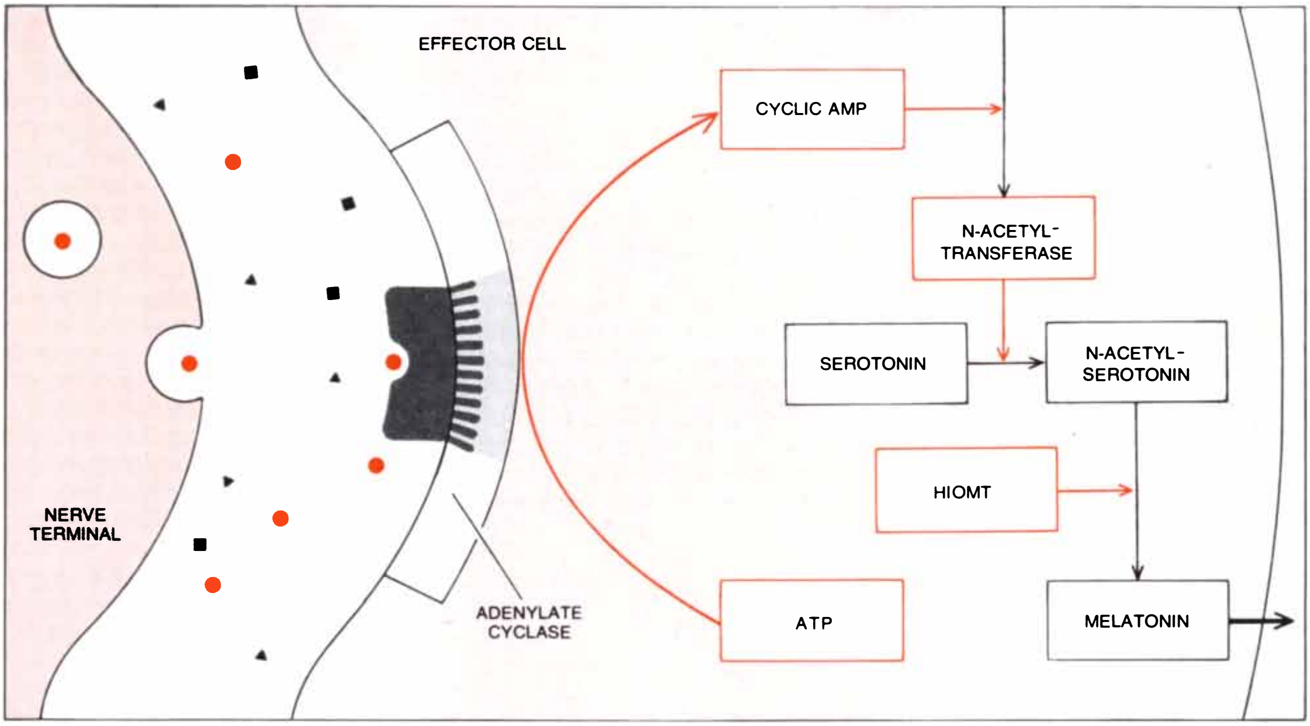
The beta adrenergic receptors turn on an effector cell, and they do so by means of adenosine 3′5′ monophosphate, or cyclic AMP, the universal "second messenger" that mediates between hormones and many cellular activities elicited by those hormones [see "Cyclic AMP," by Ira Pastan; Scientific American, August, 1972]. Investigators have traced several of the steps in the activation of the receptor by noradrenaline by studying the interaction of noradrenaline and fat cells or cells of the liver or of the pineal gland. We have found the pineal, which makes a hormone called melatonin that inhibits the activity of sex glands, particularly suitable because it is heavily supplied with nerves containing noradrenaline [see "The Pineal Gland," by Richard J. Wurtman and Julius Axelrod; Scientific American, July, 1965]. Melatonin is synthesized in a number of steps, one of which, the conversion of serotonin to N-acetylserotonin, is catalyzed by the enzyme serotonin N-acetyltransferase. It is that enzyme's synthesis that is controlled by the beta adrenergic receptor. When noradrenaline is released from a nerve innervating the pineal, it interacts with beta adrenergic receptors on the outside of the membrane of a pineal cell. Once a receptor is occupied by noradrenaline, the enzyme adenylate cyclase, on the inner surface of the cell membrane, is activated. The adenylate cyclase then converts the cellular energy carrier ATP to cyclic AMP, which in turn stimulates the synthesis of serotonin N-acetyltransferase [*see illustration on opposite page*]. This complex series of events can be turned off by propranolol, a drug that prevents the noradrenaline from combining with the beta adrenergic receptor.

The adenylate cyclase system is involved in scores of biological actions. The ability of the pineal cell to carry out its special function, the manufacture of melatonin, by utilizing the almost universal adenylate-cyclase system depends on the presence of receptors on the cell surface that can specifically recognize noradrenaline and of the enzyme hydroxyindole-O-methyltransferase, uniquely present in the pineal cell, that can convert N-acetylserotonin to melatonin.

Once the neurotransmitter has interacted with the postjunctional cell, its actions must be rapidly terminated; otherwise it would exert its effects for too long and precise control would be lost. In the cholinergic nervous system the acetylcholine is rapidly inactivated by the enzyme acetylcholinesterase, which metabolizes the transmitter. In the past 10 years it has become clear that the inactivation of neurotransmitters through enzymatic transformation is the exception rather than the rule. Catecholamine neurotransmitters are metabolized by two enzymes, catechol-O-methyltransferase (COMT) and monoamine oxidase (MAO); the latter is a particularly important enzyme that removes the amino (NH_2) group of a wide variety of compounds, including serotonin, noradrenaline, dopamine and adrenaline. There are enzyme-inhibiting chemicals that



PRODUCTION of catecholamine transmitters is accomplished by four enzymes (*color*) in sympathetic-nerve terminals and in the adrenal gland. Tyrosine, an amino acid, is transformed into the intermediate dopa. Removal of a carboxyl (COOH) group forms dopamine, which is itself a transmitter and is also the precursor of the transmitter noradrenaline. In the adrenal gland the process continues with the addition of a methyl (CH_3) group to form adrenaline.



MODE OF ACTION of a transmitter is exemplified by the effect of noradrenaline on a pineal cell. Noradrenaline (*colored dots*) released from a nerve ending binds to a beta-adrenergic receptor on the pineal-cell surface. The receptor thereupon activates the enzyme adenylate cyclase on the inside of the pineal-cell membrane. The activated adenylate cyclase catalyzes the conversion of adenosine triphosphate (ATP) into cyclic adenosine monophosphate (AMP). The cyclic AMP stimulates synthesis of the enzyme N-acetyltransferase; the enzyme converts serotonin into N-acetylserotonin. This is transformed in turn by the pineal cell's specific enzyme, hydroxyindole-O-methyltransferase (HIOMT), to form melatonin, the pineal-gland hormone that acts on the sex glands.

can prevent COMT and MAO from carrying out their biochemical transformations; when such inhibitors were administered, the action of noradrenaline was found still to be rapidly terminated. There had to be a method of rapid inactivation other than enzymatic transformation.

In order to track down such a mechanism we injected a cat with radioactive noradrenaline. The labeled transmitter persisted in tissues that were rich in sympathetic nerves for many hours, long after its physiological actions were ended, indicating that radioactive noradrenaline was taken up in the sympathetic nerves and held there. My colleagues and I designed a simple experiment to prove this. Sympathetic nerves innervating the left salivary gland in rats were destroyed by removing the superior cervical ganglion on the left side of the neck; about seven days after this operation the noradrenaline nerves of the salivary gland on the right side were intact, whereas the nerves on the left side had completely disappeared. When radioactive noradrenaline was injected, the transmitter was found in the right salivary gland but not in the left one. We also found that in cats injected with radioactive noradrenaline the transmitter was released when the sympathetic nerves were stimulated electrically. The experiments clearly demonstrated that noradrenaline is taken up into, as well as released from, sympathetic nerves. As a result of these experiments we postulated that noradrenaline is rapidly inactivated through its recapture by the sympathetic nerves; once it is back in the nerves, of course, the neurotransmitter cannot exert its effect on postjunctional cells. Leslie L. Iversen of the University of Cambridge has since shown that this neuronal recapture by sympathetic nerves is highly selective for noradrenaline or compounds resembling it in chemical structure. Recent work indicates that uptake by nerves may be the most general mechanism for the inactivation of neurotransmitters.

Regulation

Chemical transmitters in sympathetic nerves (and presumably in other nerves) are in a state of flux, continually being synthesized, released, metabolized and recaptured. The activity of nerves can also undergo marked fluctuation during periods of stress. In spite of all these dynamic changes the amount of catecholamines in tissues remains constant. This is owing to a variety of adaptive mechanisms that alter the formation, release and response of catecholamines. There are fast regulatory changes that require only fractions of a second and slower changes that take place after minutes or even hours.

When sympathetic nerves are stimulated, the conversion of tyrosine to noradrenaline in them is rapidly increased. The increased nervous activity specifically affects tyrosine hydroxylase, the enzyme that converts tyrosine to dopa, because its activity is inhibited by noradrenaline and dopamine. Any increase in nerve-firing brought on by stress, cold and certain drugs lowers the level of catecholamines in the nerve terminals. This reduces the negative-feedback effect of noradrenaline and dopamine on tyrosine hydroxylase, so that more tyrosine is converted to dopa, which in turn is converted to make more catecholamines. Conversely, when nerve activity is decreased, the catecholamine level rises, slowing down the conversion of tyrosine to dopa by once again inhibiting the tyrosine hydroxylase.

Another rapid regulation is accomplished at the nerve terminal itself, where the alpha adrenergic receptors are situated. When the alpha receptors are activated, they diminish the release of noradrenaline from nerve terminals into the synaptic cleft. When too much noradrenaline is released, it accumulates in the synaptic cleft; when the catecholamine level is high enough, it activates the alpha receptors on the presynaptic nerve terminals and shuts off further release of the neurotransmitter.

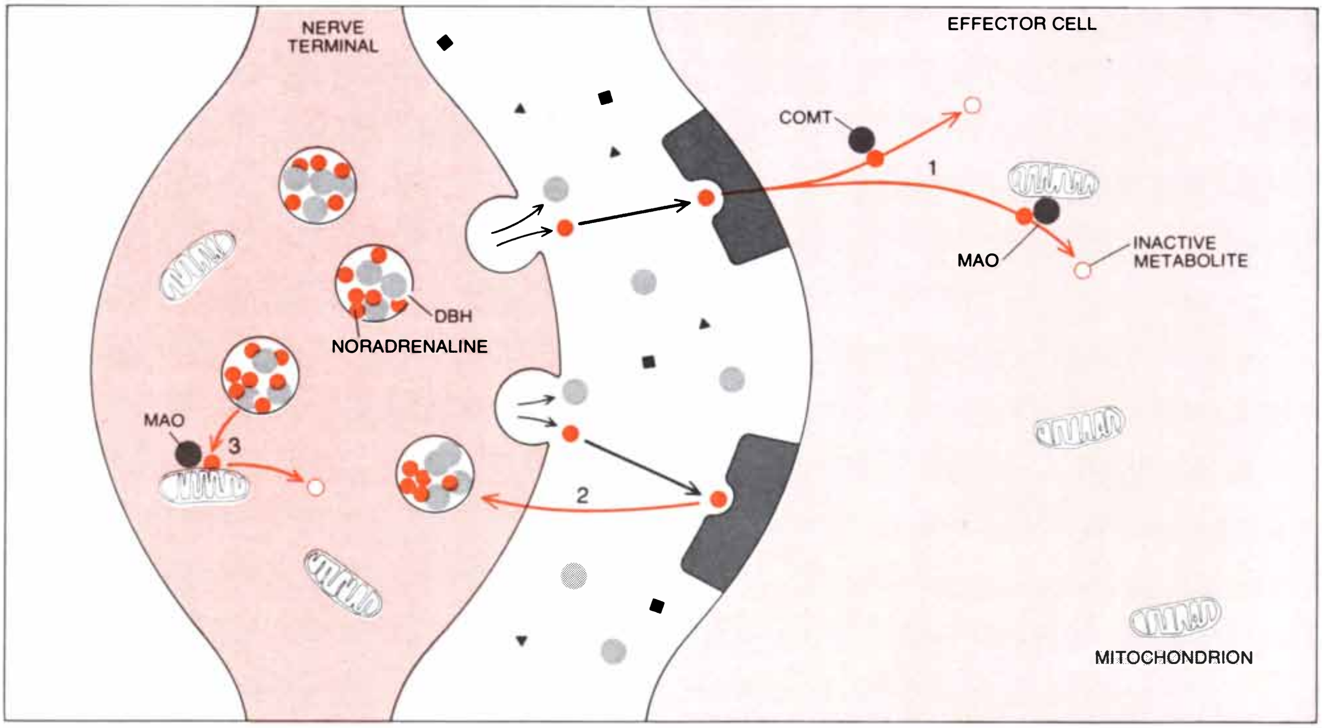
A slower regulatory process is brought on by prolonged firing of sympathetic nerves, which can step up the manufacture of the catecholamine-synthesizing enzymes tyrosine hydroxylase, DBH and (to a lesser extent) PNMT; the rise in the enzyme level enables the nerves to make more neurotransmitter. We discovered this phenomenon of increased enzyme synthesis when we gave animals reserpine, a versatile drug that lowers the blood pressure and incidentally increases sympathetic-nerve firing (which tends to raise the pressure) by a reflex action. The reserpine brought about a gradual increase in tyrosine hydroxylase and DBH in sympathetic nerves and the adrenal gland and of PNMT in the adrenal gland. Increases in these enzymes were also found in animals exposed to stress, cold, physical restraint, psychosocial stimulation or insulin injection. When the synthesis of proteins was prevented by drugs, on the other hand, there was no elevation in enzyme activity after reserpine was given. This indicated that increased nerve activity stimulates the synthesis of new molecules of tyrosine hydroxylase, DBH and PNMT; with a greater need for neurotransmitters there is a compensatory increase in the synthesis of enzymes that catalyze the making of these transmitters.

In order to learn whether the command for increased synthesis of new tyrosine hydroxylase and DBH molecules can be transmitted from one nerve to another we cut the nerve innervating certain noradrenaline cell bodies—the superior cervical ganglia—on one side. When nerves were then stimulated reflexly by reserpine, there was an elevation of tyrosine hydroxylase and DBH levels in the innervated ganglia but not in the denervated ones. The experiment showed that one nerve can transmit information to another nerve (presumably by means of a chemical signal) that causes the postsynaptic nerve to make new enzyme molecules.

Sensitivity

In 1855 the German physiologist J. L. Budge observed that when the nerves leading to a rabbit's right eye were destroyed, the pupil of that eye became more dilated than the left pupil. The phenomenon was later explained by the American physiologist Walter B. Cannon, who postulated that as a result of denervation the effector cells somehow become more responsive. He called this effect the "law of denervation supersensitivity." Subsequent work showed that denervation supersensitivity is caused by two separate mechanisms, one presynaptic and the other postsynaptic. When nerves are destroyed, presynaptic inactivation by recapture is abolished, thereby leaving the neurotransmitters to react with the postsynaptic site longer.

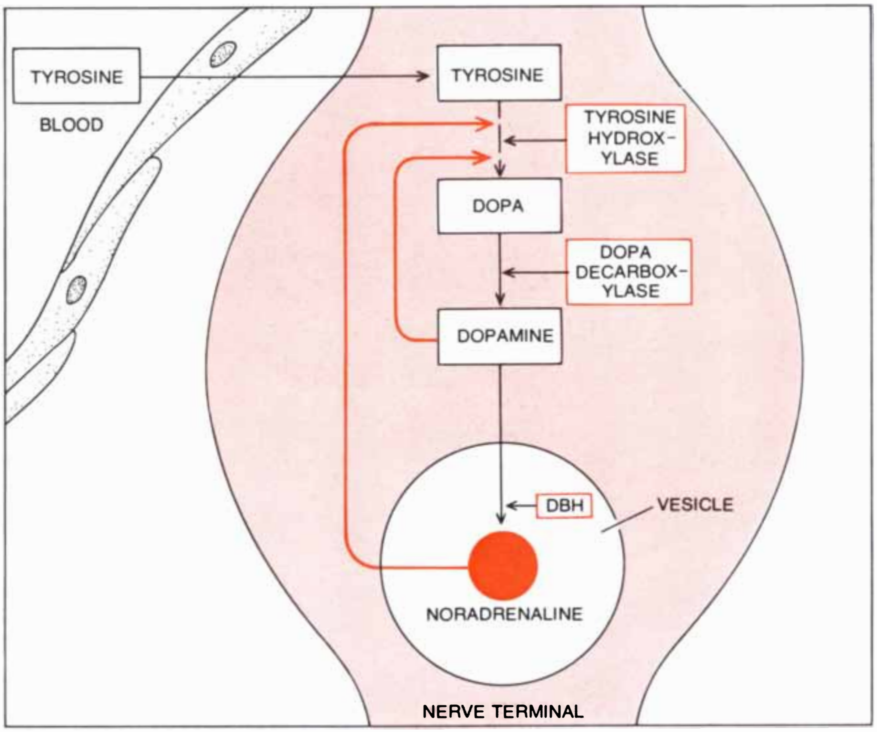
Denervation also causes a profound change in the degree of activity of the postjunctional cell. Recent work with the pineal gland in our laboratory has suggested a hypothesis for supersensitivity, and also for subsensitivity, in postjunctional cells. As we have seen, noradrenaline stimulates the synthesis of the enzyme serotonin N-acetyltransferase through a beta adrenergic receptor in the postjunctional pineal cell. When the nerves to pineal cells are destroyed (or depleted of noradrenaline by the administration of reserpine), the pineal cells become 10 times as responsive to noradrenaline; that is, when the postjunctional cell is deprived of its neurotransmitter for a period of time, it takes just



RELEASE AND INACTIVATION of the neurotransmitter noradrenaline is shown in more detail. The enzyme DBH is stored in the noradrenergic nerve terminals along with the transmitter and is released with it into the junctional gap. The noradrenaline binds to the receptors on the effector cell, eliciting that cell's response as shown in the illustration on the preceding page. Then the noradrenaline's action is terminated either through metabolism (*1*) by the enzymes catechol-O-methyltransferase (COMT) and/or monoamine oxidase (MAO), or by recapture and storage (*2*) in the presynaptic sympathetic-nerve terminal; the latter is the more important process. MAO, stored in the membrane of mitochondria, also inactivates noradrenaline that leaks out of vesicles (*3*).

one packet of catecholamines to cause the same increase of N-acetyltransferase in the cell as 10 packets of transmitters would cause in a normally innervated cell. If, on the other hand, the pineal cell is exposed to an excessive amount of catecholamines for a period of time, it becomes less responsive: a larger amount of the transmitter is required to produce the same increase in N-acetyltransferase.

These experiments suggest that changes in the responsiveness of excitable cells are the result of an alteration in the "avidity" with which the receptor binds the neurotransmitters. If the receptor is exposed to small amounts of catecholamine for some time, it reacts with the neurotransmitter easily; if too many neurotransmitter molecules bombard the receptor, it becomes less responsive. Depending on the tissue, this change in sensitivity can come within hours or days, so that it is an effective adaptive mechanism for excitable cells. It is possible that the tolerance that is often developed to a drug taken in excess may reflect subsensitivity on the part of cells that respond to the drug.



RAPID REGULATION of catecholamine synthesis is accomplished by a feedback mechanism: a buildup of dopamine and noradrenaline inhibits (*colored arrows*) the activity of tyrosine hydroxylase, which catalyzes the first step in the synthesis. An increase in nerve activity reduces the amount of dopamine and noradrenaline in the terminal, removing the inhibition; tyrosine hydroxylase activity increases and more transmitter is synthesized.

Role in the Brain

The brain has billions of nerves that talk to one another by means of neurotransmitters. Neurobiologists are just beginning to unravel the complex biochemistry and physiology of chemical transmission in the brain. Many different neurotransmitters function in brain neurons, but because there are more precise methods of measuring catecholamines and drugs are available that perturb their formation, storage, release and metabolism, we know more about brain catecholamines than about the other neurotransmitters. Fluorescence photomicrography and drugs that selectively destroy catecholamine-containing nerves have made it possible to locate the noradrenaline, dopamine and serotonin cell bodies and trace the pathways of their axons and nerve endings [*see illustrations on page 71*]. The cell bodies of the dopamine-containing nerves are in the area of the brain stem called the substantia nigra, whence the dopaminergic axons course through the brain stem, many of them terminating in the caudate nucleus. The dopamine-containing tracts in the caudate nucleus play an important role in the integration of movement.

The elucidation of the biochemistry and pharmacology of dopamine in the brain has led to the development of a powerful treatment of a crippling disease, Parkinsonism. The Swedish pharmacologist Arvid Carlsson noted in 1959 that when reserpine was given to rats, it sharply reduced the dopamine content of the caudate nucleus in the brain and also caused a Parkinson-like tremor. The administration of dopa, a dopamine precursor that can get from the blood into the brain more easily than dopamine, reversed the tremors. These findings prompted Oleh Hornykiewicz, who was then working at the University of Vienna, to measure the content of dopamine in the brain of patients who had died of Parkinson's disease. He found that there was virtually no dopamine in the caudate nucleus. The finding led directly to a major therapeutic advance by George C. Cotzias of the Brookhaven National Laboratory: when dopa, the dopamine precursor that can cross the blood-brain barrier, is administered, it makes up the dopamine deficiency and effectively relieves the symptoms of Parkinson's disease. This is a good example of how basic research can sometimes lead rapidly to a new treatment for a disease.

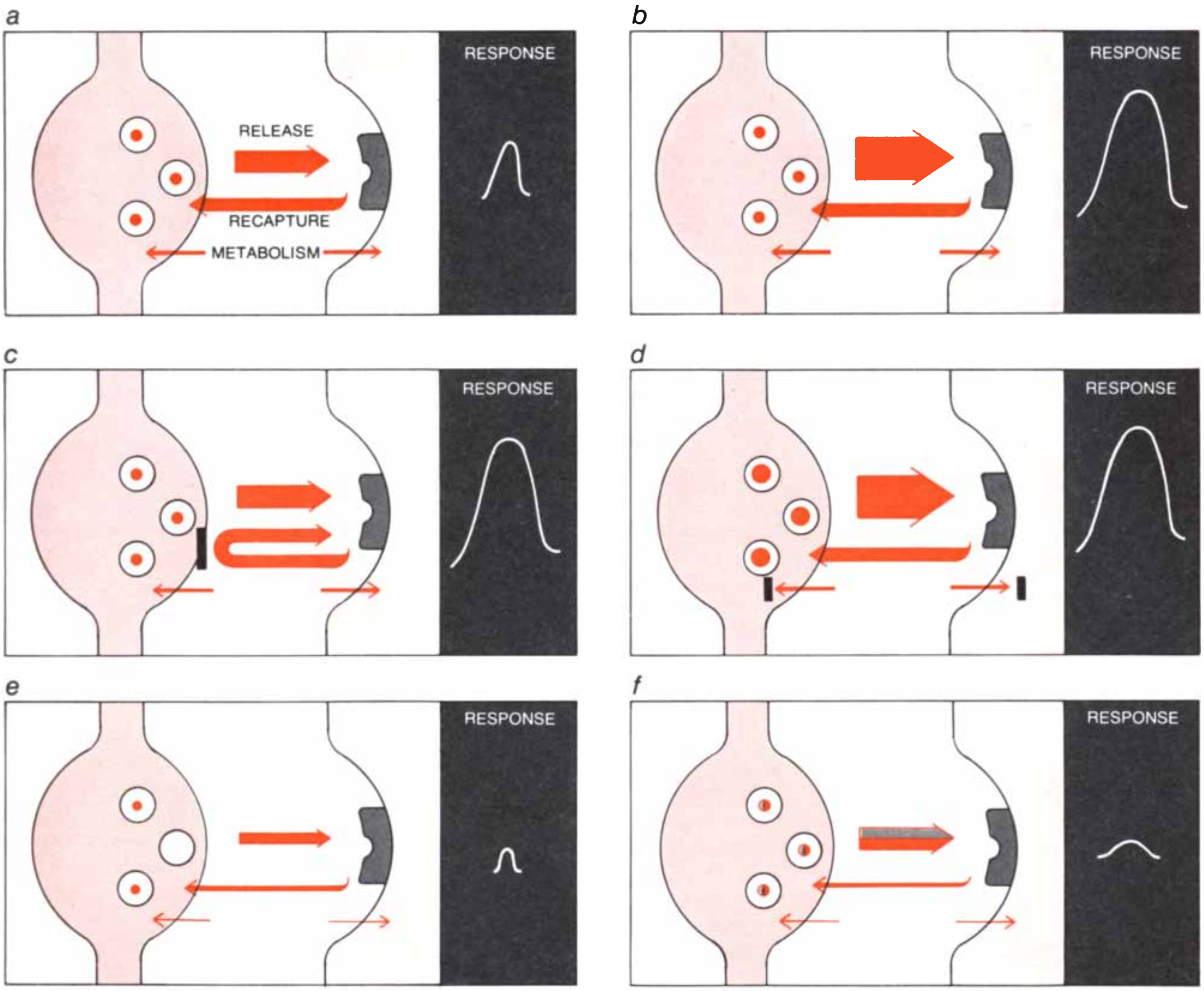
There are two main nerve tracts containing noradrenaline in the brain, the dorsal and ventral pathways. The cell bodies of the noradrenaline-containing tracts are found in the lower part of the brain in the area called the locus ceruleus, or "blue place." Noradrenaline-containing nerve tracts are highly branched and reach many parts of the brain. Among the areas they innervate are the cerebellum and the cerebral cortex, which are concerned with the fine coordination of movement, alertness and emotion. Another part of the brain innervated by noradrenaline neurons is the hypothalamus, which controls many visceral functions of the body such as hunger, thirst, temperature regulation, blood pressure, reproduction and behavior. Manipulation of the noradrenaline levels in the brain can change many of the functions of the hypothalamus, particularly the "pleasure" centers. Noradrenaline tracts also appear to be involved in mood elevation and depression. Recently nerves containing adrenaline have also been observed in the brain stem. The next few years should show whether these adrenaline-containing tracts also control emotion, mood and behavior.

Drugs have been powerful tools for probing the action of neurotransmitters. As our knowledge concerning neurotransmitters has broadened, so has our understanding of the action of drugs on behavior and on the cardiovascular and motor systems; the two trends have interacted nicely. In the early 1950's pharmacologists recognized that the hallucinogenic agent lysergic acid diethylamine (LSD) not only resembles serotonin in chemical structure but also counteracts some of its pharmacologic actions (by occupying sites intended for serotonin). Several workers therefore proposed that serotonin must have something to do with insanity. Other hallucinogenic agents such as mescaline and amphetamine, on the other hand, are related in structure to noradrenaline. In the mid-1950's clinical investigators were learning that chemicals such as chlorpromazine could mitigate psychotic behavior, and that monoamine oxidase inhibitors and imipramine and related drugs could relieve depression. At about the same time it was observed that reserpine, which was proving valuable not only for hypertension but also for schizophrenia, markedly reduced the levels of noradrenaline and serotonin in the brain. The observations combined to suggest that these drugs exerted their actions on the brain by interfering with neurotransmitters. When my colleagues and I found that radioactive noradrenaline can be taken up and released from nerves, we were in a good position to investigate how a drug influences the disposition of injected radioactive transmitters.

Effect of Drugs

The first compound we examined was cocaine, a potent stimulant that can produce psychosis and that also intensifies the action of noradrenaline. When radioactive noradrenaline was injected into cats that had been given cocaine, the uptake of catecholamines by the sympathetic nerves was prevented, demonstrating that cocaine magnifies the effect of noradrenaline by preventing its capture and inactivation and leaving larger amounts of the catecholamine to react with the effector cell. Antidepressant drugs such as imipramine had the same effect: they blocked the uptake of noradrenaline into sympathetic nerves. By using radioactive noradrenaline we found that amphetamine, which is both a stimulant and a mind-altering drug, affects noradrenergic nerves in two ways: it blocks the uptake of noradrenaline and also promotes the release of the neurotransmitter from nerves.

Many drugs that are effective in the treatment of hypertension affect the



CERTAIN ADRENERGIC DRUGS increase or decrease the availability of noradrenaline at the adrenergic receptor. Normal release, recapture and metabolism (*colored arrows*) are illustrated, with a curve representing the normal response of a postjunctional cell (*a*). Antidepressant drugs enlarge that response in several ways, all of which increase the availability of noradrenaline at the synapse. Amphetamine does so by promoting the release of noradrenaline (*b*). Amphetamine and imipramine and related drugs block recapture (*c*); the monoamine-oxidase inhibitors interfere with inactivation through metabolism (*d*). Conversely, reserpine, which reduces blood pressure and may induce depression, reduces the response by depleting the noradrenaline in storage (*e*); alpha-methyldopa and other "false transmitters" are stored in the vesicles with noradrenaline and released with it, diluting its effect (*f*).

Our dream began with six inconceivable concepts...

...but our problems began with this rough wooden model.

The ultimate photographic system. That was our dream for the SX-70 Land camera.

"The process," we said, "must be nonexistent for the photographer who, by definition, need think of the art in the *taking* and not the making of photographs."

When we first spoke of that dream at the Royal Photographic Society of Great Britain on May 21, 1949, it sounded quite reasonable. But it began to sound quite unattainable when we sat down and actually listed the desiderata for such a system. We ultimately arrived at six basic considerations:

First, simplicity of operation. The photographer would have nothing to think of, or do, but select his subject, focus and shoot. The system would then automatically read the light, set the shutter speed and lens opening, make the exposure, eject the film from the camera, develop and fix the image, and produce a finished, protected, mounted print.

Second, nothing to time. The film would develop itself by means of a process of some kind that would invariably stop when the print achieved its full color potential.

Third, nothing to throw away. The film would develop outside of the camera, even in broad daylight. But *without* the peel-apart light-tight format that had protected its predecessor.

Fourth, a wide range of functions. The ability to photograph subjects that were close up or distant, active or still, in daylight or darkness . . . automatically.

Fifth, a reasonable price.

And sixth, an unreasonable size (only 1″ x 4″ x 7″). We established that size with our model. It looked innocent enough: just a couple of blocks of wood and some bits of metal. But before we knew it, we had a tyrant in our midst. On the one hand, the *internal* space restrictions it imposed seemed impossible. On the other hand, its pocket-size *external* dimensions were unalterable.

The resulting dilemma ultimately forced us to reevaluate (and in many instances, reinvent) every element of the picture-making process.

For example, when the optical team realized that the model provided insufficient space to house the versatile taking lens that the camera required, they requested that the camera's thickness be increased. When that request was denied, the team had to conceive, and exploit, a whole new family of lens designs. Which led to the development of an extraordinary new optical system.

It is one thing to imagine a perfect photographic system, and another to solve the problems that stand in the way.

Many times we were told that our demands on current technology were unrealistic. But, as you will see in future advertisements in this magazine, *it was precisely because we adhered to those "unrealistic" demands, that the SX-70 was successfully developed*. And why it has given us the process that we call absolute one-step photography.

That term encompasses a system that provides all six of our requirements. And more.

Half of the system is a film that (among other things) develops instant color prints of striking beauty, right before your eyes. Like the picture below.

The other half of the system is a camera that (among other things) captures subjects from infinity down to 10.4 inches without a close-up lens attachment. That can take a sequence of up to 10 photographs at intervals of only 1.5 seconds. That weighs only 24 ounces. And that is almost identical to the pocket-size dimensions of a certain, *very* rough, wooden model.

The SX-70 System from Polaroid

Polaroid®

Conversation Pieces

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

How the Days Got Their Names — On Thursday, March 2, 1972 our Pioneer 10 spacecraft left for Jupiter, the first of the outermost planets. Although Pioneer travels so fast it swept past the Moon's orbit in a mere 11 hours, the voyage to distant Jupiter (a half a billion miles away) took two years.

Pioneer 10's departure took place on a peculiarly appropriate day. Thursday, it happens, is named after Jupiter. In fact, if we look back through astronomical history, we find that every day of the week is associated with an object in our solar system.

Early astronomers named the planets after gods and goddesses, and believed that each planet "ruled" or had primary influence on one day of the week. Jupiter, they held, ruled Thursday and so named the day Jove's day, or *jeudi* in the French*. Our Anglo-Saxon forebears replaced the Roman Jove with their equivalent deity, Thor. Hence we know it as Thor's day or Thursday.

THE NAMES OF THE DAYS COME FROM THE PTOLEMAIC SYSTEM

Here, for your information, is the complete planetary week. Women's lib advocates will be pleased to note that we should thank a goddess it's Friday.

Day	*Ruling Planet/ Divinity*	*Anglo-Saxon Equivalent*
Monday	Moon	—
Tuesday	Mars	Tiw
Wednesday	Mercury	Woden
Thursday	Jupiter	Thor
Friday	Venus	Freya
Saturday	Saturn	—
Sunday	Sun	—

*Those of you familiar with the French will see the planet's names clearly in *lundi, mardi, mercredi, jeudi, vendredi, and samedi.*

Burn Coal (But Not Throats)! Must a high standard of living and low quality of life always go hand-in-hand? The argument for the case is as follows. A high standard of living requires the consumption of large amounts of energy (e.g., lights, air conditioners, cars, home appliances). In producing and using this energy, however, we pollute our environment. If the air you breathe is toxic or the water you drink causes you to retch, be happy; your discomfort is proof positive you have a high standard of living.

To add to this dilemma, our so-called clean sources of energy are dwindling fast. A logical replacement is coal, the Earth's most abundant fossil fuel. Yet coal is a major polluter. When burned, it produces sulfur dioxide, a gas noxious to lungs, eyes, and throats. In 1970, for example, the U.S. pumped around 28 million tons of sulfur dioxide into the air.

How can we burn the coal and make the electricity and light the lights and run the air conditioners without befouling our atmosphere? At TRW, our answer is to remove the sulfur from the coal ***before*** burning it. The result: clean coal and a clean environment.

Until we came upon the method, it was considered formidable to remove the sulfur content. Strong acids have little or no effect on the sulfur, most of which is locked up tightly in the iron pyrites or fool's gold molecule. Strong oxidizers dissolve the pyrites but also oxidize the coal, making it useless. Our method removes the sulfur without altering the coal matrix, and increases the heat content of the coal by cutting down on the ash content. As an added attraction, our oxidizing agent can be regenerated and recycled.

Right now, we're happy to report, the Environmental Protection Agency is supporting the development of the process to determine its effectiveness and assess its economic merit. If it lives up to specs, we'll all breathe easier.

For further information, write on your company letterhead to:

TRW®
SYSTEMS GROUP

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

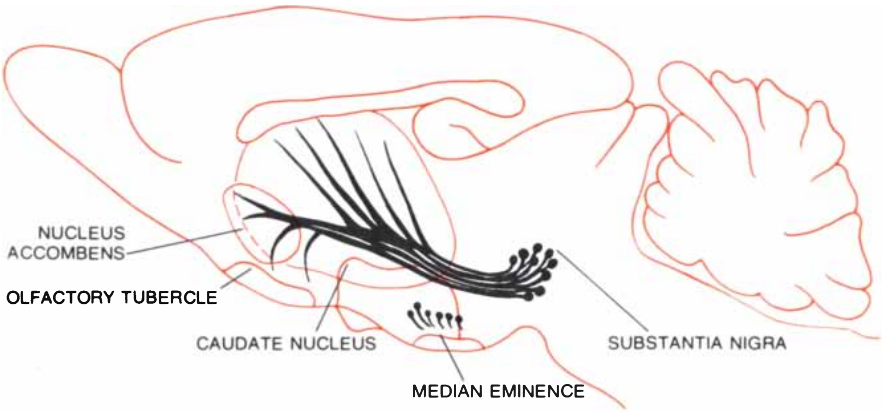
storage and release of noradrenergic transmitters. Reserpine and guanethidine reduce blood pressure by preventing the nerves that raise the pressure from storing noradrenaline. Antihypertensive drugs such as alpha methyl dopa, on the other hand, are transformed by enzymes in the nerve into substances that resemble the noradrenaline chemically. The "false transmitters" are stored and released along with natural neurotransmitters, diluting them and thus reducing their effect.

In the past 10 years many psychiatrists and pharmacologists have been struck by the fact that drugs that relieve mental depression also interfere with the uptake, storage, release or metabolism of noradrenaline. Whereas imipramine blocks the uptake of noradrenaline by nerves and amphetamine both releases noradrenaline and blocks its uptake, monoamine oxidase inhibitors, as their name implies, prevent the metabolism of the catecholamine. In other words, all these antidepressants produce similar results by different mechanisms: they increase the amount of catecholamine in the synaptic cleft, with the result that more transmitter is available to stimulate the receptor. Conversely, reserpine, a compound that decreases the amount of the chemical transmitters, sometimes produces depression. These considerations led to the proposal of a catecholamine hypothesis of depressive states, which holds that mental depression is associated with the decreased availability of brain catecholamine and is relieved by drugs that increase the amount of these transmitters at the adrenergic receptor. Although the hypothesis is not yet entirely substantiated, it has provided a valuable framework within which new approaches to understanding depression can be sought.

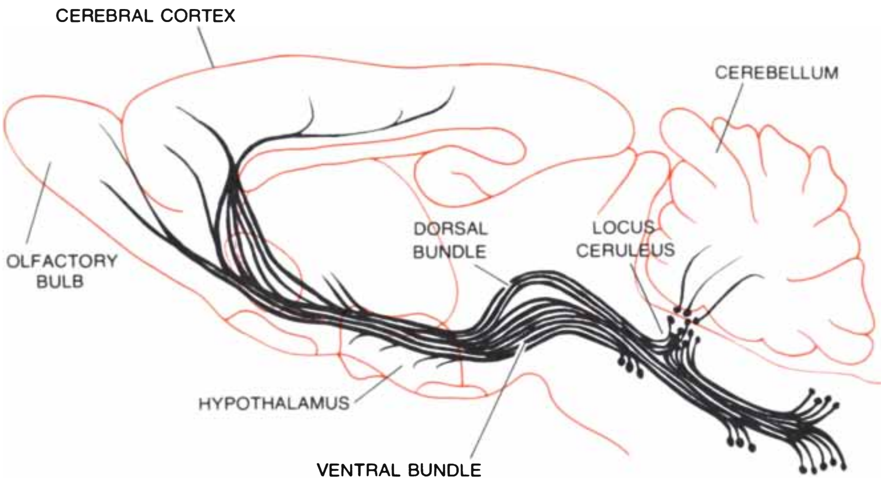
The introduction in the 1950's of antipsychotic drugs such as chlorpromazine and haloperidol revolutionized the treatment of schizophrenia, dramatically reducing the stay of schizophrenics in mental hospitals and saving many billions of dollars in hospital care. Research in the past decade has shown that antipsychotic drugs also exert their effect on the catecholamine neurotransmitters. Carlsson had observed that antischizophrenic drugs caused an increase in the formation of catecholamines in the rat brain, and he formed the hypothesis that this was owing to the drug's ability to block dopamine receptors. Work by other investigators has confirmed and extended this hypothesis. Antipsychotic drugs do block dopamine receptors in the brain, and there is a strong association between the blocking ability of various drugs and their capacity to relieve schizophrenic symptoms. These findings point clearly to the involvement of the dopaminergic nerves in schizophrenia.

Amphetamine has also helped to clarify the nature of schizophrenia. Taken repeatedly in large amounts, amphetamine produces a psychosis manifested by repetitive and compulsive behavior and hallucinogenic delusions that are indistinguishable from the symptoms exhibited by paranoid schizophrenics. Amphetamine releases catecholamines from nerves in the brain to stimulate both noradrenergic and dopamine receptors. After doing experiments with two forms of amphetamine Solomon H. Snyder of the Johns Hopkins University Medical School hypothesized that the schizophrenia-like psychosis the drug induces is due to excessive release of dopamine. The ability of antischizophrenia drugs, which block dopamine receptors, to relieve symptoms of amphetamine psychosis is consistent with this hypothesis.

Although there have been rapid advances in our knowledge of neurotransmitters in the past 20 years, much remains to be discovered about these compounds. Only a few of the chemical transmitters of the brain neurons have even been characterized. The role of neurotransmitters in behavior, mood, reproduction and learning and in diseases such as depression, schizophrenia, motor disorders and hypertension is beginning to evolve. If only the present trend toward reducing the funds committed to research support can be reversed, exciting new discoveries about neurotransmitters should soon be made, many of which will surely contribute directly toward the treatment or cure of some of man's most tragic afflictions.



DOPAMINE TRACTS, the main bundles of nerves containing dopamine, are shown (*black*) in a drawing of a longitudinal section along the midline of the rat brain. The cell bodies are concentrated in the substantia nigra, and the axons project primarily to the caudate nucleus. A dopamine deficiency in that region causes Parkinsonism, which can be treated with dopa.



NORADRENALINE TRACTS arise primarily in the locus ceruleus and reach many brain centers, including the cerebellum, cerebral cortex and hypothalamus. Illustrations on this page are based on maps made by Urban Ungerstedt of Royal Caroline Institute in Sweden.

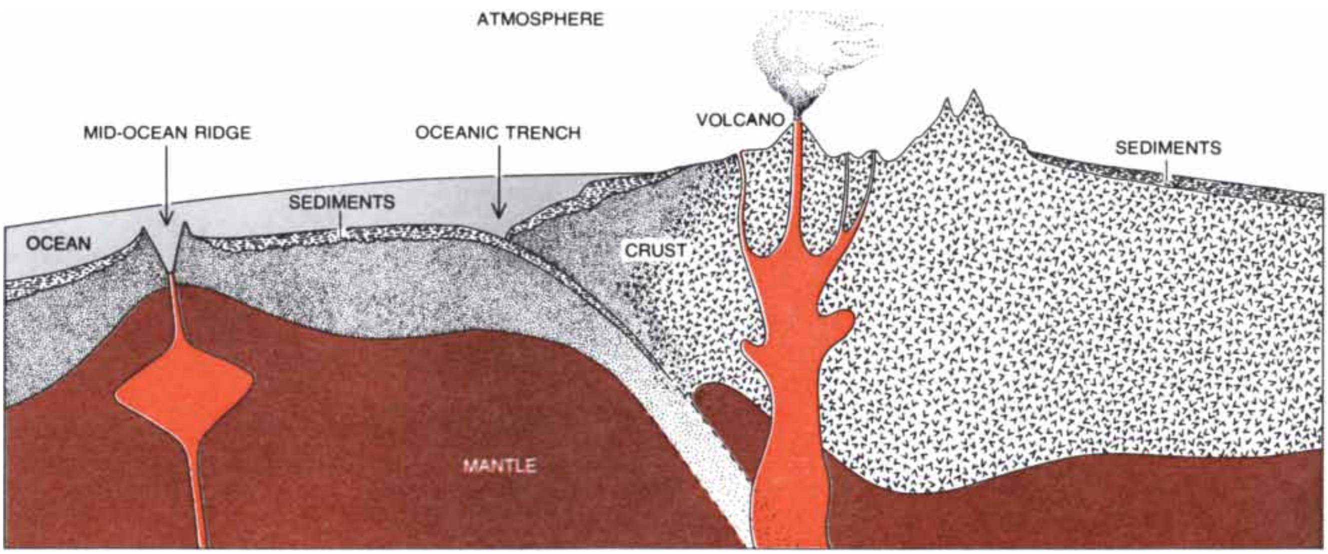
The Steady State of the Earth's Crust, Atmosphere and Oceans

The main divisions of the earth's surface are linked by vast geochemical cycles to form a stable worldwide system in which the inputs and outputs of the key constituents are in balance

by Raymond Siever

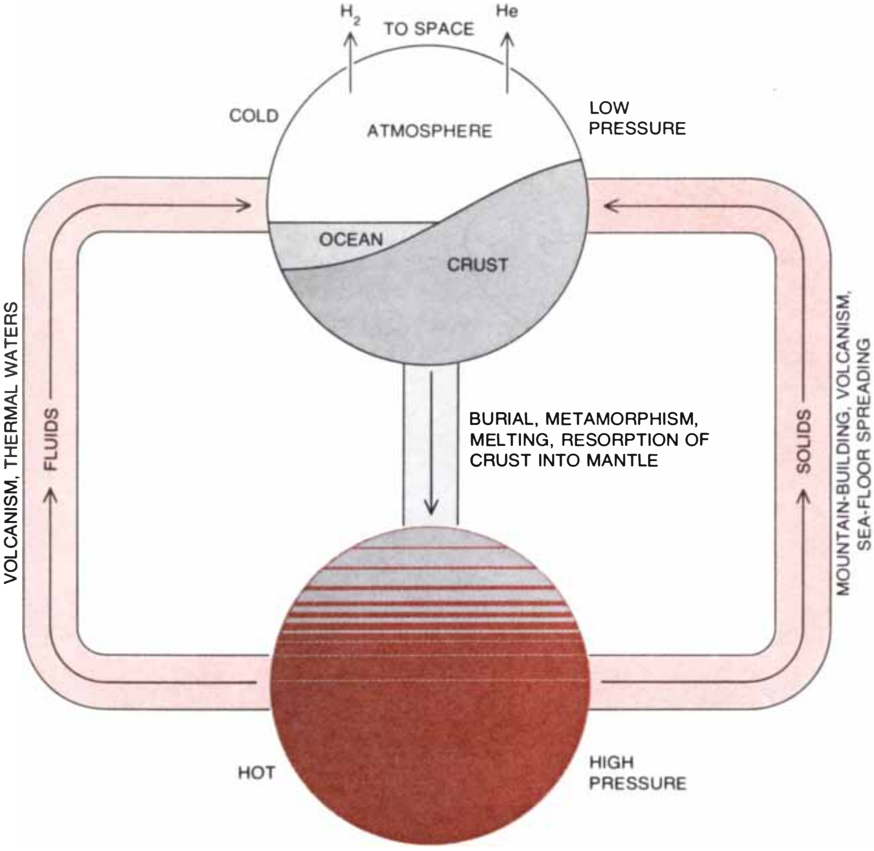
In trying to understand the operation of any large and complex natural system, such as the interaction of the solid surface of the earth with the atmosphere and the oceans, it sometimes pays to back off from the mass of detail and take a fresh look at the overall system. This kind of perspective does not come from looking at the earth from the moon (although comparisons with other planets are helpful) but rather from sitting here on the earth and sorting out the nature of the general processes at work on a worldwide scale from the specific mechanisms that operate at particular times and places in isolated components of the system. Perhaps the most useful way to make such an overall model is to compare the inputs and outputs of the major chemical components of the system as a whole and then to deduce the kinds of chemical reactions and physical processes that control the balance. This mode of analysis is described in modern jargon as the general-systems approach; more than a century before systems engineers were calling themselves that, however, geologists were using the same approach to deal with problems such as the origin of the ice ages.

In an analysis of this type much depends on how one defines the limits of the system and the states of matter in it. For the purposes of this article I shall include the atmosphere, the oceans and all other bodies of water on or near the solid surface of the earth, the rocks and soils at the surface, and the sedimentary, metamorphic and igneous rocks in the crust of the earth, together with the parts of the upper mantle that interact with the crust [*see illustration below*]. The fluid phases of this system include the gases of the atmosphere, the comparatively small amounts of gases



GEOPHYSICAL PROCESSES by which the crust, atmosphere and oceans are brought in contact with one another are summarized in this cross-sectional representation of the outer shell of the earth. The deep crust and the upper mantle are connected with the atmosphere and the oceans by the emanation of gases and the extrusion of lava from volcanoes and from the mid-ocean-ridge system. The rocks of the crust are recycled through the mantle at subduction zones underlying the oceanic-trench system. Sediments are recycled by being deeply buried and metamorphosed or melted in the deep crust or the upper mantle. Igneous, metamorphic and sedimentary rocks are uplifted in mountain-building and exposed to the atmosphere and hydrosphere, where they can be eroded and chemically weathered. The products of erosion and weathering are carried to oceans by the flow of ground water, streams and rivers.

trapped in the crust and the liquid water contained in the oceans, lakes and rivers of the world, together with the large amount of subsurface water found in buried sedimentary rocks. The chemical components of the system that I shall rely on for "tracers" include many of the most abundant elements: sodium (Na), potassium (K), calcium (Ca), magnesium (Mg) and chlorine (Cl), all normally present as ions; carbon (C), present either in the form of organic matter, principally formaldehyde (CH_2O), or as one of the carbonate species, that is, carbon dioxide (CO_2) and its close relatives carbonic acid (H_2CO_3), bicarbonate (HCO_3^-) and carbonate (CO_3^{--}); silicon (Si), chiefly in the form of silica (SiO_2), and hydrogen (H) and oxygen (O), most commonly in the form of water (H_2O), ionic hydrogen (H^+) and molecular oxygen (O_2). I shall ignore such important elements as sulfur, nitrogen and aluminum as well as a large number of less abundant elements, not because they do not tell important stories but because it is possible to illustrate the general nature of the system with what I have chosen.



STEADY-STATE MODEL of the closed-reaction system represented by the earth's crust, atmosphere and oceans is presented in this simplified diagram. The main components of the system are two chemical reactors, a leaky, hot, deep one and a cool surface one. The two reactors are linked by a variety of pathways that are controlled by differences in temperature and pressure. Only hydrogen and helium are light enough to escape the system.

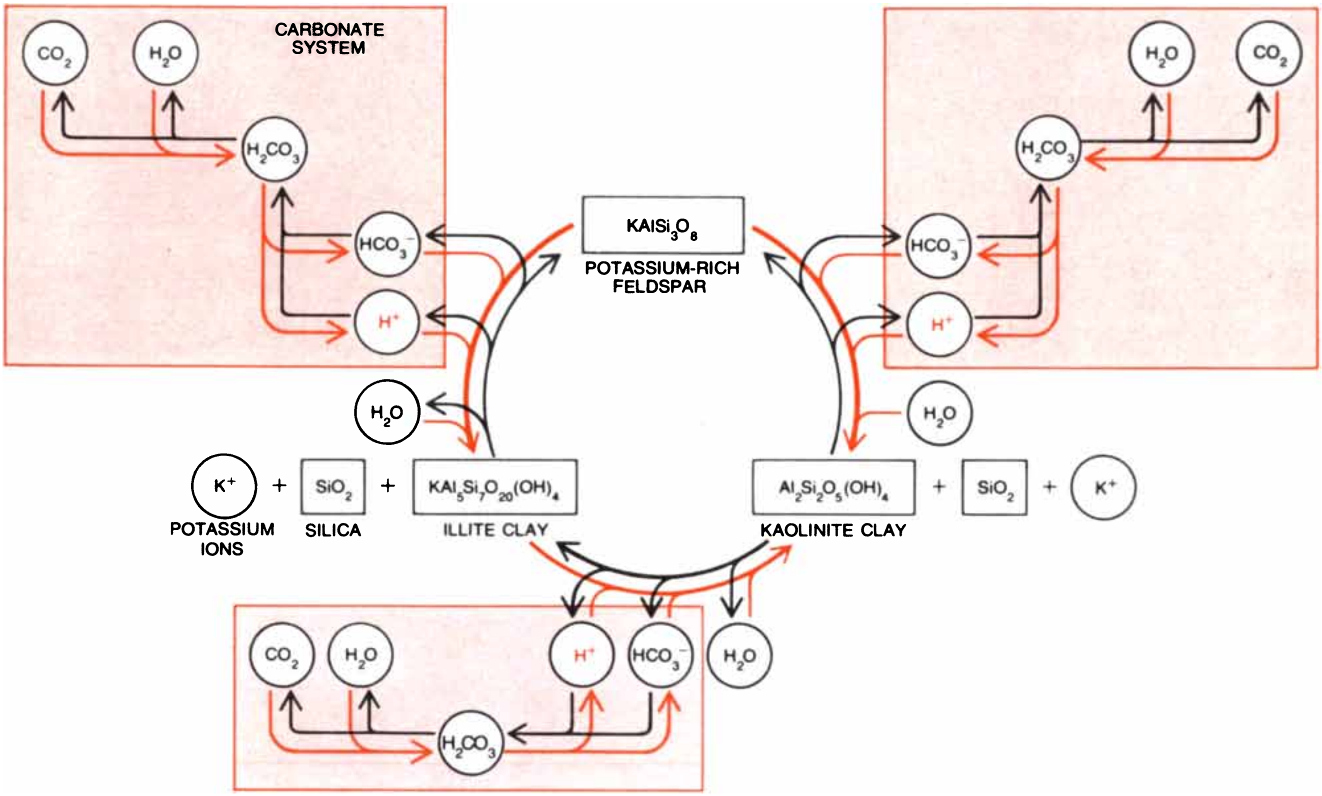
The analysis starts with an assumption of balance between inputs and outputs. The entire system represented by the earth's crust, atmosphere and oceans is further assumed to be generally stable now and to have been so over geologically long periods in the past. In short, this giant system is regarded as being in a steady state.

There is justification for the steady-state assumption in comparing the materials that enter and leave the world oceans every year. In 1952 Tom F. W. Barth of the University of Oslo made the steady-state assumption, estimated the influx of solids and dissolved material into the oceans and came up with the concept of the residence time of an element, which he defined as the total amount of the material present in the system (in his case the oceans) divided by the rate of either the influx or the efflux, the two being equal in the steady state. In 1958 Edward D. Goldberg and Gustaf O. S. Arrhenius of the Scripps Institution of Oceanography calculated residence times for the elements in the oceans from efflux data alone, that is, the rate at which such components leave the oceans in the form of sediment laid down on the ocean floor. They found a remarkable coincidence with Barth's residence times for many elements. This matching of rates of influx and efflux is a powerful argument for the assumption that the oceans are in a steady state at present and that they have been so in the recent past.

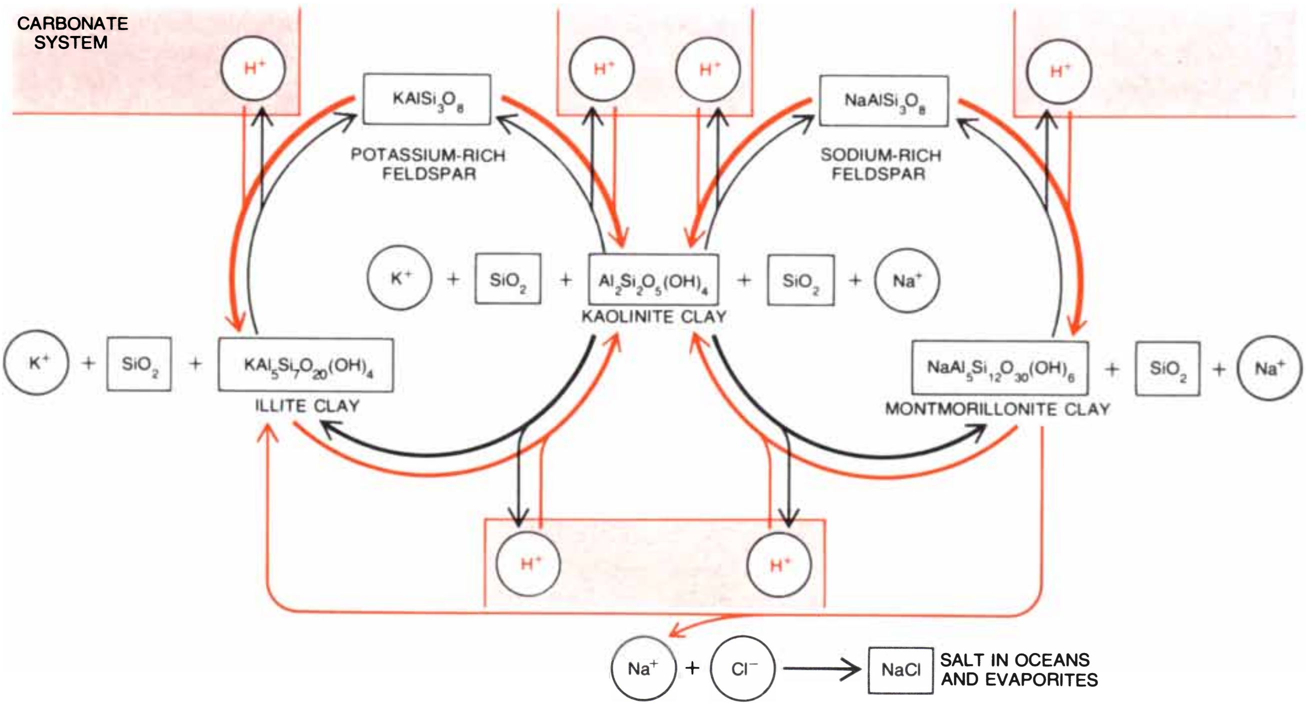
It is less easy to demonstrate in this way that the atmosphere is precisely balanced with respect to its components or that the solid surface of the earth is stable in its balance of mountain-building and erosion, but one is emboldened to do so by the geological record. Ancient rocks look just like those being made today; moreover, roughly the same relative ratios of rock types are characteristic of all large time segments of the world-averaged geological column. As a first approximation let us make the assumption that the entire outer shell of the earth has been stable for at least most of the well-known geological past; that covers the 600 million years or so since the beginning of the Cambrian period.

The primary input of solids into this system is through mountain-building, which exposes deeply buried rocks to the erosive attack of the atmosphere and the hydrosphere through the mechanism of weathering, the physical and chemical process by which rocks decay. Weathering as a chemical process is the sum of several reactions that result in the fixing of hydrogen ions, the removal of carbon dioxide and oxygen from the atmosphere, and the hydration, carbonation and oxidation of the reduced metals found in oxide and silicate minerals of igneous origin. A typical weathering reaction involves the weathering of a potassium-rich feldspar ($2KAlSi_3O_8$), a major constituent of many igneous rocks, to form kaolinite ($Al_2Si_2O_5(OH)_4$), an aluminum-rich clay abundant in many weathered soils: $2KAlSi_3O_8 + 2H^+ + 2HCO_3^- + H_2O \rightarrow Al_2Si_2O_5(OH)_4 + 2K^+ + 4SiO_2 + 2HCO_3^-$.

The hydrogen ions required for this reaction are supplied by the carbonate system, in which carbon dioxide (CO_2) in the atmosphere combines with rainwater to form carbonic acid (H_2CO_3), which then ionizes to form hydrogen ions (H^+) and bicarbonate ions (HCO_3^-). The result of the weathering reaction is to take the hydrogen away from the carbonic acid, thereby immobilizing the carbon dioxide as bicarbonate in solution and fixing the hydrogen and water in the solid kaolinite clay. At the same time the



CYCLICAL NATURE of feldspar weathering (*color*) and its opposite, silicate reconstitution (*black*), is evident in this illustration. After the feldspar, which is almost entirely of igneous origin, weathers to form various kinds of clay, the clays cycle through weathering and reconstitution with only a small return to feldspar. The silica left over after the amount involved in the clay-to-clay cycle is subtracted from the original amount derived from feldspar weathering appears in the geologic record as chert. The hydrogen ions required for the weathering reactions to go are supplied primarily by the carbonate system, in which carbon dioxide in the atmosphere combines with rainwater to form carbonic acid, which then ionizes to produce hydrogen and bicarbonate ions. The particular feldspar shown in this example of a typical weathering reaction is a potassium-aluminum silicate called orthoclase ($KAlSi_3O_8$).



PARTITIONING of potassium ions (K^+) and sodium ions (Na^+) during the weathering and reconstitution of silicate rocks goes through several stages. Originally potassium and sodium are well separated in the primary feldspars, but they are released to solution by weathering, eventually finding their way to the oceans. They can later become separated again in the course of silicate reconstitution, but because montmorillonite clay, which is rich in sodium, gradually (in about 100 million years) transforms to illite clay, which is rich in potassium, the potassium ends up stored in the clay whereas the sodium is dissolved in the oceans or is stored in the sedimentary salt deposits known as evaporites. Only the hydrogen-ion role of the carbonate system is emphasized here.

potassium ions (K^+) and the silica (SiO_2) become part of the dissolved load of the soil waters, eventually finding a path to the ocean by way of ground waters and rivers. Thus the unaltered silicate, feldspar, can be thought of as a hydrogen-acceptor and the free carbonic acid as a hydrogen donor.

The weathering of limestone accomplishes much the same chemical result as the weathering of silicates; hydrogen ions are used up in this reaction as well: $CaCO_3 + H^+ + HCO_3^- \rightarrow Ca^{++} + 2HCO_3^-$. In the limestone-weathering reaction the carbonate ion (CO_3^{--}) of the calcite mineral of the limestone is the hydrogen-acceptor. It now appears that most of the chemical weathering going on at the surface of the earth is the weathering of limestone, and that the amount of silicates chemically weathered is small in proportion.

A different kind of weathering reaction is the simple oxidation of a reduced metal, such as iron. Thus magnetite (Fe_3O_4) can be said to combine with oxygen to form hematite (Fe_2O_3), ignoring for the moment the complex chain of intermediate reactions by which this overall result is achieved. The amount of oxygen gas that is required to make this reaction go is orders of magnitude less than the amount of oxygen present in the atmosphere. The discrepancy not only accounts for the rapid oxidation of reduced metals in rocks when they are first exposed at the earth's surface but also suggests that the level of oxygen in the atmosphere is not at all controlled by the mineral equilibrium between reduced and oxidized states of metals such as iron.

The question of where the carbon dioxide in the atmosphere came from was considered by W. W. Rubey in 1951 in a now classic paper on the origin of seawater. Rubey showed that comparing the composition of igneous rocks with the sediments derived from them by weathering indicated a surplus of certain compounds in the sedimentary rock column, the oceans and the atmosphere. These substances—mainly carbon dioxide, hydrogen chloride (HCl) and sulfur dioxide (SO_2)—he called the "excess volatiles." The excess volatiles were produced by the outgassing of the interior of the earth through active volcanoes, hot springs and geysers. Most of the excess carbon dioxide that has accumulated as a result of outgassing has reacted with calcium dissolved from weathered igneous and metamorphic rocks and is now buried as calcium carbonate ($CaCO_3$) in limestone. The excess chloride has accumulated in the sea or has been buried in fossil salt beds.

Let us now take an overall look at the system and try to find an explanation for the driving force behind the chemical weathering reactions. The easiest way to do this is to consider a model in which the deep interior of the earth, the lower portions of the crust and the upper parts of the mantle are linked in a cycle that is controlled by differences in temperature and pressure [*see illustration on page 73*]. Rocks melt in places deep in the crust and in the upper parts of the mantle as a result of the high temperature and pressure there; in this sense the earth's interior resembles a "leaky" hot reactor. The leakage involves fluid phases, such as water, carbon dioxide and hydrogen chloride, and gaseous species of the elements sulfur, nitrogen and phosphorus—in short, all the excess volatiles that have appeared at the surface of the earth through geologic time. The leakage is vented to the surface, where it can be observed and measured in the vicinity of volcanoes and hot springs. In addition, there is ample evidence from the geochemistry of deeply buried sediments that there must also be a venting of some of these volatiles into the lower parts of deep metamorphic and sedimentary rocks; such venting may never reach the surface. The degassed solids that stay in the depths are stable at those high temperatures and pressures. Partly as a result of the loss of volatiles and partly as a result of cooling, these chemicals solidify to form the deep igneous rocks of the earth's crust. Eventually, through the agencies of sea-floor-spreading, mid-ocean-ridge formation and mountain-building, the solid products of the high-temperature reactor are brought to the surface, where they encounter much lower temperatures and pressures.

The fluid phases that escape to the surface stay in the atmosphere, condense or react chemically at the surface of the earth. The chemical reaction products—weathered clays and sediments—are held in storage there. Hydrogen is the only element light enough to leave the atmosphere of the earth quickly and be lost into space. Helium escapes too, but much more slowly than hydrogen. Except for these two elements (and an occasional spacecraft) one can consider the surface of the earth, including the atmosphere, as a closed system: a cold reactor into which are deposited the products of the hot reactor deep in the earth.

As the newly exposed igneous rocks come into contact with the cooled and condensed fluid phases, a low-temperature reverse reaction takes place between the reaction products of the high-temperature reaction. The dominant constituents of the fluids, mainly carbon dioxide and water, react with the silicates and oxides of igneous rocks at temperatures from 500 to 1,000 degrees Celsius lower than those at which they formed in the interior. As a result of the lower temperature and pressure, hydrated and carbonated phases form that are unstable at high temperatures and pressures.

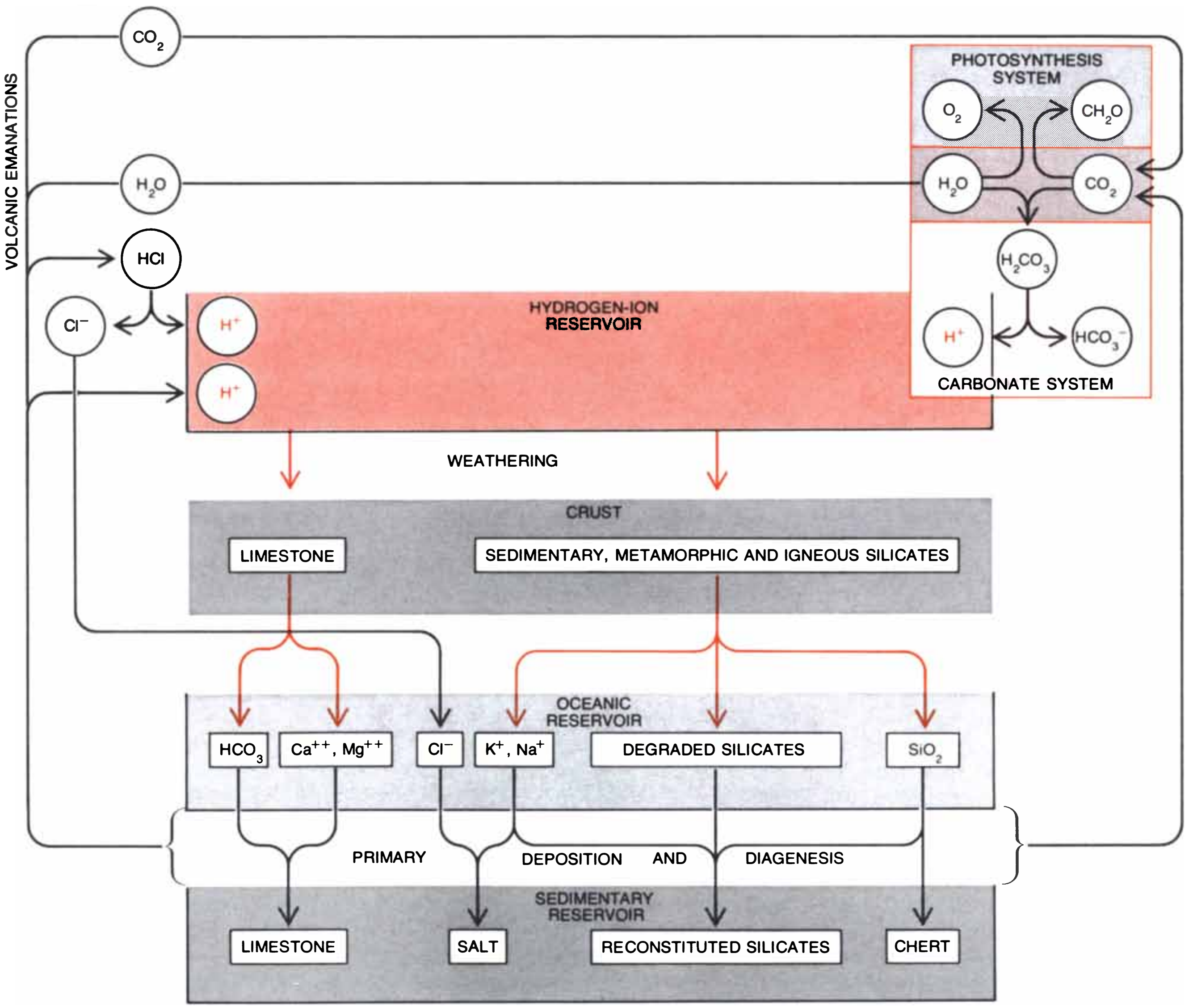
A simple mineral transformation that shows this reversibility is the process by which calcite, a form of calcium carbonate, reacts with quartz, a form of silica, at high temperatures during metamorphism to form the pyroxene mineral wollastonite ($CaSiO_3$) plus carbon dioxide. In the outcome carbon dioxide escapes to the surface. At the surface, when the wollastonite is exposed to the air containing carbon dioxide, the reaction goes in reverse and two typically sedimentary minerals are formed: calcite and quartz. Metamorphism and the resorption into the mantle of the edges of crustal plates in deep-sea trenches is simply the return of the solids to the hot reactor.

Thus the weathering process can be visualized as one side of a large closed system—the earth's mantle, crust, atmosphere and oceans—in which there is a temperature and pressure gradient. Whatever reacts at the hot end reacts in the reverse direction when it is moved to the cold end. A slightly different way of looking at the entire process was devised by the late L. G. Sillén of the Royal Institute of Technology in Stockholm, who described it as a giant acid-base neutralization reaction in which the escaping volatiles combine with water to form acids, which in turn react with the reduced metal silicates and oxides of igneous and metamorphic rocks.

Now let us focus on the balance between the ocean and the atmosphere with respect to the disposition of the products of weathering and erosion: the particles and dissolved matter brought from the continents to the sea. Weathering is reaction with carbon dioxide; the result is the removal of carbon dioxide from the atmosphere as a gas and the sequestering of it in dissolved form as bicarbonate ions in water. These ions are carried by rivers to the oceans, the negative bicarbonate ions being balanced largely by the positive ions of elements dissolved in the weathering of rocks. Now, however, in order to maintain the constraint of the original hypothesis—that everything is in dynamic balance and thus the amount of carbon dioxide in the atmosphere is constant—carbon dioxide must somehow be "exhaled" back into the atmosphere. Only a minor part of that steady return is from volcanic gases, because the turnover budget of carbon dioxide from weathering is very fast compared with the total amount accumulated over many billions of years from the outgassing of the mantle. The major part must be the return of dissolved bicarbonate ions to gaseous carbon dioxide. The only way that can be accomplished is to have the carbon dioxide combine with available hydrogen ions by some reaction so that it can again form carbonic acid and then carbon dioxide and water.

Where can one get the needed hydrogen? Some appreciation of the role that sedimentation may play comes from the same chemical reaction that describes the weathering of limestone. If that reaction is read in reverse, as the reaction of positive calcium ions with negative bicarbonate ions to form calcium carbonate, it describes the precipitation of limestone. The fact that much of this precipitation takes place biochemically in the production of the calcareous (calcium-containing) shells of marine invertebrates should not obscure the general nature of the reaction. The important aspect of the precipitation of calcium carbonate is that it releases hydrogen ions, which can then combine with the bicarbonate ions to form carbon dioxide and water.

Because we know that the weathering of limestone is the largest part of chemical weathering, and therefore the major immobilizer of carbon dioxide, it might be instructive to hypothesize the balance if limestone were the only rock that controlled the carbon dioxide of the atmosphere. In that event for every mole, or gram molecule, of carbon dioxide used up in weathering and hence removed from the atmosphere one mole of limestone would be dissolved. In the balancing precipitation of limestone a mole of carbon dioxide would be returned to the atmosphere for every mole of limestone deposited on the sea floor. As long as



PRINCIPAL GEOCHEMICAL CYCLES operating at present at the surface of the earth are summarized in this flow chart as they relate to three distinct chemical reservoirs: the hydrogen-ion reservoir, the oceanic reservoir and the sedimentary reservoir. The pathways of various "tracer" elements and compounds are traced through the stages of limestone and silicate weathering (made possible by an input from the hydrogen reservoir), dissolution in the oceanic reservoir and deposition (including precipitation and diagenesis) in the sedimentary reservoir, which releases hydrogen ions and carbon dioxide molecules for recycling. The all-important reservoir of hydrogen ions is maintained primarily by the carbonate system and secondarily by the hydrogen-chloride, or HCl, system.

limestone precipitation exactly balances limestone weathering there would be no change and the carbonate would simply be constantly recycled through the atmosphere and oceans. If the precipitation of limestone were to lag, however, a gradual reduction of carbon dioxide in the atmosphere would ensue.

One might think that this reduction would be counterbalanced a bit by carbon dioxide from the interior of the earth. The rate of outgassing is very slow, however, and could not operate over time intervals of less than hundreds of millions of years, whereas the residence time of carbonate ions in the ocean is much shorter, approximately 100,000 years. No, the system takes care of itself. Any reduction of limestone precipitation would cause a build-up in the concentration of calcium ions and carbonate ions in the ocean: the negative feedback mechanism would reside in the fact that the solubility of calcium carbonate would be so exceeded in the ocean that ultimately the rate of precipitation would start increasing and the balance would be reestablished.

Some hydrogen ions are available from the hydrogen chloride that comes out of the interior of the earth, but the amount that is available at any given time is small, as can be seen by comparing the amount of hydrogen chloride with the amount of carbon dioxide outgassed. All the chlorine that has appeared at the surface of the earth through geologic time, ultimately derived from outgassed hydrogen chloride, has either accumulated in the oceans or has been deposited in sedimentary rocks as salt beds or briny pore waters. Some of it has been recycled through the deep crust and mantle and so must be subtracted from the budget. The total amount of chloride is small compared with the total amount of carbon dioxide, however; because the residence time of chlorine is so much longer than that of carbon dioxide there cannot be enough of it at any given time to provide much acid.

Another way of restoring hydrogen ions to the carbonate system, thus allowing the return of gaseous carbon dioxide to the atmosphere, is to turn weathering reactions around and have them run in the opposite direction. This is analogous to the same process described above in which limestone precipitation is viewed as the reverse of limestone weathering. For the silicates, however, it is a more complicated process in which sodium, potassium, calcium and magnesium ions combine with silica and the products of clay minerals, such as kaolinite, to form new silicates that are richer in silica and poorer in hydrogen (and usually also poorer in water). For example, the reaction by which kaolinite is made into another clay, illite, gives off carbon dioxide.

Transformations of one clay mineral into another are well known as one part of the complex of postdepositional chemical and physical changes, called diagenesis, by which freshly deposited sediment ultimately becomes hard rock. One can view such clay reactions as silicate reconstitution, a process counterbalancing weathering, which is silicate destruction [*see top illustration on page 74*]. The cycle is normally incomplete; although feldspars, the most abundant minerals of the earth's crust, weather extensively to form clays, only a small fraction of the total clay reacts all the way back to reconstitute feldspar in sedimentary rocks. In the sedimentary part of the crust it is only the clay-to-clay transformations that are truly cyclic. During the metamorphism and recycling of sediments and metamorphic rocks through an igneous phase the cycle is completed to the feldspar stage. Hence the entire cycle that involves carbon dioxide is complete only when processes are included that are ordinarily thought to have no relation to the oceans and the atmosphere. One of the implications of the cycling for the sedimentary record is that for every mole of feldspar weathered and not balanced by diagenetic feldspar precipitating in sediments, 1.6 moles of silica are left over to find their way into the rock record as chert (flint) or other varieties of sedimentary silica.

The partition of sodium and potassium and their differing abundances in clays are also part of this general cycle, since these alkali metals are released in large amounts by the weathering of feldspars, their primary source in igneous rocks [*see bottom illustration on page 74*]. How does this subcycle operate? Although kaolinite is originally formed as a weathering product of the two most abundant feldspars, potassium-rich feldspar (called orthoclase and microcline) and sodium-rich feldspar (albite), reconstituted clays incorporate the two elements differentially. Illite gets the potassium, and montmorillonite the sodium. In this respect the earth does almost as well as an analytical chemist in separating the two elements. Ultimately the sodium finds a different home in the oceans and in the salt beds called evaporites; the evidence from the rock record is that montmorillonite is gradually transformed to illite over a period of about 100 million years. Both sodium and potassium reconstitute into feldspars during metamorphism or igneous melting. Thus for a mass balance to operate so that igneous rocks keep their average sodium-to-potassium ratios it is clear that in this process there has to be some appropriate mixture of the sedimentary part of the crust, which includes clays, evaporites, some saline pore waters and seawater.

If silicate reconstitution is an important part of the "buffering" process by which the hydrogen-ion concentration of the atmosphere and the ocean is maintained with respect to carbon dioxide, silica, alkali metals and alkaline earths, one ought to be able to find out just where and when such transformations take place. Where do weathered clays have their first opportunity to get into the appropriate geochemical environment? That place ought to be one that has abundant sodium, potassium, calcium, magnesium and silica in an alkaline environment that would soak up hydrogen ions as they are produced by reconstitution reactions. The mixing of river waters and their suspended weathered clays with seawater in estuaries and deltas would at first seem to be the best guess. The concentration of dissolved silica would decrease if reconstitution took place and so can be used for a tracer of the reaction. Work done in most such environments, however, shows that there is little if any silicate reconstitution.

Proceeding on the assumption that there may be too little time for such a sluggish reaction to take place in rapid transit through an estuary, one might then look at the bottom sediment of the oceans. Here one finds what appear to be the right conditions and more time. Some years ago I took a series of oceanographic cruises during which I collected sediment cores and squeezed out their interstitial pore waters just so that I might be able to find evidence of reconstitution reactions from changes in the composition of the waters in contact with the clays. I found that examination of both the waters and the mineralogy of the sediments failed to show any extensive reactions of the kind postulated. In fact, it was apparent that the opposite happened to a small degree, that is, mechanically eroded particles of feldspar transported with the dominantly clayey sediment continued to "weather" on the sea floor, enriching the pore waters with potassium. The shells of silica-secreting organisms (diatoms and radiolarians) appeared to be dissolving to give high concentrations of dissolved silica instead of the concentration's being reduced through precipitation or incorporation with clays in reconstitution reactions. I have been able to discover by electron microscopy and X-ray diffraction (as have others who have worked in this field) that there are tiny amounts of reconstituted silicates to be found in marine sediments, but those amounts seem to be associated with areas where undersea volcanism has played an important role.

Disappointment turned me back to my experience with ancient sedimentary rocks found on the continents, where mineralogical studies of sandstones, shales and limestones have made it abundantly clear that silicate transformations have in fact taken place. Large amounts of the potassium-rich clays (such as illite) and the magnesium-rich clays (such as chlorite), diagenetic feldspars and free silica in the form of chert and diagenetic quartz are characteristic of sediments now found in deep sedimentary basins. There the passage of time, deep burial and reaction with pore waters of varied composition have led to the formation of diagenetic assemblages of minerals. As sediments become buried in such basins they gradually move into regions of higher temperature and pressure that accelerate the reactions in the direction of reconstitution. This tendency implies that there may be a steady, but extraordinarily slow, exhalation of carbon dioxide diffusing upward from the continents as a result of the diagenesis of sedimentary rocks and their associated pore waters.

If one looks carefully for the sites of maximum diagenesis, in terms of the total quantity of rocks involved as well as the intensity of the process, it becomes apparent that the continental margins are the site of the thickest accumulations of sediments, both in marginal geosynclines and in the continental shelves and rises. It is in these same belts that major regional episodes of mountain-building and rock metamorphism have taken place. Here too are the sites of subduction zones, where thick accumulations of oceanic sediment are carried by downward-moving lithospheric plates, crumpled and scraped off at trench walls. Diagenetic and metamorphic chemical reactions at continental margins supply the hydrogen ions that are ultimately responsible for the continued buffering of the oceans and the atmosphere.

Evidence is now accumulating from the deep-sea drilling program that there must be deep-water circulation in the continental shelves and probably the rises as well. A general view of groundwater movement in the sediments of coastal plains is that the formation waters diffuse into the oceans over a long time scale. Thus the hydrogen ions released from both silicate reconstitution and carbonate precipitation find their way into the oceans to buffer them. Comparatively shallow sedimentary accumulations on the continents, far from the oceans, do not connect with the system as a whole except over very long time scales; in addition they are volumetrically unimportant in buffering the large system.

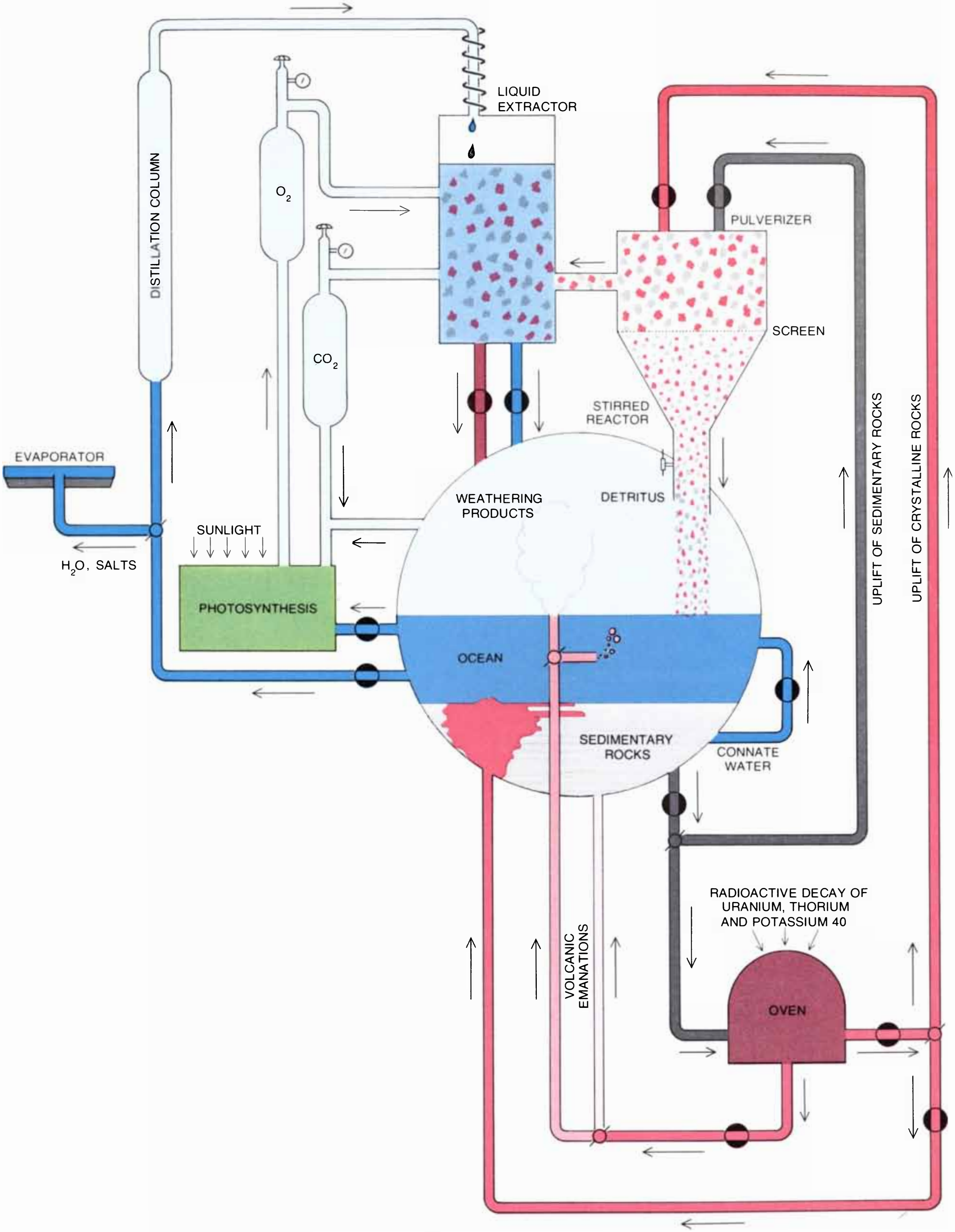
The entire process can be summed up in a flow chart of weathering, sedimentation and reconstitution with respect to the hydrogen-ion reservoir, the oceanic reservoir and the sediment reservoir [*see illustration on page 76*]. Still another way of showing this system is by comparing the earth's geochemical system to a giant chemical-engineering plant, a concept suggested at a research conference some years ago by Kenneth S. Deffeyes, now of Princeton University. The provocative features of such a hypothetical plant are the multiplicity of valves and switches that control the flow of the system and the speed with which the system can respond to some kind of perturbation, oscillation or unsteady imbalance if there are violent movements of any of the switches or valves [*see illustration on opposite page*]. The gas regulators of the carbon dioxide and oxygen tanks are of special importance to man, indeed to all respiratory life.

The oxygen regulator is governed dominantly by photosynthesis and only secondarily by weathering. If an earth empty of green plants were left to itself for a geologically long time, the oxygen of the atmosphere would be gradually used up in oxidizing the reduced metals brought to the surface by mountain-building and igneous activity. The carbon dioxide regulator is probably much more controlled by the weathering system than it is by photosynthesis. Both the diagram and the flow chart are of course incomplete, since I have omitted the important subcycles of nitrogen and phosphorus, two elements that are strongly related to the origin of life and to biological processes in general. Even though the details differ, however, the general pattern of operation is the same for all the reactive elements.

This analysis leads one to look for possible different states of the system in the past by examining the kinds and proportions of sediments in the geologic record as compared with those of the present. Where today the limestone subcycle of weathering and precipitation is far more important than the silicate system in buffering carbon dioxide, this may not have been so during all of geologic time. Because both limestone precipitation and silicate reconstitution supply hydrogen ions, their sum should be constant in order to keep the atmosphere and the oceans steady in carbon dioxide. Hence past times where there was a worldwide decrease in limestone precipitation and weathering should have produced rock accumulations with more reconstituted silicates in proportion to limestone than modern or recent deposits do.

Another consequence of the steady state is that salt-bed deposition and silicate reconstitution should have a constant sum for the alkali-metal content of the oceans to remain constant. There are many other balances that reflect possible changes in past operation of the system; all depend on a detailed analysis of the world distribution of rock types for a given age. Such analyses have been undertaken only in the past few years, most notably by A. B. Ronov and his co-workers in the Institute of Geochemistry and Analytical Chemistry in Moscow.

The earth's past is the business of geologists, but the future is the business of all of us. In order to know how the tremendous injection of carbon dioxide into the atmosphere in the past century from the burning of fossil fuels will affect the entire system, we shall need to know how rapidly the system responds and what the immediate consequences will be. Although the reserves of oxygen in the atmosphere are so huge that we need not fear for the next few millenniums, the balance of photosynthetic oxygen production and oxygen demand by organisms, decaying organic matter and reduced metals may be displaced detectably by man's activities. Photosynthetic oxygen production may be affected somewhat, perhaps by pollution influencing the productivity of algae in the sea or by land developers cutting down extensive tropical forests such as those of Brazil. How important will the rate of oxygen use by reduced metals in rocks be when rates of erosion are drastically increased by man's landscaping activities? What is the rate of oxygen use by all the rusting metal of the world's scrapheaps? Until we understand more fully the workings of this large system in a quantitative way we shall not know how much we are tampering with it.



PROVOCATIVE ANALOGY portrays the surface geochemical system of the earth as a giant chemical-engineering plant. The main source of energy for the plant is the radiant energy from the sun; another energy source is the decay of radioactive elements in the "oven" representing the interior of the earth. A striking feature of such a hypothetical plant is large number of valves and switches.

Opel. It doesn't need a lot of gas to be exciting.

The Germans are a demanding people. They expect perfection of most things—themselves, their work, their products and, of course, the automobiles they drive. And more often than not, they get it. Or at least come close.

From the country that gave you the Gullwing and the economy import

It is not surprising, therefore, that some of the fastest, most powerful cars in history have come from Germany. It is also not surprising that some of the most frugal, gas-pinching cars are likewise German in origin.

The optimum compromise: economy *and* spirit

It is equally logical that the best selling car in such a nation must embody a combination of both. Performance *and* economy. Emotion *and* frugality Art *and* Science.

The Opel Manta Rallye is such a blend. Its precision 1.9 litre, cam-in-head 4-cylinder engine and 4-speed transmission allow it to cope handily with the rigors of the German Autobahn. And for urban situations, there's an automatic transmission available. But in either case, the powerplant is fundamentally rugged and simple Hence, maintenance and fuel consumption are held to practical minimums.

Old World practicality

Virtually everywhere you look, Manta Rallye is an exercise in sound, no-nonsense thinking. The hood is painted a businesslike flat-black, for efficiency as well as looks. The instrumentation is as sensuous to the eye as it is plentiful to watch. A large tachometer checks engine speed, and there are warning lights and gauges for amps and oil pressure. The front seats, in best German fashion, recline individually at the twist of a knob. And the steering wheel carries a skin of soft vinyl.

It works

The total effect is that of function. The car looks like it works. And it does work. It works because of things like front and rear stabilizer bars in the suspension and an efficient forced-air ventilation system. It works because of power front disc brakes and dozens of other things, little and big, that help make it a genuine pleasure to drive. It works because, well, it *has* to, to be the best selling car in Germany

And the price?

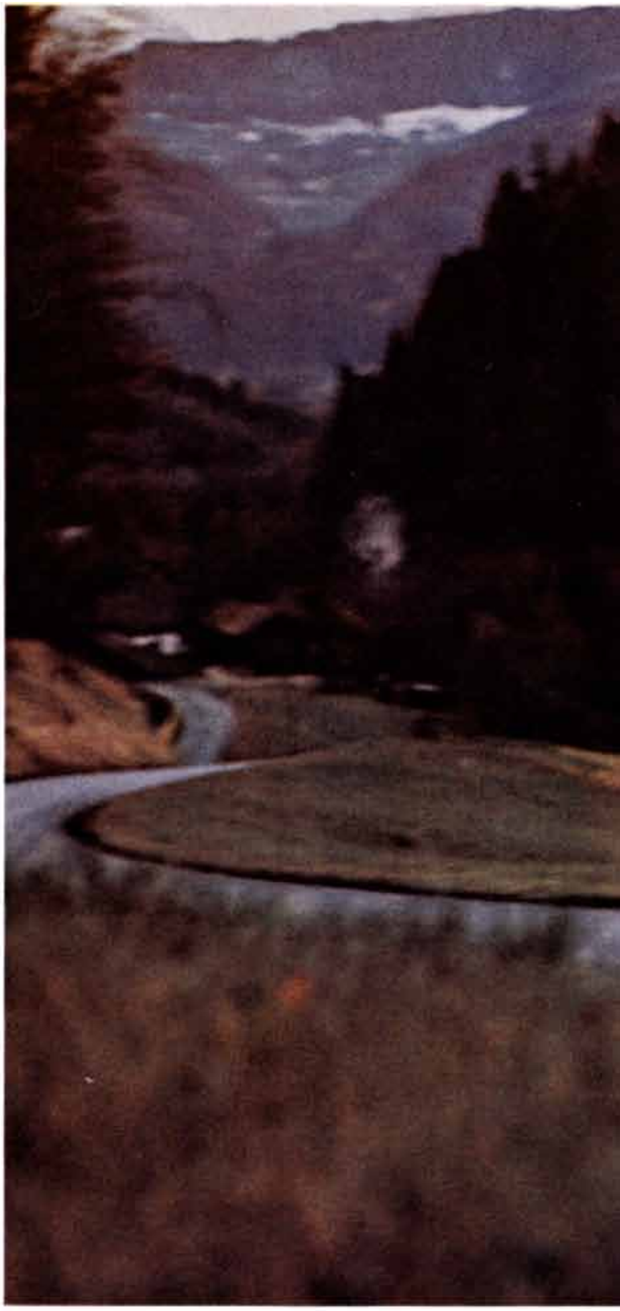
The price of an Opel Manta Rallye in America is $3,584.22,* including road wheels, special instrumentation, and radio.

The standard Manta Coupe is priced at $3,298.50,* the Manta Luxus and Opel Sportwagon at $3,535.50* and $3,521.50,* respectively

See and drive one at any of more than 2,200 Buick/Opel dealers throughout America.

**Manufacturer's Suggested Retail Price, including dealer new vehicle preparation charges; other optional equipment, state and local taxes and transportation charges from Port of Entry are additional.*

Opel. The best selling car in Germany.



ROTATING CHEMICAL REACTIONS

Certain colorful reactions visibly organize themselves in space and time as a spiral wave. They suggest that many other chemical reactions have a geometrical aspect that needs to be investigated

by Arthur T. Winfree

Most of us have grown up thinking of chemical reactions as things that happen to a combination of stuffs in a test tube, thereby changing them in a fairly direct and uniform way to some other combination of stuffs. We never saw test tubes flashing different colors like a neon sign or differentiating into interdependent regions pursuing different reactions. Of course, no one doubts that such things can happen in chemical systems. After all, the world is made of chemicals. Richly structured reactions are as common as flames or growing plants or frog eggs in a pond. Such complicated and interesting reactions, however, usually seem to involve the transfer of heat across surfaces, or physical movement, or membranes separating diverse materials, or electric currents flowing from one region to another. Every schoolchild knows that reactions of ordinary chemicals in a homogeneous solution is a rather dull business, favored at best by a single uniform change in color.

It was presumably because of such impressions that A. M. Zhabotinsky and his colleagues at the Institute of Biological Physics near Moscow puzzled over a reaction discovered by B. P. Belousov in 1958. In Belousov's reagent, ions of the metal cerium catalyze the oxidation of an organic fuel by bromate in water. Unlike the overwhelming majority of familiar reactions, this particular one has no stable steady state: it oscillates with clock-like precision, changing from yellow to colorless and back to yellow again twice a minute.

Although Zhabotinsky's experiments were published only in Russian during the 1960's, they aroused considerable interest among Western biophysicists and biochemists studying rhythmic phenomena in living organisms. I met Zhabotinsky at a gathering of such people in Prague a few weeks before the Russian occupation of Czechoslovakia in the summer of 1968. I toyed with the reaction long enough to improve my own ideas of how to investigate the 24-hour "clock" oscillation in the brain of the fruit fly, but in the excitement of biological discoveries I soon forgot the chemicals behind a shelf of culture media.

Then in 1970, in a short article in *Nature,* Zhabotinsky and A. N. Zaikin reported that they had observed circular chemical waves propagating through a modified version of the oscillating reagent. These vivid blue waves emerge periodically, forming concentric rings like a bull's-eye target pattern around isolated "pacemaker" points in a shallow dish of the reddish-orange liquid. Where waves from two pacemakers collide they do not interpenetrate like ripples in a pond; they vanish like colliding grass fires. In successive collisions waves from the higher-frequency pacemaker encroach step by step closer to the lower-frequency pacemaker. Eventually they arrive at the slower pacemaker and entrain it, reducing by one the number of target patterns in the dish.

What is the pacemaker source of the target pattern? Is there something unusual at those points, or do concentric rings arise purely by diffusion and the kinetics of chemical reactions in pure reagent? Pacemakers have diverse periods, suggesting that they arise from diverse causes. They commonly appear on interfaces between the liquid and the air or between the liquid and its container. Often a floating dust mote or a scratch on the glass floor of the Petri dish can be seen in the microscope. If the liquid is in contact with a strong oxidant or certain metallic alloys, at that point the reddish-orange reagent turns blue at regular intervals, sending out ring-shaped waves like a pacemaker, suggesting that local chemical impurities in the solution can be pacemakers. One can also introduce pacemakers by the time-honored technique of crystallographers: ruffling one's beard over an open dish, or sprinkling in a pinch of laboratory floor dust.

Pacemakers can be eliminated by carefully filtering the reagent into a scratch-free dish lined with silicone. Reagent so prepared and left undisturbed continues to oscillate, but it does so almost homogeneously, in bulk. Is such a filtered reagent still able to conduct waves? It is indeed: if a droplet of reagent from the blue part of a wave in another dish is added to the homogeneous solution, it "infects" the filtered reagent, and the infection propagates as a single ring expanding at the standard speed. Thus the absence of target patterns is due to a lack of pacemaker sources for circular

SPIRALS OF CHEMICAL ACTIVITY form in a shallow dish of red "Z reagent." A blue ring was induced by touching the surface of the solution with a hot filament, and then the dish was rocked gently to break the ring. The free ends of the fragmented circular wave curl around a pivot near each end point, winding up into spirals that have a uniform spacing between waves. Wherever two waves collide head on, both vanish. Each entire spiral rotates with a period (τ_0) of about one minute. Sequence of photographs reads from left to right; pictures were taken at times of 0, ½, 1, 1½, 2, 3½, 4½, 5½, 6½, 7, 7½ and 8 minutes. Dishes shown are about actual size. Bubbles are carbon dioxide, a product of the chemical reaction. Depth of the liquid in all photographs on this page and the next is 1.4 millimeters.

"BULL'S-EYE" TARGET PATTERNS are formed by rings of chemical activity in the Z reagent. They are emitted at diverse but regular intervals by pacemaker points, which seem to be particulate impurities. Rings from faster pacemakers gradually encroach on slower pacemakers until the slower sources are taken over and entrained. Photographs are at intervals of half a minute.

waves. It seems that the pacemakers are removable particulate impurities.

It also turns out that by having a little less acid and a little more bromide in the same recipe, the solution's bulk oscillation can be eliminated without affecting its ability to conduct waves of chemical activity. What remains is a stable reddish-orange solution that I shall call Z reagent in honor of Zhabotinsky and Zaikin. Z reagent does nothing when it is left alone, but when it is stimulated by an "infectious" droplet of a blue wave from another dish or by the touch of a heated needle, a single sharp blue ring propagates through it at a steady rate of a few millimeters per minute.

The details of the chemical mechanism and kinetics in the Z reagent are only now being unraveled. Elaborating on the earlier studies by Zhabotinsky, Zaikin and V. A. Vavilin, Richard J. Field and Richard M. Noyes of the University of Oregon and E. Körös in Budapest have recently proposed a scheme for Belousov's reagent involving reactions among 11 substances. Field and Noyes have further shown that their scheme can be approximated by a sequence of reactions involving only three substances that would spontaneously oscillate.

Their sequence of reactions goes something like this. The important basic constituents of the solution are bromide and bromate (two forms of bromine compounds), malonate (an organic fuel) and the indicator dye phenanthroline (containing iron). The dye is red when its iron atom is reduced to the ferrous form with two electrons in its outer electron shell, and it is blue when it is oxidized to the ferric form with three electrons in its outer electron shell. The concentration of bromide ions in the solution determines which of two sets of reactions will dominate. The first set of reactions is associated with a high concentration of bromide, with reduced (ferrous) iron; the solution is therefore red. The second set of reactions is inhibited by bromide and is associated with oxidized (ferric) iron; the solution is therefore blue.

When the concentration of bromide is high, it is used up in the first set of reactions to brominate, or add bromine to, the malonate to form bromomalonate. As the concentration of bromide falls to a threshold level, the second set of reactions starts to dominate and the last vestiges of bromide are consumed. Meanwhile the bromate takes over the bromination of the malonate; simultaneously it oxidizes the iron atom in the indicator dye, changing it from the ferrous state to the ferric, so that the phenanthroline complex is transformed from red to blue. The ferric phenanthroline then oxidizes the accumulated bromomalonate. This reaction releases bubbles of carbon dioxide and so much bromide that the second reaction sequence shuts off and the first begins anew [*see top illustration on page 86*].

Similar reactions presumably govern the excitability of the nonoscillating Z reagent made with extra bromide and less acid. Here the red and blue reactions will balance with the concentration of bromide just above the threshold: red dominates until some external disturbance intervenes. Following such a disturbance the blue frontier propagates by consuming the slight amount of bromide diffusing from the adjacent red liquid, thus allowing it likewise to switch to the blue stage. Once switched, the reactions revert on their own to the red stage and await the next triggering disturbance.

Since many laboratories here and abroad are taking a vigorous interest in these reactions, it seems likely that the chemical processes will soon be exposed in complete quantitative detail, making it possible to custom-tailor more exotic variants. Nothing more exotic is needed, however, if all one wants is strange phenomena to play with while thinking about spatially self-organizing reactions. The phenomenon that intrigued me is the creation of spiral waves of chemical activity in the Z reagent.

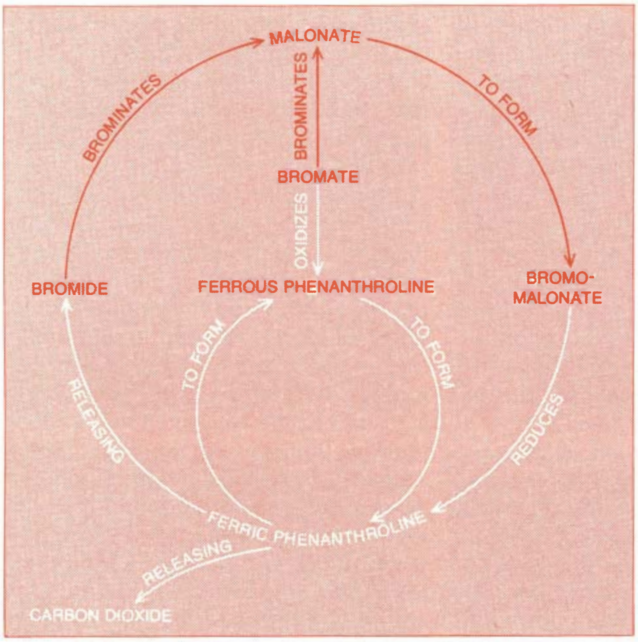
Segments of a circular wave front propagating through the reagent can be made to vanish by gently shearing the liquid, say by slightly rocking the dish. The remaining fragmented waves begin to curl up around a pivot near each end, eventually winding into spirals having a uniform spacing of about two millimeters between waves. The entire pattern rotates once every 20 to 60 seconds, depending on the recipe [*see illustration on page 83*].

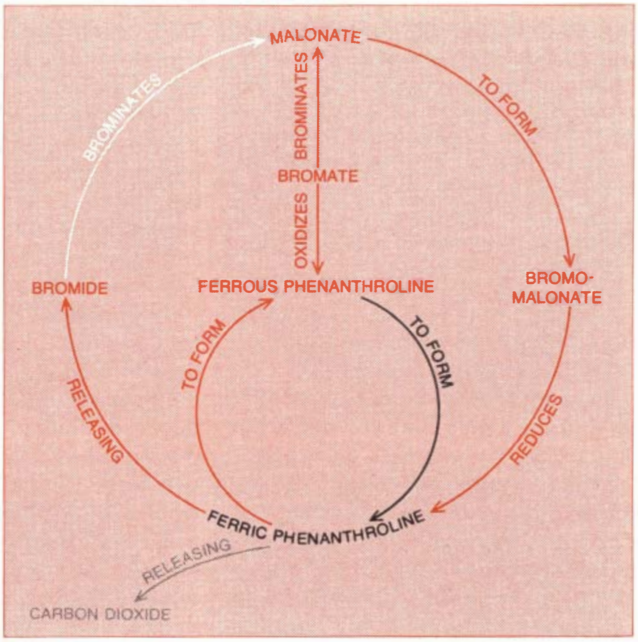
A note from Zhabotinsky revealed that he too had encountered spirals. In fact, Russian mathematicians and biophysicists trying to understand flutter and fibrillation in the heart had studied spiral waves in mathematically idealized excitable media years before the chemical medium was discovered. The Russian workers had become interested in the problem through a paper in the 1946 *Archives of the Mexican Institute of Cardiology* by the mathematician Norbert Wiener and the cardiologist Arturo Rosenblueth. They examined a model of electrically excited waves circulating around obstacles in heart muscle. Their early analysis was elaborated both mathematically and biologically in the 1960's, largely through the emphasis of I. S. Balakovskii and V. I. Krinskii on spiral waves. Could these analyses provide insights into the dynamical organization of spiral waves of chemical reaction in a solution?

In some ways they could. This line of analysis, however, is built on the notion that the state of an excitable cell changes in only one important way: it can be anywhere in a continuum of states varying from "rest" through "excited" to "refractory" and back to "rest." That may be a fair description of biological cells, and perhaps even of the propagation of waves in the Z reagent. By its inherent restriction to a single variable, however, it obstructs inquiry into the interplay of diffusion and reactions that actually generates waves in the rotating middle of a spiral. Any model relying only on a single variable requires that there must be an infinitely steep concentration gradient in the middle of a rotating wave. Therefore we must have a model that incorporates two variables. I shall return to this point in greater detail.

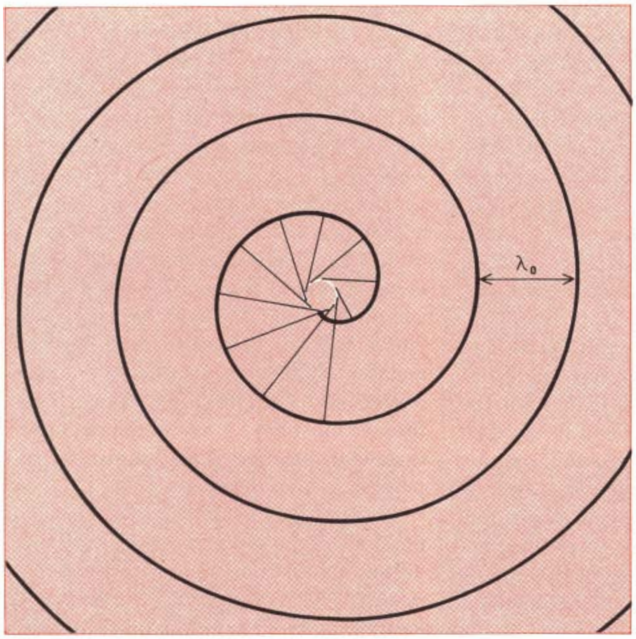
For the sake of simplicity, let us begin not with the initiation of a spiral but with a spiral that is fully mature and symmetrical. If one's attention is centered on a small section of the spiral, one could describe the wave near that section by saying: (*a*) "Each segment of wave front propagates perpendicular to its length at nearly the same velocity, v_0." A "global," or overall, description of the spiral could be: (*b*) "The wave rotates rigidly with period τ_0, as though it were engraved on a spinning phonograph rec-

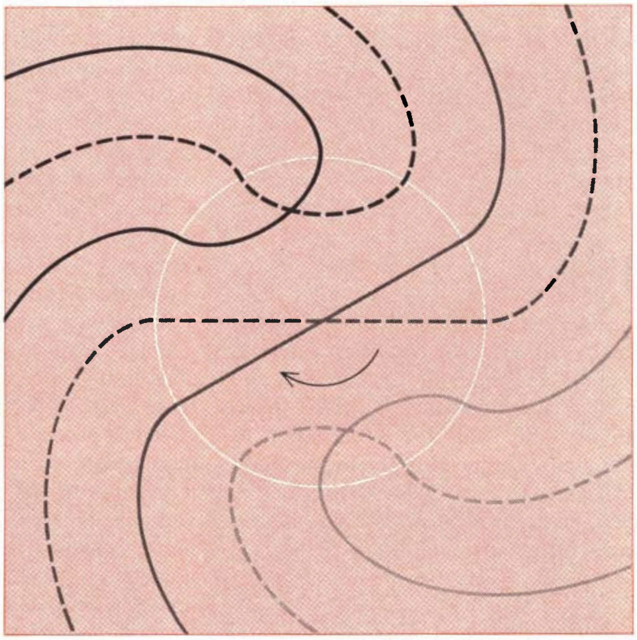
ELONGATED RINGS DECAY into a pair of symmetrical involute spirals at about the 12 o'clock position in the dish. At five o'clock an elongated ring source shortens toward perfect symmetry. Just before the sources at nine o'clock, 12 o'clock and two o'clock closer to the center would have emitted circular rings, they vanished. Photographs were made at intervals of one minute.





TWO SETS OF REACTIONS can account for the oscillation of the *Z* reagent from red to blue and back to red. The mechanism can be followed in skeleton form by describing reactions among bromide, bromate, malonate and iron phenanthroline, which serves double duty as a catalyst and an indicator dye. The concentration of bromide determines which of the two sets of reactions will dominate in a certain region. In the first set (*left*) the bromide and bromate both brominate (add bromine to) the malonate to form bromomalonate. During this process the phenanthroline dye is red, with its iron atom in the ferrous form. If the concentration of the bromide drops below a threshold level, then the second set of reactions (*right*) starts to dominate. The last vestige of bromide is consumed and the bromate takes over the bromination of the malonate. Simultaneously it oxidizes the iron atom in the indicator dye, changing it from red ferrous phenanthroline to blue ferric phenanthroline. Accumulated bromomalonate now reduces ferric phenanthroline back to its red form ferrous phenanthroline, releasing bromide and carbon dioxide. High concentration of bromide shuts off this reaction sequence and restarts the red stage. Other substances such as oxygen in the air are also involved but they are not shown.





INVOLUTE SPIRALS IN *Z* REAGENT have characteristic geometrical properties. Any line drawn perpendicular to the advancing spiral wave (*left*) and extended back toward the center of the spiral will be tangent to a small circle (*white*) at the center. This small circle delimits the spiral's core. The circumference of the circle is the same length as the spacing λ_0 between waves of the spiral. The visible wave front is a "concentration isobar," that is, a contour line of equal chemical concentration within the solution. Along any isobar each tiny glob, or small volume, of reagent has the same concentration as the one next to it. Outside the spiral core all concentration isobars are involute spirals, and the isobars of different substances run parallel to one another. Inside the core (*right*), however, the isobars cross over one another. Isobars are shown for two different reactants *A* (*solid lines*) and *B* (*broken lines*). Black lines represent highest concentration, dark gray lines a lesser concentration and light gray lines the lowest concentration.

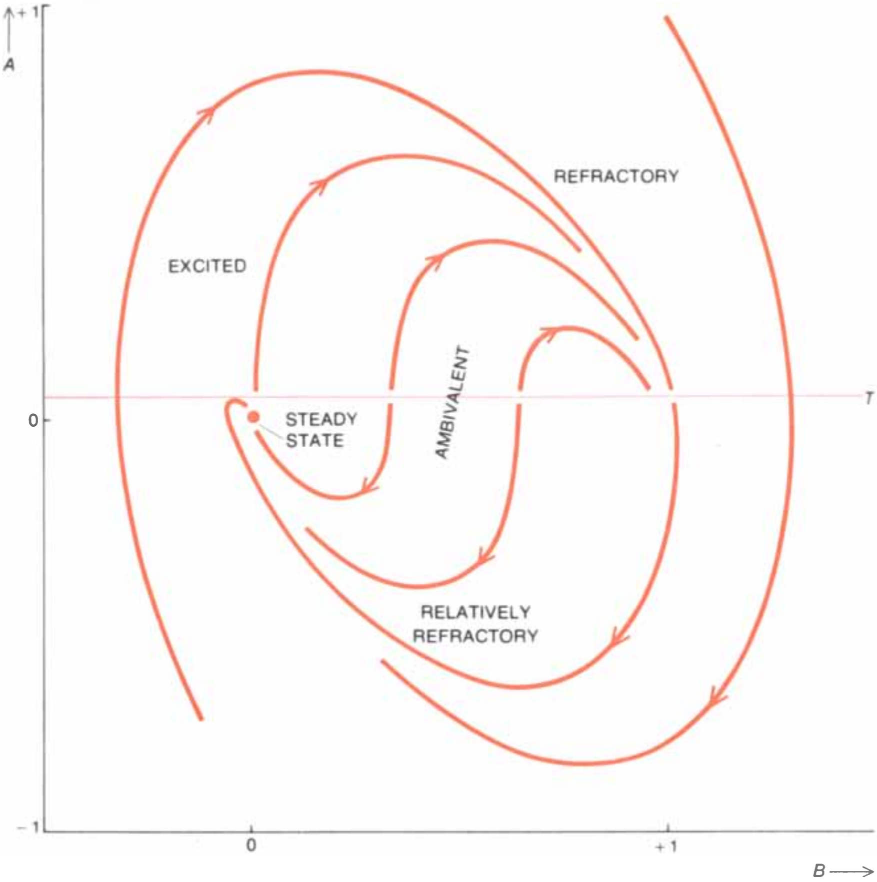
ord." Combining the local description (*a*) with the global description (*b*) according to a geometrical argument, one can derive that the perpendicular spacing, λ_0, between adjacent turns of the spiral should be the same everywhere, namely $v_0\tau_0$. Moreover, the inward extension of any line drawn perpendicular to the spiral wave should be tangent to a small circle of circumference λ_0 near the center [*see bottom illustration on opposite page*]. The spiral wave is the involute of that circle: the curve traced by a pen tethered to the circle by a taut string and winding around it. If the spacing between waves, λ_0, is just over two millimeters, the circle is about three-fourths of a millimeter in diameter.

Just ahead of the wave front the phenanthroline indicator is uniformly red (reduced); just behind the wave front it is blue (oxidized) and gradually fading back to red. At the wave front itself some substance (presumably the bromide ion) apparently passes through a critical concentration that shifts the balance of oxidation and reduction. The visible wave front is the geometric locus of points of a particular concentration; let me call it a concentration "isobar," borrowing the meteorologist's word for a locus of uniform pressure on a weather map. Every volume element, or tiny compact volume of reagent, along the isobar has the same concentration. Let me drop "volume element" in favor of the shorter "glob," meaning a volume small enough that any differences in the concentration of chemicals within it can be ignored.

If it is true that globs at each distance behind the visible wave front have nearly identical chemical composition, it logically follows that the concentration isobars of all substances involved must also be involutes of the small central circle. As the concentration isobars rotate around the circle, each glob along a particular isobar goes through the same cycle of changes in composition: red, blue, red, blue and so forth.

But here a paradox emerges: It is impossible that globs near the center of a rotating concentration pattern should go through that same cycle of changes in each rotation. If they did, it would mean that there are arbitrarily steep concentration gradients near the pivot, a situation that could not persist for long in the face of molecular diffusion. What went wrong?

This paradox stems from the one I referred to above in saying that a chemical model involving only one variable will not suffice to explain a spiral wave.

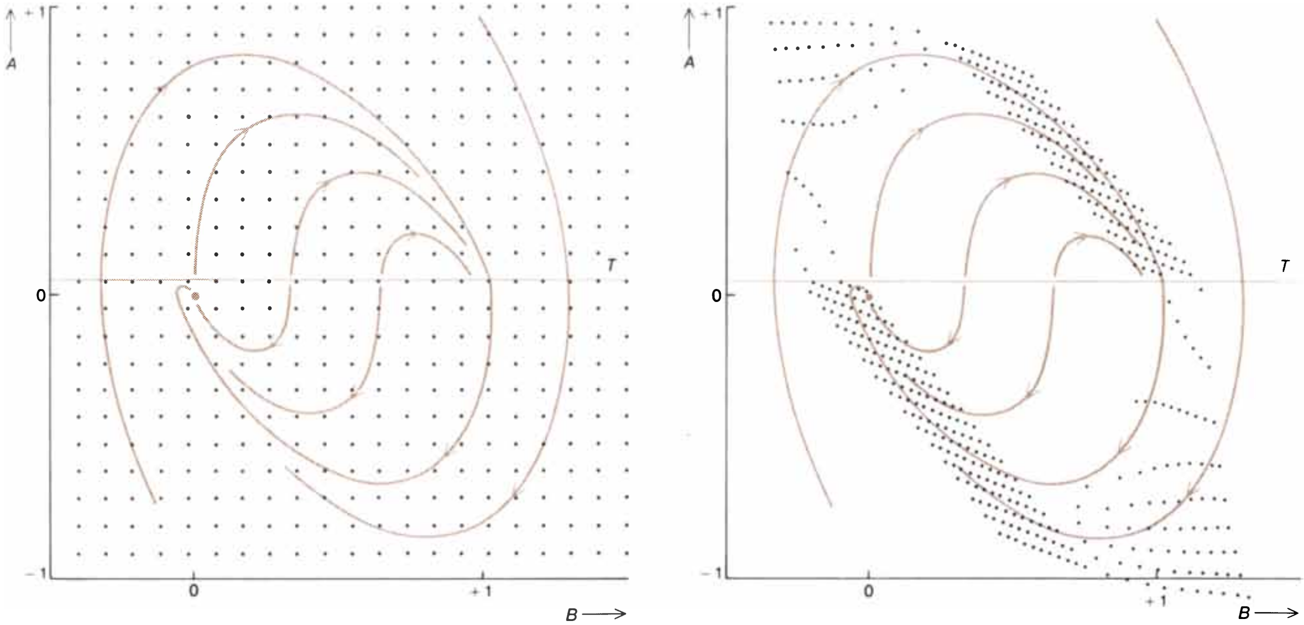


CONCENTRATION SPACE is a useful device for pictorially showing the reaction kinetics of A and B. The position coordinates (A,B) of each point on the diagram represent a possible composition (A,B) of a glob in the excitable medium. Thus an increase in the concentration of A corresponds to the upward direction; a decrease in B corresponds to the leftward direction. The reaction equations specify how A and B change from a given set of initial values, and thus specify a trajectory (*colored arrows*) through each point (A,B). In the illustration B changes at a rate KA (where K is some constant of formation), and A changes at the rate of $-(A+B)$. The steady state, to which all "positive" and "negative" concentrations are referred, is at the point where $A=B=0$ (*colored dot*). Above the threshold T in the excited region A's rate of change becomes $-(A+B)+1$. The concentration of A increases quickly for a short while before all trajectories return to steady state.

Inside the magic central circle, in the spiral's core, the local description of the spiral (*a*) and the global description (*b*) are incompatible. In mathematical terms, the differential equation engendered by these two descriptions has no solution closer than a distance of $\lambda_0/2\pi$ from the center of the spiral. And what becomes of the concentration isobars in that region? They cannot just end; every glob has some chemical composition, and the composition of the solution must vary continuously in space.

Suppose it were possible to make a tour of inspection clockwise around the boundary of the spiral's core. We would proceed in order through globs in each phase of the wave cycle, from the triggering of red to blue through the relaxation of blue back to red. Since for each substance involved "what goes up must come down," each concentration isobar must be encountered once as we proceed from low concentration to high concentration in our clockwise tour, and once as we continue back around from high concentration to low concentration again. These two arms of the same isobar must somehow connect smoothly through the spiral's core [*see bottom illustration on opposite page*].

These isobars, however, are moving: the entire pattern rotates rigidly around the spiral's core. Hence the concentration is increasing in globs on one of the arms and decreasing in globs on the other arm. Such behavior can only be accounted for if some second substance has a different concentration on the two arms of a single isobar. Thus the second substance, although its isobars are also involutes connecting through the spiral's core, must have a different pattern of rising and falling concentration along the



CHEMICAL REACTION ALONE, without diffusion, propels globs (*small black dots*) along trajectories through concentration space. Here 441 globs are set up with diverse initial compositions at a time zero (*left*). (Grid corresponds to a small area of medium .7 millimeter on a side in real physical space.) Each glob independently follows its appropriate trajectory (*right*) since in the computation diffusion between adjacent globs has been omitted. Seven and a half seconds later they are accumulating at the steady state.

circular boundary of the core. The result is that inside the core, concentration isobars are not parallel as they are outside it: they cross through one another.

This is the critical feature of the core that distinguishes it from the spiral's outer regions. The outer regions merely propagate the disturbance created by the rotation of the core. Along the isobars outside the core, the composition of each glob can be specified by a single number something like the phase in Wiener's cycle of states from rest to excited to refractory to rest. Inside the core the composition of each glob appears to vary independently in two separate ways. If that indeed is what is happening, then by artificially setting up appropriately crossed gradients of concentration in the Z reagent it should be possible to create a spiral core that as it rotates emits a spiral wave. And experimentally that seems to be true.

But why is it true? What are the basic processes that engender this colorful mode of spatial and temporal organization in a reaction that is also capable of simply oxidizing malonate in a monotonously slow, uniformly red steady state? There are two factors collaborating to generate the spirals, each of which by itself would result only in a homogeneous steady state. One is the purely physical process of molecular diffusion; the other is the purely chemical process of reaction. In combination the two give rise to an activity that is almost lifelike.

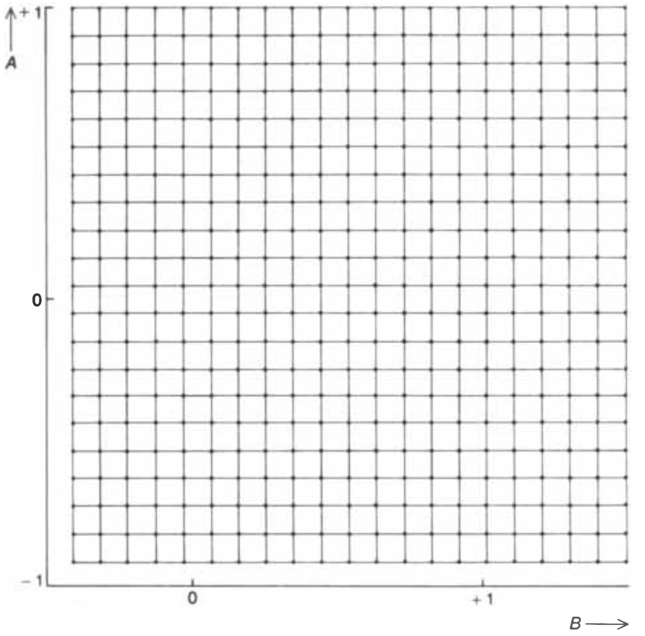
First let us look closely at diffusion. Exactly what is happening? At the molecular level diffusion is the consequence of particles wandering from a crowded region to a less crowded region more often than the other way around, simply because there are fewer to wander from the less crowded region. Diffusion causes the concentrations of substances in two adjacent globs of solution to change toward each other at a rate proportional to their difference in concentration. The process stops only when no difference remains. In the world of concentrations diffusion is the great equalizer—an innocent-seeming suspect for the instigation of spatial structure in chemical solutions. All the more reason to look with care!

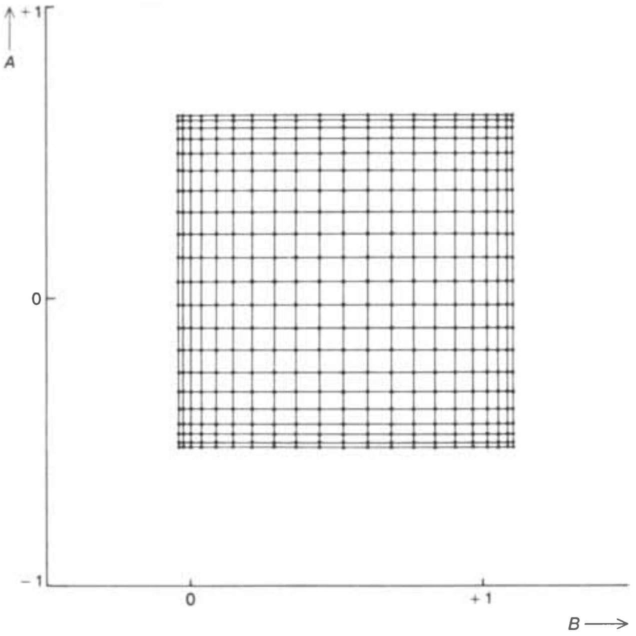
Let us look next at reaction, the conversion of populations of colliding molecules from one chemical form to another. Forget about physical space, concentration gradients and diffusion; think instead about an isolated, homogeneous glob. Within the glob concentrations are changing at rates determined through the equations of reaction kinetics. What kind of reaction kinetics characterizes an excitable medium?

Analysis of quantitative models by computer shows that much of the behavior of spiral waves is almost independent of the chemical details. Instead it derives largely from the qualitative features of excitability. Let us then adopt the simplest two-variable caricature of the dynamics of an excitable reaction. Let A and B be the concentrations of two substances, measured as positive or negative deviations from their respective steady-state concentrations. Let us postulate the following kinetics: B changes at the rate KA, where K is some constant of formation, and A changes at the rate $-(A + B)$. Thus at the steady state A and B are zero, as required by our definition. When A, however, exceeds some threshold or trigger concentration T, its rate of change becomes $-(A + B) + 1$. This stipulation provides the needed excitability.

If K is large enough, the corresponding equation behaves like a mechanical clock, with A and B associated with the position and the momentum of the escapement. Until A is pushed above T ("tick") nothing happens; then it comes back down ("tock") and overshoots the equilibrium point past T again ("tick, tock, tick . . ."). Alternatively, if K is a little smaller and A is interpreted as the membrane voltage in a living cell, then the equation is a stripped-down version of the complex differential equations describing the excitability of nerve and muscle tissue, the equations for which Alan L. Hodgkin and Andrew F. Huxley were awarded a Nobel prize in 1963.

In nerves "tick" is followed only by "tock" and then a return to the steady state. The rate equation for A can be supplemented by a term representing the "diffusion of voltage" from nearby regions along a long fiber of such an excitable medium. In this way John Rinzel





DIFFUSION ALONE, without reaction, in the same 441 globs is started with the same initial compositions as the ones on the opposite page at time zero *(left)*. The difference here is that the globs are linked to one another, each link representing their connection by molecular diffusion. Diffusion tends to equalize concentrations of *A* and *B* in adjacent globs, so that they draw together in concentration space. Thus the grid contracts on itself like an elastic membrane as shown at right seven and a half seconds later.

and Joseph Keller of the Courant Institute of Mathematical Sciences in New York have obtained analytic solutions for the essential features of a "tick-tock" pulse propagating in a one-dimensional nerve fiber.

The important implication for our case, however, is that if *A* and *B* are interpreted as being concentrations of two substances subject to diffusion at equal rates, the same equation mimics the excitability of the *Z* reagent.

Now let us exercise our geometrical intuition on a pictorial summary of these kinetic equations. The essential device is to let the concentrations *A* and *B* be the position coordinates of a tiny glob of reagent in an imaginary "concentration space." Each point (*A,B*) in this space represents a possible composition of the glob as defined by its concentration coordinates *A* and *B* [*see illustration on page 87*].

The chemical composition of a reacting solution cannot be pinned in place like a dead butterfly. It changes as *A* and *B* are synthesized and destroyed in reactions whose rates are governed by the local concentrations of both *A* and *B*. In geometrical terms one can imagine a little arrow at each point in the concentration space showing how *A* and *B* would jointly change in the next instant. By following the arrows one sees in one timeless glance the entire gamut of possible kinds of behavior in the isolated glob. Starting from any initial composition, the future of its concentrations of *A* and *B* unfolds along trajectories through concentration space.

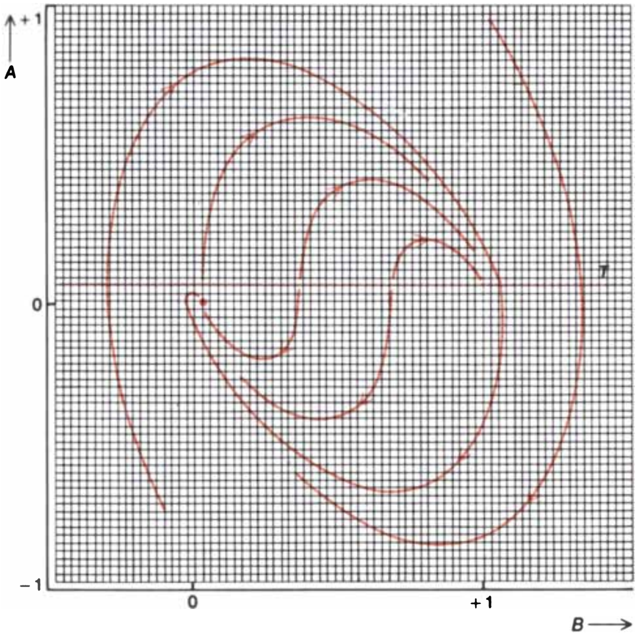
By following the arrows in the illustration on page 87 it is possible to see how *A* and *B* change simultaneously, starting from any initial composition in an isolated homogeneous glob of excitable reagent. It is evident that "all roads lead to Rome," that is, to the steady state where both *A* and *B* equal zero. From just beyond the threshold *T*, however, the roads make a wide excursion on the way back. This excursion takes a glob through an excited phase at high concentrations of *A*. The result is infectious, because the diffusion of molecules of *A* into an adjacent glob at rest will bring its concentration of *A* too above the threshold.

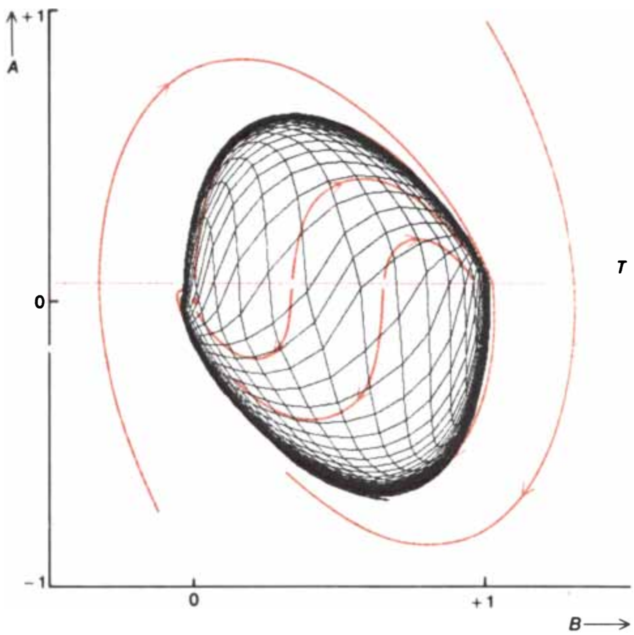
While the glob is above the threshold, the synthesis of *B*, being proportional to the concentration of *A*, proceeds faster than usual. As *B* accumulates, the rate at which *A* decays is increased. The concentration of *A* comes down through a refractory phase in which subsequent contact with excited globs cannot do much to alter the inevitable recovery back to the steady state. The last stages of recovery are relatively refractory because only an extraordinary "shot in the arm" of *A* could immediately raise the recovering glob above threshold concentration again.

The abrupt color transition in the *Z* reagent from red to blue is apparently triggered when the concentration of *A* exceeds the threshold. Reagent whose composition is near the steady state is red.

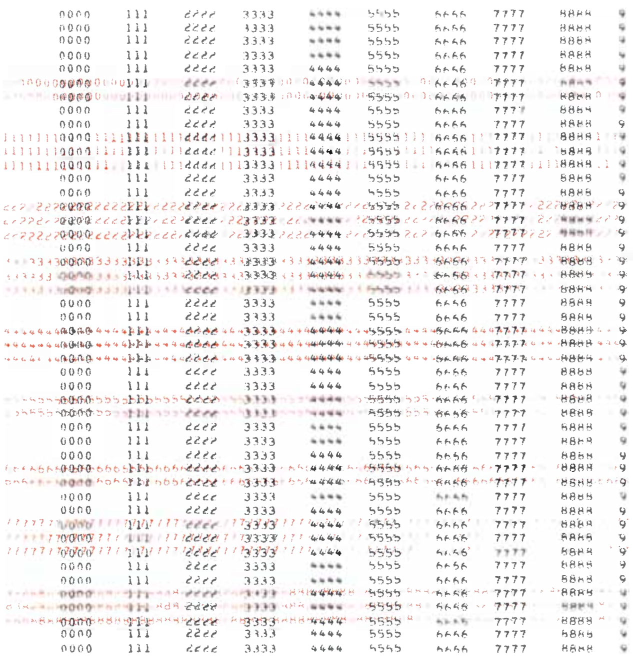
To understand the structure of a rotating wave one must keep track of the spatial distributions of *A* and *B*. Imagine each tiny glob of reagent in a small square dish mapped into concentration space: the composition (*A,B*) of each glob in the dish determines its coordinates (*A,B*) in concentration space. Since adjacent globs have a similar composition, the mapping is smooth, without holes or discontinuities. Suppose the concentration of *B* increases from top to bottom in the dish, and *A* increases from left to right so that there are gradients of concentration crossing through ambivalent compositions in the middle of the dish. The kinetics of chemical reaction now moves each glob (not physically but in concentration space) along a certain trajectory. The movement is generally clockwise around the ambivalent region in the center except for a narrow range of compositions between the steady state (where both *A* and *B* equal zero) and the threshold *T*. In that region the arrows point counterclockwise back to the stable steady state.

If there were no diffusion, globs above the threshold would continue around clockwise while those just below the threshold would fall back to the steady state through a decreasing concentration of *A*. Instead globs initially just below the threshold have their content of *A* supplemented by diffusion from their neighbors now well above the threshold;





REACTION AND DIFFUSION TOGETHER collaborate to support sustained activity in an excitable medium. Here a computer drawing shows what happens in a section of medium 3½ millimeters square using a fine grid of 5,625 globs in concentration space. Initially the globs are arranged to achieve crossed concentration gradients (*left*). After a time equivalent to 315 seconds the grid is rotating stably. Globs on the rim periodically approach the steady state, but then they are pulled across the threshold by diffusion from their neighbors. Globs in middle of grid are stretched across the ambivalent region (*right*) and change relatively little.



COMPUTER SIMULATION shows in real physical space the results of chemical reaction and diffusion together in the imaginary concentration space (*see illustration above*). Initial concentration isobars of A and B are crossed in a square dish of excitable medium 3½ millimeters on a side. Concentrations are indicated by digits from 0 to 9. The concentration of A (*numbers in black*) increases from left to right in the square dish and the concentration of B (*numbers in color*) increases from top to bottom. Middle of the dish corresponds to the ambivalent region in concentration space.

SPIRAL WAVE FORMS IN COMPUTER SIMULATION shown here after a length of time equivalent to 315 seconds. Again the concentration of A is indicated by the numbers in black and B by the numbers in color. The solid black line indicates where A is increasing through the threshold and where the sharp blue wave front would be in the Z reagent. Concentration isobars of A and B are parallel away from the spiral's core (*circle in color*) but cross over one another within core itself. Successive computer printouts are alike except for a rotation. Period of rotation τ_0 is 30 seconds.

thus they are pulled up over the threshold in spite of themselves. Once that has happened they provide the same service to their neighbors, and so the process continues. Like a perpetual-motion machine, the dish of reagent cartwheels around the ambivalent region in concentration space.

The diffusion of molecules between adjacent globs confers on the reagent a kind of coherence in concentration space that acts to average out local variations in the reaction rates: the narrow range of compositions in which isolated globs backslide toward the steady state is smoothed over and the general tendency for them to rotate in concentration space prevails. Moreover, at the same time that the image of the reagent rotates in concentration space, the concentration isobars physically rotate in the reagent.

This situation evolves only if suitably steep concentration gradients are provided at the outset. If the reagent is initially almost homogeneous throughout the dish, then its mapping into concentration space is quite compact. In this case diffusion only assists the general tendency for the reagent to remain in a spatially homogeneous, red steady state.

Such arguments as these must and can be made with greater care. For example, they do not necessarily work if the diffusion coefficients of A and B are too unequal; instabilities then crop up to destroy the simplicity of pictorial reasoning. In order to check whether a stable spiral wave can indeed evolve from excitable kinetics coupled with unrestricted diffusion, as I have argued above, one needs only to follow the equations from moment to moment in each of many tiny globs in the dish, adding at each step a small mathematical term describing the exchange of molecules between adjacent globs in proportion to their concentration difference. In principle it is all very simple, but in practice it is so tediously repetitive that a computer proves indispensable for doing the bookkeeping and for graphing the results.

Soon after sketching out the necessary programs I was fortunate to encounter Patrick Murphy, an undergraduate student at Purdue University with unusual talents for efficiently programming digital computers. Murphy carried out numerical solutions of the simple model I have outlined, and the results bear out our expectations from the model [*see illustrations on pages 88 through 90*]. Similar rotating waves result from a variety of more complicated kinetic schemes as well, just so long as they retain the essential features of excitability: trajectories of a glob in concentration space must converge toward a single steady state, not far from a threshold beyond which trajectories find their way back to the steady state only after a long, high-speed odyssey through extreme concentrations.

To me this interpretation seemed reasonable until I discovered an anomaly in some old photographs of waves in Z reagent. It had often happened that before a thin layer of reagent organized itself into involute spirals the prevailing periodic waveforms resembled greatly elongated rings and spirals [*see bottom illustration on page 84*].

The elongated spiral resembles the ordinary involute spiral except that its inner end point is not close to a pivot. Instead of rotating once with the period τ_0, the end point shoots back and forth with the period τ_0 along a slitlike arc that can be as long as 10 or 20 millimeters. More perplexing, the elongated rings emerge not from a pacemaker point but from a slit that divides itself in half first to the right and then to the left (completing a ring), then to the right and back again, also with the period τ_0.

Both kinds of elongated source can fragment into a greater number of less elongated ring and spiral sources. In such cases parity, defined as the number of clockwise spirals minus the number of counterclockwise spirals, is conserved. During the interludes between such fragmentations elongated ring and spiral sources continually shorten toward greater circular symmetry. The spiral source contracts to the familiar core, spinning out an involute wave. The ring source, however, does something that seems to require consultation with a Zen master: without warning, it simply vanishes an instant before its rings achieve perfect symmetry! Waves that were already emitted, from the outermost and most elongated one to the innermost and most nearly circular one, propagate away from the center at a constant velocity, leaving an enlarging disk of red quiescence where a moment ago blue waves were being created at precisely regular intervals.

Since color betrays the concentration of the reactant phenanthroline, it seems obvious that these waves depict the evolution of a pattern of concentration isobars. Yet I could not come up with an interpretation of these anomalies that was consistent with plausible kinetics and continuity in two-dimensional space. What had gone wrong? The way out of the dilemma seems obvious in retrospect: abandon the "self-evident" principle that color contours reflect chemical concentrations. They do in a strictly two-dimensional medium, but in Z reagent a depth of as little as a millimeter might be important. If there is spatial structure in depth, then the color seen in projection through that depth would indicate only the average concentration of the reactants in each vertical column of fluid.

To test this possibility I squeezed a droplet of Z reagent between two parallel Plexiglas plates separated by three tiny rollers so that waves could be sheared and broken by rolling one plate over the other. Elongated sources arose and gradually decayed, but only in layers thicker than about the diameter of a spiral's core. In thinner films none appeared, or one might say that they decayed instantly to form waves whose free tips immediately began to curl into the usual involute spirals. Thus the two-dimensional model of reaction and diffusion may need only to be generalized for service in three-dimensional space.

Imagine that spiral cores can lie tilted almost onto their sides in layers of liquid thick enough to accommodate them. Then in three dimensions the core must extend like a thread through the liquid. The wave emerging from it must be a scroll unwinding from the thread: the involute spiral is only a transverse section through a scroll wave [*see top illustration on page 93*]. Some effort with a sketch pad shows that in projection through a layer of liquid two or three times thicker than the diameter of a spiral's core, a scroll wave whose core slants from one interface to the other would resemble an elongated spiral emerging from a slitlike source. A scroll whose core touches only one interface at two places by bending around like a *U* would resemble a nest of elongated rings around a slitlike source. All such waves would have the period, τ_0, of the scroll's rotation.

What happens when the scroll's core lies exactly parallel to a nearby interface? If it lies close enough, it seems to gravitate into the interface and vanish. A reason for such behavior is offered below, but first let us look at its implications for the decay of elongated sources.

If the scroll core of an elongated ring source comes too close to the opposite interface and gravitates into it, there remain a pair of spiral sources, each about half as long as the original. One rotates clockwise and the other rotates counterclockwise. The ring source is assigned the value 0 for its parity, as defined above; the spiral rotating one way has a parity of 1 and the spiral rotating the other way has a parity of −1. In this event, then, parity is conserved: 0 = 1 − 1. If the core of an elongated spiral wave (parity = 1) passes too close to one interface on its way to the other interface, it may similarly break into a shorter elongated spiral source (parity = 1) and an elongated ring source (parity = 0). Again parity is conserved: 1 = 1 + 0. If a not very elongated ring source (parity = 0) gravitates into the same interface that its two ends are touching (which is a possibility if the ends are close together and the source is generating nearly symmetric rings) then the ring source may vanish altogether. Parity is conserved: 0 = 0.

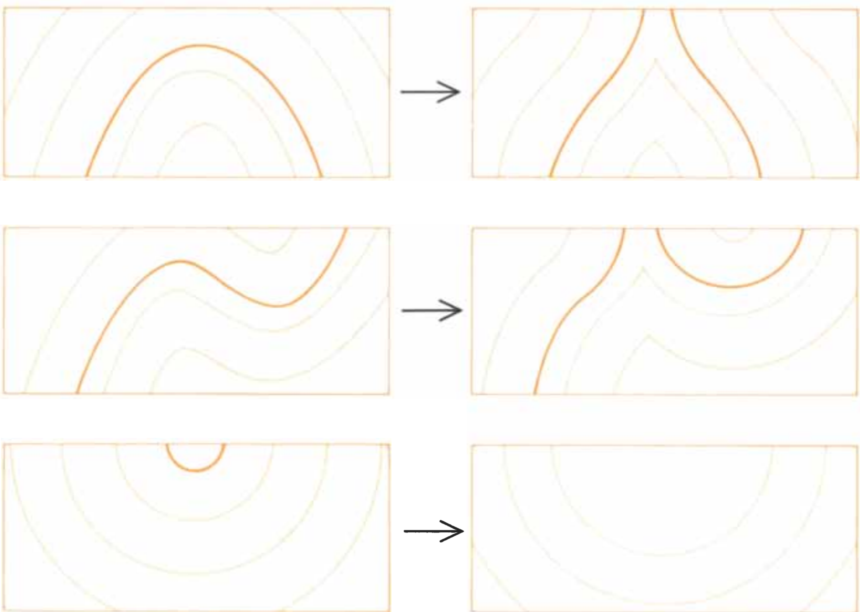
But why should the scroll's core be attracted to its destruction on nearby interfaces? We saw that a square dish of idealized reagent with the concentration gradients of *A* and *B* crisscrossing through the ambivalent region will stably rotate in concentration space, since the trajectories taken by the globs reacting without any diffusion diverge enough from the ambivalent region to counterbalance the homogenizing effect of diffusion. By way of contrast, reagent whose concentration gradients do not overlap the ambivalent region of concentration space simply lapses into homogeneous quiescence, its globs following trajectories that all point roughly the same way. Of course, between these two extremes there are intermediate cases where the reagent rotates a short time as it gradually pulls off to one side of the ambivalent region before swirling into the steady state. To create such cases, the dish of reagent is started off-center in concentration space, with one side too close to the ambivalent region. To put it another way, the spiral's core is started too close to the side of the dish. As the reagent drifts across and past the core region, the core drifts into the side of the dish. It seems that at least in this simple two-dimensional model the spiral core is indeed attracted to nearby boundaries of the medium where diffusion acts asymmetrically.

If in three-dimensional real space the scroll core wanders between parallel interfaces, sometimes coming too close to one of them, it might be expected to pull into the interface along such areas, fragmenting the thread from which the scroll wave emerges. As we have seen, such an event would account for the elongated sources and the way they decay. To find out whether any particular analytic model actually *does* these things, it would be necessary to numerically integrate the equations for reactions and diffusion on a three-dimensional grid of globs, starting from suitably arranged initial concentrations. Moreover, a particularly fine grid would be required because we are looking for an instability and must be sure that it is not an artifact of our dividing the reagent into discrete cells. Such a numerical simulation is prohibitively time-consuming even on the fastest computers. Let us instead turn back to the chemical reagent as an analogue computer.

A binocular dissecting microscope does much to confirm one's faith in these rather fanciful speculations. Visual observation, however, can seldom extend through many rotations of a scroll wave. The slightest thermal disturbance or mechanical agitation destroys the developing pattern, if a loosening bubble of carbon dioxide does not do so first. Furthermore, interpreting the pattern in three dimensions while trying to focus the microscope on moving fields of transparent color also leaves much to the imagination. Could these waves be somehow fixed and cut into thin sections for serial reconstruction, in the time-honored style of classical embryology?

Reagent can be gelled by colloidal silicon dioxide without much altering the processes of reaction or diffusion; if it were then possible to stop the reaction completely at an interesting moment, slicing the reagent with a freezing microtome would do the rest. I found that even liquid nitrogen cooled the gel too slowly to prevent it from crystallizing, thus segregating the various chemicals and destroying the pattern.

Reagent that had spilled on the laboratory bench seemed to support waves reasonably well among the paper fibers of my notebook. A variety of newspapers, cigarette papers, filter papers and toilet tissue offered little improvement. Finally I tried Millipore filters, composed of chemically inert cellulose esters riddled with interconnected empty tunnels less than a micron in diameter. They absorbed the reacting liquid and exhibited waves as crisp and colorful as any yet seen. The mechanical binding of the fluid in the tiny tunnels prevents the liquid from moving quickly. Waves propagate, with sharp fronts, at unvarying speed, even as the filter is being handled freely. When the filters are stacked like pancakes, they adhere closely and conduct waves in three dimen-



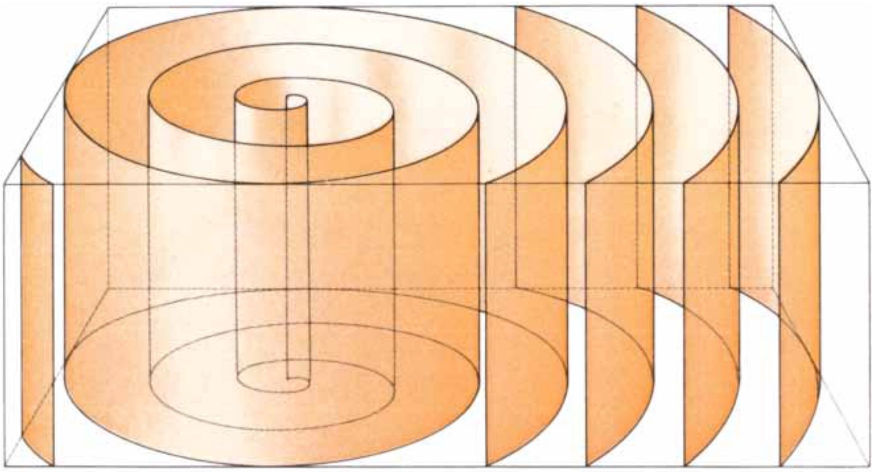
SCROLL CORE MEANDERS through the reagent from one interface to another and back again (*top left*), as in the bottom illustration on the opposite page. The dark colored line is the scroll axis viewed from the side and the light colored lines are emerging waves. If the core pulls into the upper interface (*top right*), there remain two oppositely rotating elongated spiral sources. If the core threads its way from one interface to another (*middle left*) appearing as an elongated spiral source as in the middle illustration on the opposite page, it may fragment into a shorter spiral source and an elongated ring source (*middle right*). An elongated ring source whose axis is very short (*bottom left*) may gravitate into the adjacent interface and vanish (*bottom right*), leaving an expanding area of red solution.

sions just as the unbound liquid does. The stack can be peeled open to expose its contents in cross section.

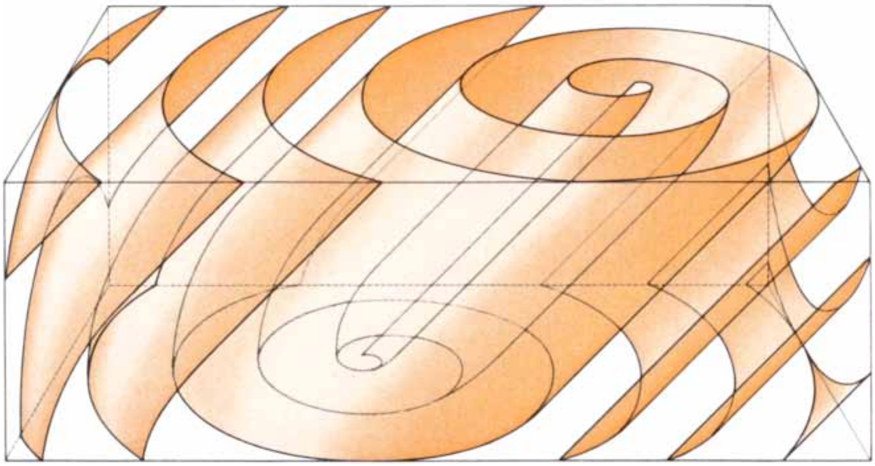
At about this time John DeSimone, David L. Beil and L. E. Scriven at the University of Minnesota discovered how to chemically fix waves of Z reagent in collodion membranes. Although I could not use their method in the Millipore filters, its mere existence encouraged a month of fiddling. Two fixatives emerged. In one process cold saturated salt water stops the reaction and precipitates the red ferrous phenanthroline in the filter while everything else dissolves out. The filter is then floated on dilute solution of iodide in a jar of iodine vapor. As the ferrous phenanthroline redissolves it diffuses to the surface and combines physically with the iodine, precipitating as a metallic-looking golden film on the surface. The other procedure is simpler, quicker and more reliable, although less colorful: the reaction is stopped and ferrous phenanthroline is fixed in a cold solution of 3 percent perchloric acid while everything else diffuses away. The dried filters are then cleared in oil.

This technique has inherent limitations, such as the difficulty of handling stacks of filters more than about a millimeter high, and the fact that invisible carbon dioxide bubbles promptly begin to separate the filters in many places. Moreover, opening the stack ends the experiment because the connection between waves traversing adjoining filters is broken. The easiest scroll to induce in a medium stratified horizontally is a horizontal scroll ring: a scroll wave whose long axis is closed in a ring around a vertical "hole axis." The scroll ring is the *U*-shaped scroll wave joined to its mirror image and flipped onto its side [*see illustration on next page*].

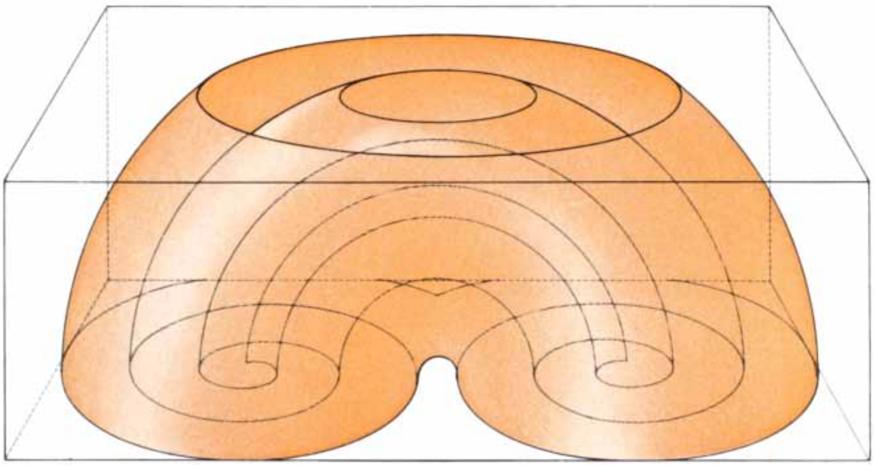
How likely is it that a fine thread of unique kinetic organization should meander through a reacting solution and exactly join onto its other end, forming a ring? Intuition says not very likely. When the phenomenon is examined in chemical terms, however, it seems quite natural. Consider a volume of solution in which the concentration of *A* varies continuously from place to place in a complicated, almost random way. The liquid is stratified by its *A* concentration levels into a series of crinkly, convoluted sheets, each of them a two-dimensional isobar of fixed concentration. The same is true of *B:* the concentration isobars of *B* divide the liquid into some other interpenetrating series of curvy surfaces. Where both concentrations are simultaneously passing through the am-



SCROLL WAVE emerging from a scroll axis standing vertically upright would be seen as an involute spiral if one looked down into the solution from the top. Each colored surface represents abrupt red-to-blue frontier of wave as it propagates outward through the reagent.



ELONGATED SPIRAL could be a scroll wave emerging from a scroll axis tilted on its side in the reagent. Seen from above, the core would blur into a long slitlike source in the center.

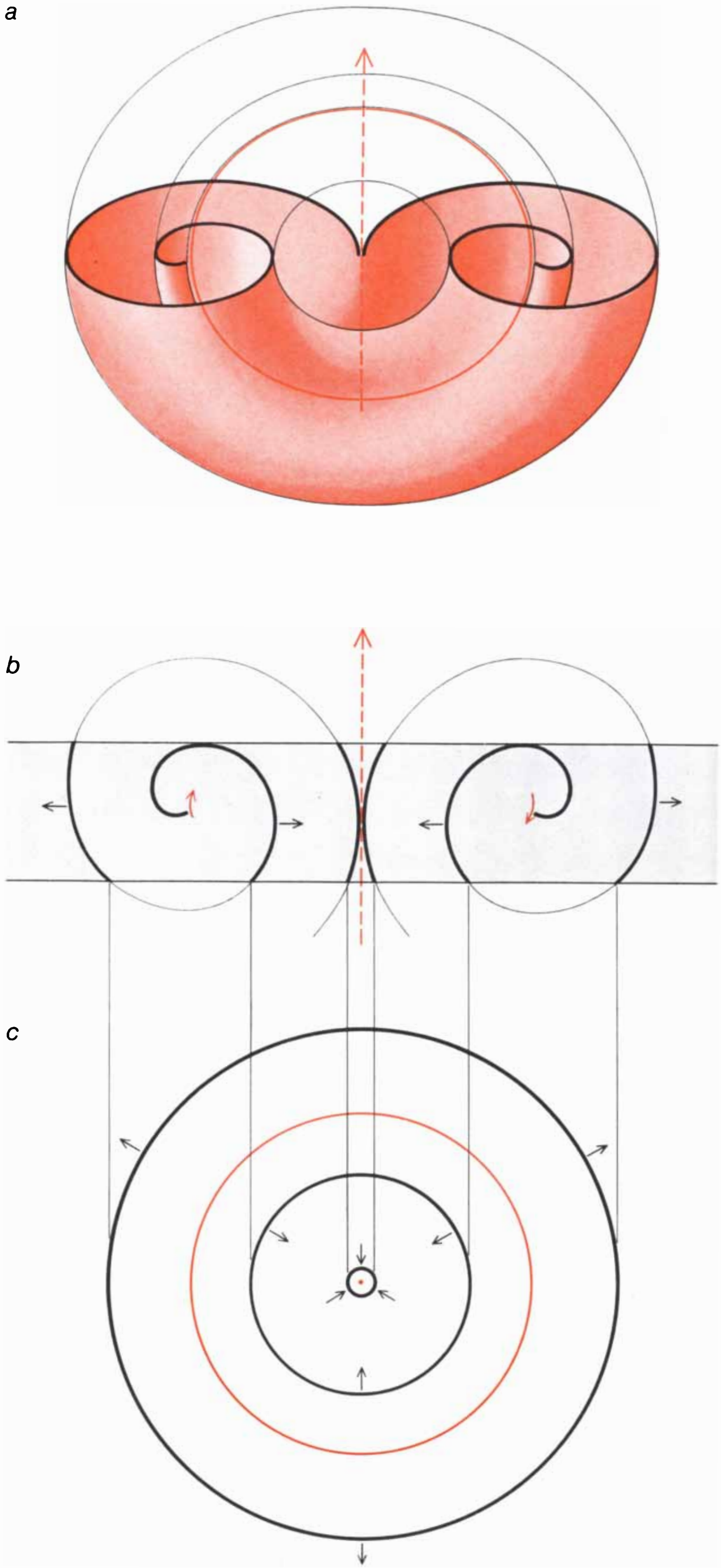


U-SHAPED SCROLL AXIS would emit scroll waves that would appear from above to be concentric waves emerging from a slitlike source dividing itself first to the left and then to the right as the core rotates on its axis. If the two legs of the *U* gradually drew closer together, making the *U* smaller, then the waves would appear to grow more and more circular. Finally, if the *U* gravitated into the same interface as it shrank, then the source would abruptly and mysteriously vanish just before the waves it emitted achieved perfect symmetry.

bivalent region, there we have a bit of scroll core. The scroll core follows the intersection of those critical isobars of *A* and *B*. Moreover, if two curvy surfaces intersect at all in three-dimensional space, they typically do so along closed rings, or fragments of closed rings interrupted by the boundaries of the medium. It appears that if reasonably simple concentration gradients are imposed on a volume of Z reagent, the typical configuration of any resulting scroll core should be a series of long threads, many of which close in rings. The illustrations on the preceding page can thus be viewed as fragments of scroll rings interrupted by the walls of the dish.

Imagine a horizontal section through a scroll ring; here all the waves are circles and the innermost ones propagate inward. These inward-propagating waves will result when, at each interval τ_0 a turn of the scroll strikes the interface broadside and splits into two circles, one propagating inward and one outward. Such structures are often formed in Z reagent that has been stirred gently, but they are typically small and hemmed in by more complicated waves. Is there a way to observe conveniently large scroll rings in isolation?

It is possible to do this experiment with Millipore filters soaked in Z reagent. A cylindrical wave can be triggered in a stack of five filters by touching the center of the stack with a hot filament. When the resulting cylindrical wave (which appears circular on the top filter) has propagated to the desired diameter, a second stack of five filters, still homogeneously in the red steady state, is layered on the top. Where the cylindrical wave abuts on uniformly red reagent, the right kind of crossed concentration gradients are set up. Because the entire experiment is symmetrical around the vertical axis of the stack, the scroll core (if it is stable) must extend in a horizontal circle. A short, predictable



HORIZONTAL SCROLL RING (*a*) with a circular scroll axis (*solid colored line*) and a hole axis (*broken colored line*) through the "doughnut hole" can be induced in a stack of Millipore filters soaked with the *Z* reagent. If the stack of filters (*gray area*) were cut in a vertical plane (*b*), the scroll ring would look like twin involute spirals. (How much of each spiral is contained in the filters depends on height of stack.) If filters were peeled apart in horizontal plane (*c*), scroll ring would look like a series of concentric rings whose wave fronts would propagate in directions indicated by arrows.

time later, and then at intervals τ_0, a circular wave erupts into view on the top filter, almost directly above the place where the cylindrical wave was when the two stacks were joined. Then the wave splits into an inward-propagating circle and an outward-propagating circle. The periodicity of the event betrays the rotation of a scroll core at the source of these waves. The absence of free ends and involute spirals on the visible surfaces of the top and bottom filters indicates that the scroll core must remain in the stack and therefore must close on itself in a ring.

If this stack now contains a horizontal scroll ring, what can we expect to see by peeling apart the Millipore filters? Two features should be conspicuous. First, the inward and outward rings should abruptly appear or vanish in pairs, wherever the wave front is horizontal and propagating vertically. Second, a unique unpaired ring should be found that simply vanishes somewhere inside the stack, where the wave's inner edge approaches the scroll's core. Both features appear in the Millipore filters from stacks fixed and stained after several rotations of the scroll.

Such scroll rings, in which the core meanders through a 360-degree turn and joins up with itself without once twisting 360 degrees around its long axis or tying itself into any knots, may be only the simplest and most readily recognizable variety of closed scroll forms. I see no reason why more elaborately twisted, knotted and linked scroll rings should not exist. If they do, they are apparently unstable or evolve only from initial concentration gradients that are seldom set up by accident.

As the reader must appreciate by now, these tiny, delicate and rapidly moving patterns of color are hard to observe in adequate detail. Moreover, the mathematical description of excitable media does not lend itself to easy solutions. Consequently the story I have told here cannot be construed as a rigorous proof that a stable and homogeneous steady state on the one hand, and scroll rings on the other, are the characteristic modes of self-organization in excitable media. The story is built up of fragmentary evidence loosely held together by a sticky matrix of theory. Nonetheless, one conclusion does emerge clearly, and I hope the reader is already persuaded: that the principles of pattern formation in chemical and biochemical systems are susceptible to the interweaving of experimental and theoretical attack in a new and delightful way.

ICE-AGE HUNTERS OF THE UKRAINE

In the span between 75,000 and 10,000 years ago men hunted other mammals in the rigorous environment of eastern Europe. At some sites they amassed mammoth bones to build shelters

by Richard G. Klein

Modern man has lived mostly in an ice age. To be specific, the kind of human being we know today arose late in the Pleistocene epoch, or Great Ice Age, that came to a close some 10,000 years ago. The Pleistocene was an epoch of spectacular glacial activity, but it was not one of uniformly cold climates. There were many separate glacial episodes followed by interglacial periods; the total number of these alternations is not known exactly.

During each glacial episode enormous sheets of ice advanced to cover much of North America and Europe. When the ice retreated, the climate during the interglacial interval approximated that of today. The next to last major advance, known to scholars as the Penultimate Glacial (Riss II in the Alpine sequence), marked the end of the Middle Pleistocene some 125,000 years ago. The ensuing 115,000 years are the Upper Pleistocene, a period that includes the Last (Riss-Würm) Interglacial and the Last (Würm) Glacial. At some time during the Last Glacial, perhaps between 45,000 and 35,000 years ago, modern man (*Homo sapiens sapiens*) seems to have made his first appearance.

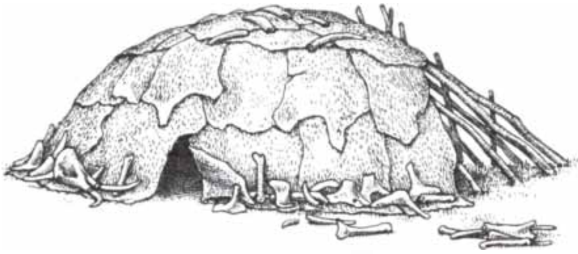
The systematic study of Pleistocene man began only in the middle of the 19th century. Its first focus was in France. Today, after more than a century of investigation, both the quantity of the discoveries made in France and the high quality of the research undertaken there have placed the French data foremost in most accounts of Pleistocene man. The student learns of the progressively more recent Acheulean, Mousterian, Aurignacian/Perigordian, Solutrean and Magdalenian cultures, whose remains have been found stratigraphically superposed in various French cave sites. He might easily conclude either that the rest of the world was sparsely populated during the Pleistocene or that few ice-age sites outside France have been investigated.

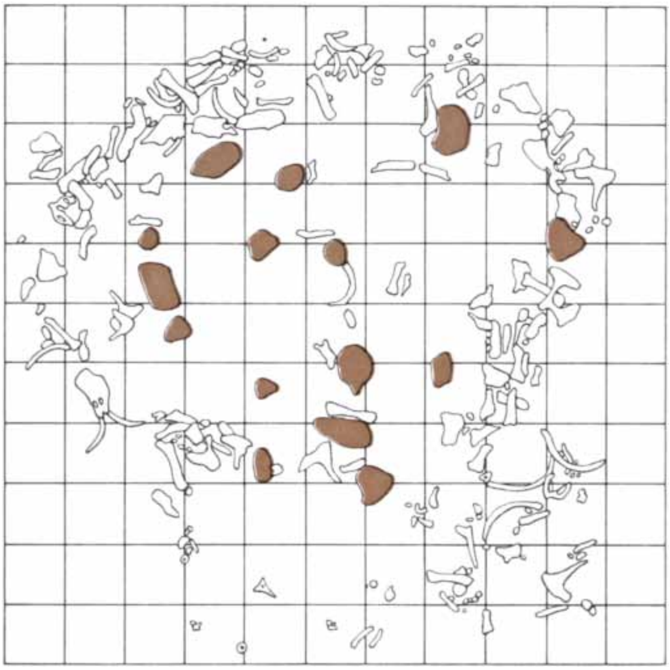
In both cases the reverse is true. For example, the first occupation sites of Pleistocene man to be discovered in central and eastern Europe, many of them spectacularly rich, were unearthed almost as long ago as the first finds in France. Language barriers have resulted in these sites' being poorly known elsewhere, but the information they contain is vital to an understanding of how early man survived and perhaps even thrived under ice-age climatic conditions in Europe. The point is perhaps best demonstrated by what archaeologists have learned over the past century from a group of nearly 100 Pleistocene sites in and around the Ukraine.

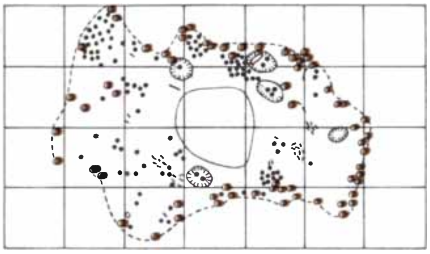
The earliest Pleistocene occupation sites found anywhere in the European U.S.S.R. are roughly between 80,000 and 75,000 years old. They date back to the end of the Last Interglacial [*see illustration on page 100*]. It is possible that older sites will be discovered someday; in Hungary, Poland and Czechoslovakia traces of early man have been discovered that are hundreds of thousands of years old. Alternatively it is possible that it was only at the end of the Last Interglacial that early man achieved the cultural capacity to survive the harsh climate of the Ukraine. Even today the Ukraine's winters are significantly colder than those of its western neighbors. In any event the vast majority of the Ukrainian ice-age occupation sites belong to the Last Glacial; they are between 75,000 and 10,000 years old. Many of the sites include separate occupation levels, and the intervals between the successive occupations sometimes amount to thousands of years.

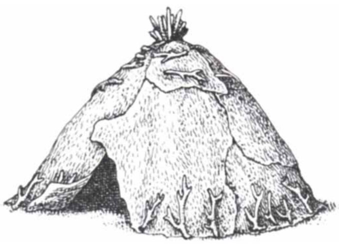
Most of the Ukrainian sites lie in the main river valleys of the region [*see illustration on page 98*]. The reason is not only that ice-age men frequented river valleys but also that valleys are places where conditions favor the accumulation of sediments and thus the burial and preservation of ancient occupation sites. Natural erosion or the activities of civilization often expose the buried sites; as is true elsewhere in the world, many of the most important Pleis-

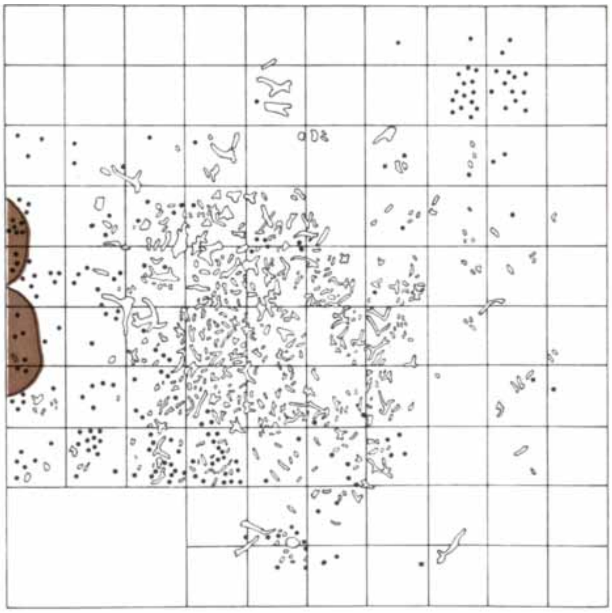
THREE ICE-AGE SHELTERS, built on an ancient terrace of the Dniester River, are shown on opposite page in plan (*far right*) and in hypothetical reconstruction. The first (*top*) was unearthed in the fourth level of the site known as Molodova I; 15 hearth areas (*color*) were surrounded by a rough oval of mammoth bones. Carbon-14 analyses of hearth charcoal show the fourth level is more than 44,000 years old; the associated artifacts are Mousterian. The reconstruction suggests that the shelter consisted of a wood framework covered with skins; mammoth bones evidently helped to hold the skins in place. A second shelter (*middle*) was unearthed in the third level of an adjacent site, Molodova V. A perimeter marked by 64 postholes (*color*) enclosed a single large hearth. Analysis of hearth charcoal shows the third level is about 13,000 years old; the associated artifacts are Upper Paleolithic. The reconstruction suggests a teepeelike wood frame for the shelter; the covering of skins was evidently secured with large wood pegs. The third shelter (*bottom*) is from the next higher level at the same site; the two hearths (*color*) are to one side of the occupational debris. The number of reindeer antlers suggests that these substituted for mammoth bones as a means of securing the skin covering of the teepeelike structure. Carbon-14 dating suggests that it was occupied 1,000 years after the one below it. The grids are one-meter squares.









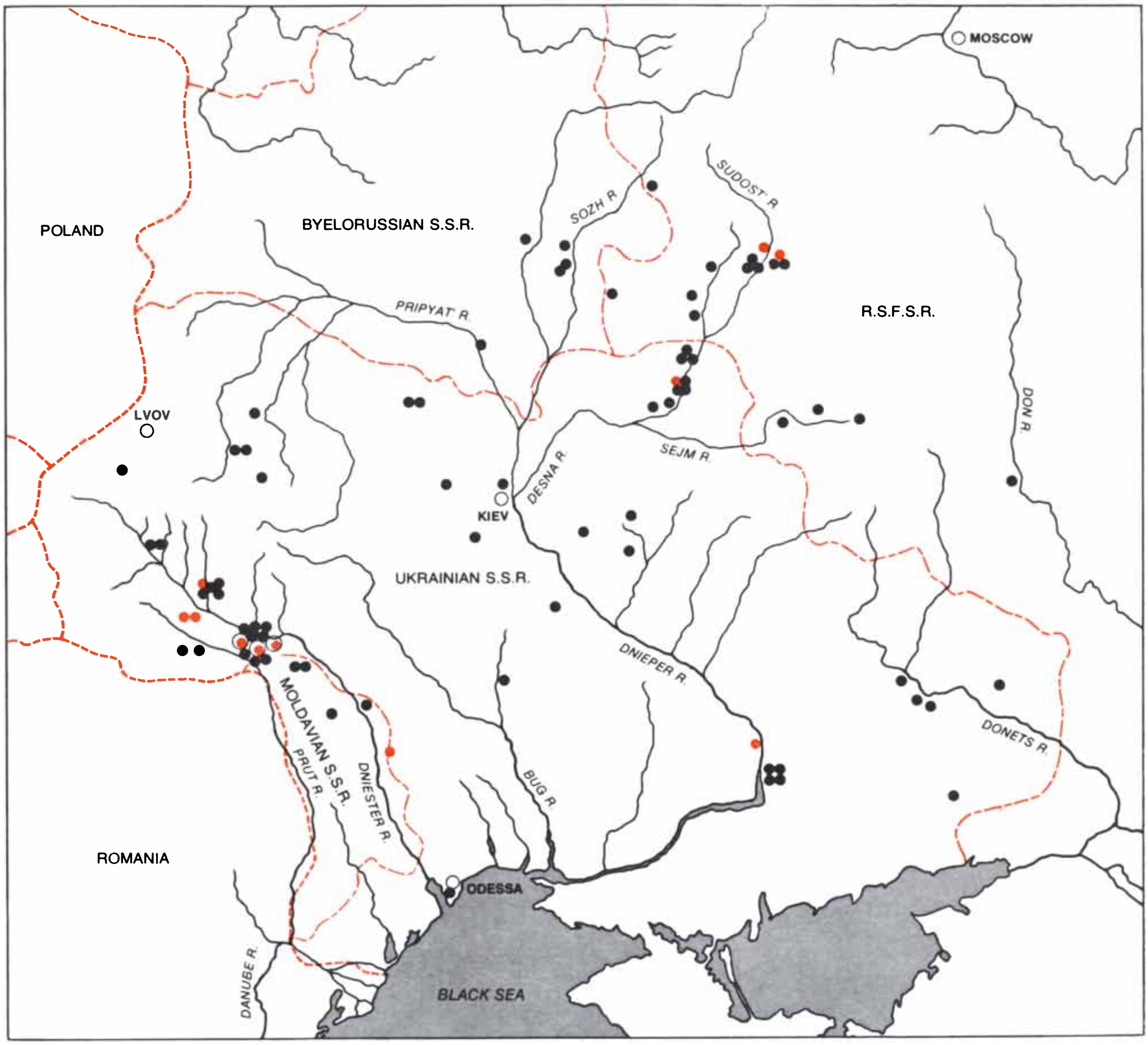


tocene discoveries in the Ukraine were made when road builders or clay diggers happened to uncover strata that contained animal remains and artifacts.

One Ukrainian site that proved to consist of a number of superposed occupation levels is located near the village of Molodova on the Dniester River 150 miles southeast of Lvov. It is one of a cluster of sites in this location and has been designated Molodova V. The most recent of its occupation levels was buried more than 10,000 years ago. The oldest levels, which lie between six and eight meters below the surface, were last inhabited more than 45,000 years ago.

The ages of most Ukrainian sites have been determined primarily by geological analysis. This is partly because of the high cost of carbon-14 dating in the U.S.S.R. and partly because age determinations in excess of 30,000 or 40,000 years are technically difficult to secure by means of carbon-14 assay. At Molodova V, however, a detailed geological analysis by I. K. Ivanova of the Soviet Commission for the Study of the Quaternary period (the period from the present back to the beginning of the Pleistocene) has been supplemented by a large number of carbon-14 determinations.

As at most other Ukrainian sites, the evidence of human occupation is stratified within deposits of sand and silt that were carried to the base of a slope by the combined action of runoff and gravity [*see illustration on page 101*]. Here the base of the slope is an alluvial deposit laid down at a time when the bed of the Dniester was substantially higher than it is today. This old floodplain can be traced downstream to the point where it grades into marine deposits that were formed when the level of the Black Sea was as high as it is today, or even higher. That, of course, would have been during the Last Interglacial, when just as today the level of the world's seas was not low-

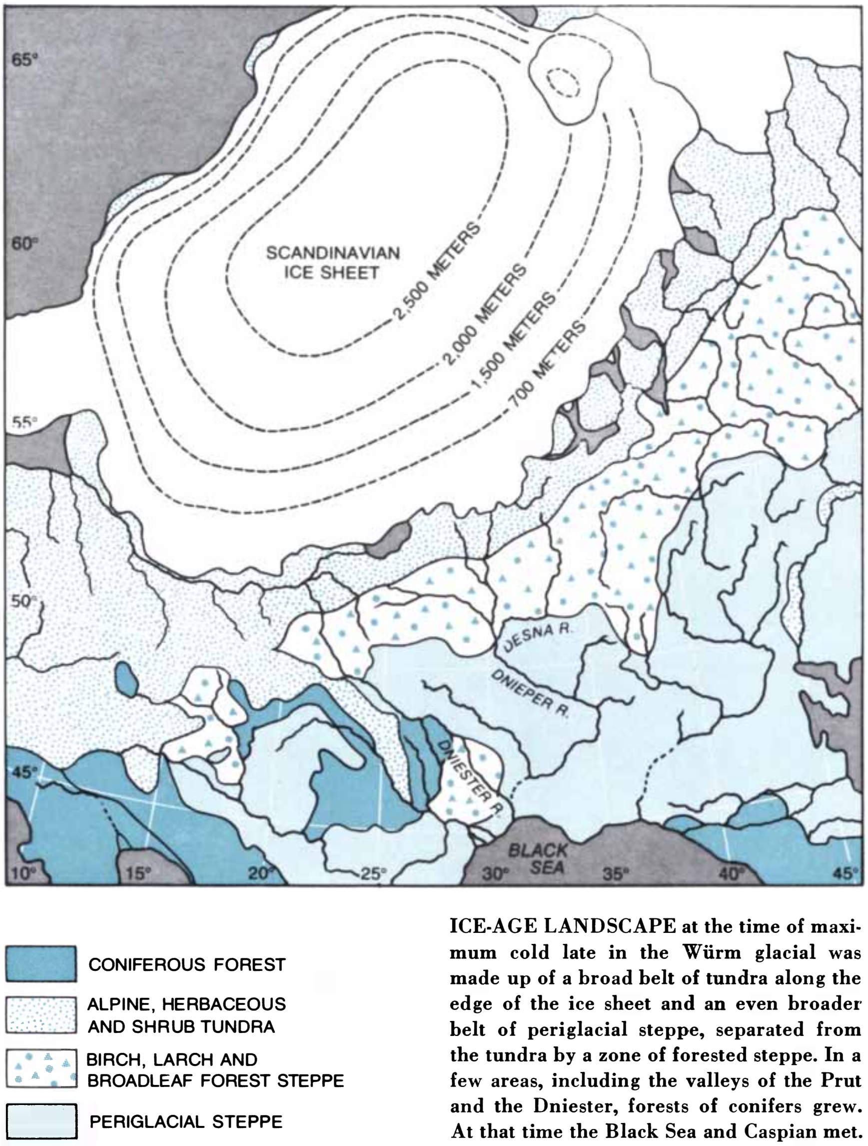


UPPER PALEOLITHIC
MOUSTERIAN
BOTH CULTURES

MAJOR PALEOLITHIC SITES in the Ukraine and vicinity lie in river valleys such as that of the Dniester, the Dnieper-Desna region and the Don or near lesser streams and tributaries such as the Prut, the Bug, the Pripyat', the Sejm, the Sozh and the Sudost'. Because many of the sites lie close together their positioning on this map is not exact. Eight of them (*color*) contain Mousterian artifacts only; 72 (*black*) contain only artifacts of the Upper Paleolithic period. At Molodova I and V and at one other site both cultures are found.

ered by the impoundment of vast quantities of water in the ice sheets that covered the northern continents.

As Ivanova's analysis made clear, because the alluvial floodplain had been formed during the Last Interglacial the sands and silts that rest on it are necessarily younger. That the overlying deposits in fact accumulated during the subsequent Last Glacial is suggested by several kinds of evidence. First, some of them appear to have been formed when moisture-laden earth slid downslope over a frozen substratum. The process, known as gelifluction, is commonplace today in subarctic and arctic zones of permafrost, or permanently frozen subsoil. Permafrost could have extended this far south in the Ukraine only during the Last Glacial. Second, the shells of snails that normally inhabit quite cool environments are present in the deposits, and the animal bones unearthed in human occupation levels are those of such subarctic and arctic species as the reindeer and the arctic fox. Moreover, some of the charcoal in the ancient hearths is from the kinds of coniferous trees found today only in much colder environments. Finally, the carbon-14 determinations at Molodova V, although lacking complete internal consistency, indicate that the youngest of the human occupation zones is more than 10,000 years old. Therefore all the occupation levels predate the end of the Last Glacial (conventionally set at some time about 8000 B.C.). The artifacts and other cultural remains sandwiched among the silts at Molodova V are indisputably the work of men who lived during the Last Glacial.



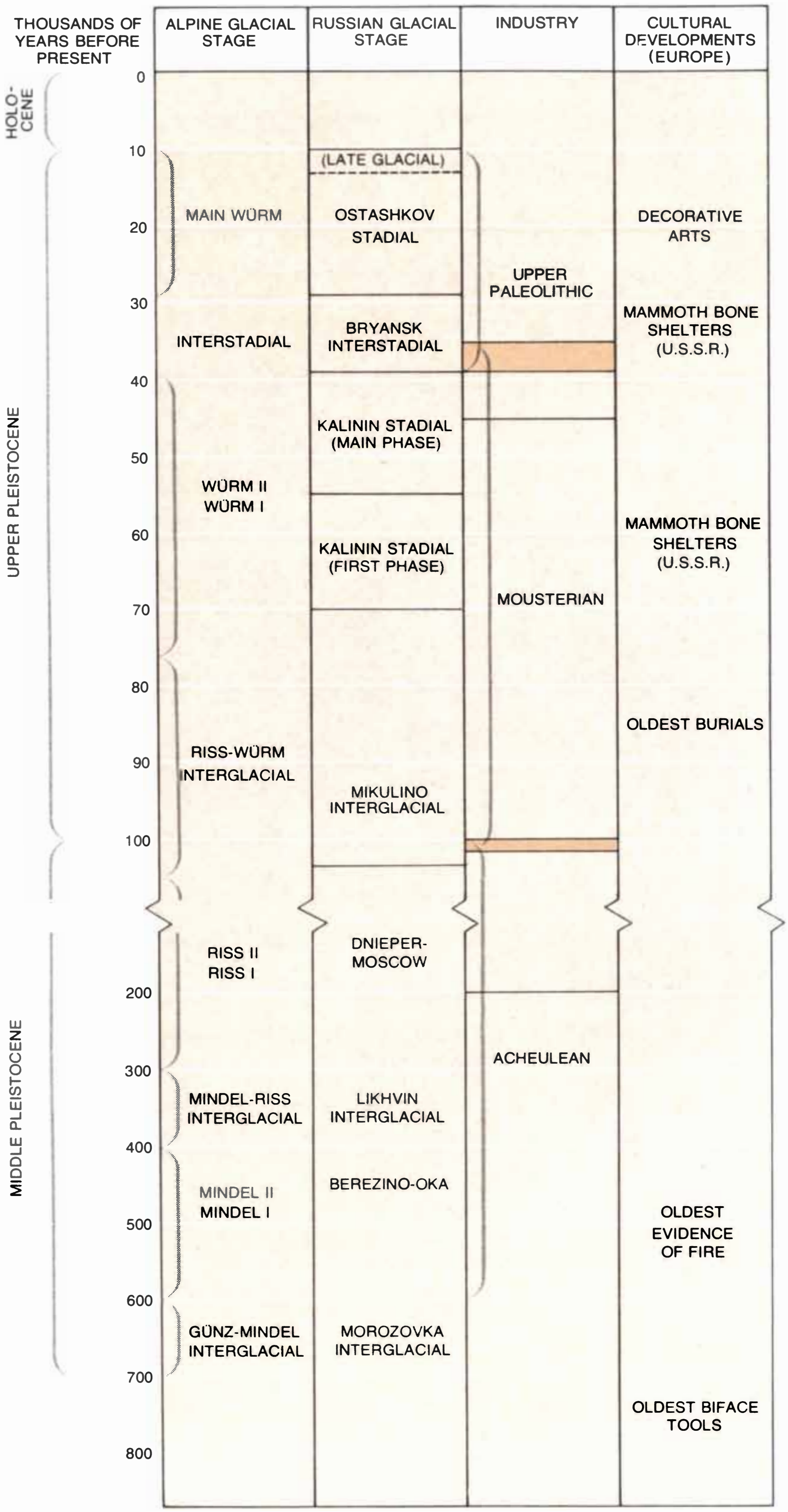
ICE-AGE LANDSCAPE at the time of maximum cold late in the Würm glacial was made up of a broad belt of tundra along the edge of the ice sheet and an even broader belt of periglacial steppe, separated from the tundra by a zone of forested steppe. In a few areas, including the valleys of the Prut and the Dniester, forests of conifers grew. At that time the Black Sea and Caspian met.

Ivanova's analysis of the Molodova V profile demonstrates a further important point: the Last Glacial was a time not of consistently low temperatures but of fluctuating climate. On several occasions the movement of sand and silt downslope toward the river terrace slowed or stopped altogether. Whenever this happened, the surface layer of sediments was subjected to weathering that led to the development of a soil. The times when deposition slowed or ceased and soils formed evidently represent intervals during the Last Glacial when the climate briefly grew warmer. Substantiation of these episodes is provided by the absence from the soils of Molodova V of the same cold-loving snails that are found in the unaltered sands and silts.

Various kinds of data collected both at Molodova V and at numerous other ice-age sites in the Northern Hemisphere suggest that the Last Glacial was composed of three main parts. The first part began with the onset of glacial conditions between 75,000 and 70,000 years ago and lasted until about 50,000 years ago. The second part, consisting of a long interval of fluctuating, milder but not really warm climate, lasted from 50,000 to about 30,000 to 25,000 years ago. The third part, which included the cold maximum of the Last Glacial, then began; this episode continued until 10,000 years ago.

Much of the same evidence that indicates the subdivisions of the Last Glacial in and around the Ukraine also makes possible the reconstruction of the landscape. For example, buried pollen grains show that the region was covered by a kind of steppe vegetation characterized by plants reflecting very cold and dry conditions. The plant communities that were present have no exact modern counterparts; they are called periglacial steppe by Russian investigators. Studies of sediments show that the periglacial steppe was underlain by permafrost that extended almost as far south as the Black Sea. Together with the pollen data, this suggests that the average January temperature, which is below freezing everywhere in the Ukraine today, was then eight or nine degrees Celsius lower.

It may seem surprising that ice-age men could have found a region with such a rigorous climate at all hospitable. The fact is that the periglacial steppe supported a rich and varied fauna. The Ukrainian summers, although short, were warm enough to allow an abundant growth for grazing animals. The winters, although long and cold, were dry; the summer growth was not buried under deep snow and the animals could feed the year round. The bones of no fewer than 13 species of herbivore, large and small, have been found in the Pleistocene sites. Based on the relative abundance of the bones, the primary prey of



CHRONOLOGY OF GLACIAL STAGES in the Middle and Upper Pleistocene in the Alps (*left*) and their equivalents in European Russia is shown in thousands of years before the present. Until 100,000 years ago each scale division indicates a 100,000-year interval; thereafter each represents a 10,000-year interval. The earliest Mousterian stone tools known in western Europe are more than 100,000 years old. The Mousterian may not have reached Russia, however, until 75,000 to 80,000 years ago. In Russia tools of the Upper Paleolithic kind probably replaced those of the Mousterian kind between 35,000 and 40,000 years ago.

the Ukrainian hunters were the reindeer and the wild horse in the north and the steppe bison in the south. The distinction between the game taken in the north and that taken in the south probably reflects some unrecognized variation in local vegetation rather than any preference of the hunters.

The other herbivores pursued as game included the woolly rhinoceros, the aurochs, the musk-ox, the saiga antelope, the red deer, the roe deer, the "giant" deer (*Megaloceros*), the moose and either the wild goat or the wild sheep. Rarely, however, are the bones of any of these animals found in large quantities. The bones of carnivores are even rarer. Only the arctic fox and the wolf are represented in significant numbers. The largest and most dangerous of the Pleistocene carnivores, the brown bear and the lion, are virtually unrepresented.

Even less common are the bones of birds and fishes. It is possible that the ice-age occupants of the Ukraine lacked the technology to deal with these animals, but other explanations can be imagined. An intriguing one is that the river valleys, where virtually all the known sites are located, were mainly inhabited only in winter, when frozen rivers prevented fishing and the most desirable birds—migratory waterfowl—had flown south. The fierce winter winds may have made life outside the river valleys all but impossible.

In summer the hunters may have moved away to the uplands between the river valleys. These areas were the grazing lands for the great herds of game on whose meat (possibly dried or smoked for winter use) the occupants of the Ukraine depended. The reason such upland camps have remained undiscovered may be that they were so transitory as to have left little behind to catch the archaeologist's eye or that few of the remains of upland camps have been exposed by road builders or quarriers.

Mammoth bones are virtually the hallmark of ice-age sites in the Ukraine. In more than one instance it was the unearthing of these large and distinctive remains that first made unwitting excavators realize they had discovered an ice-age site. The mammoth bones are present in large numbers. For example, those dug up at eight sites in the Dnieper-Desna basin represent a minimum of 500 individual mammoths. Yet not one Ukrainian site contains clear evidence that a mammoth was actually killed there. Indeed, there is evidence to the contrary: chemical tests of the mammoth bones found at one Dnieper site indicate that the animals probably lived and died in different millenniums. Again, many of these same bones, and the bones from another Dnieper site, had been gnawed by carnivores. The implication is clear. Ice-age men in the Ukraine are at least as likely to have gathered the bare bones of long-dead mammoths as to have hunted and killed the huge beasts. But why would they have wanted the bones?

The answer is clear-cut and surprising. At virtually every Ukrainian site where mammoth bones are common most of them are found in patterned arrangements that look like the ruins of shelters, many of them quite ambitious. The use of mammoth bones as a construction material is further indicated by the presence in disproportionate number of particular bones: skulls, tusks, mandibles, pelvises, scapulas and certain limb bones.

One item of animal-bone evidence indicates the adaptability of the ice-age hunters to winter conditions. It is found among the bones of the wolf, the arctic fox and the hare. Skeletons of these animals are often found intact except that the paws are missing. Groups of paw bones, in turn, are found in other parts of the site. By analogy with modern practices the animals appear to have been skinned for the warmth of their pelts.

In summary, then, the animal remains at the Ukrainian ice-age sites seem to have been accumulated by successful Pleistocene hunters whose diet was rich in protein and fat, who valued furs and who built substantial shelters. The shelters were heated too; there are quantities of charcoal, indicative of hearths, and bits of burned bone as well. All of this should have made the ice-age winters easier to bear.

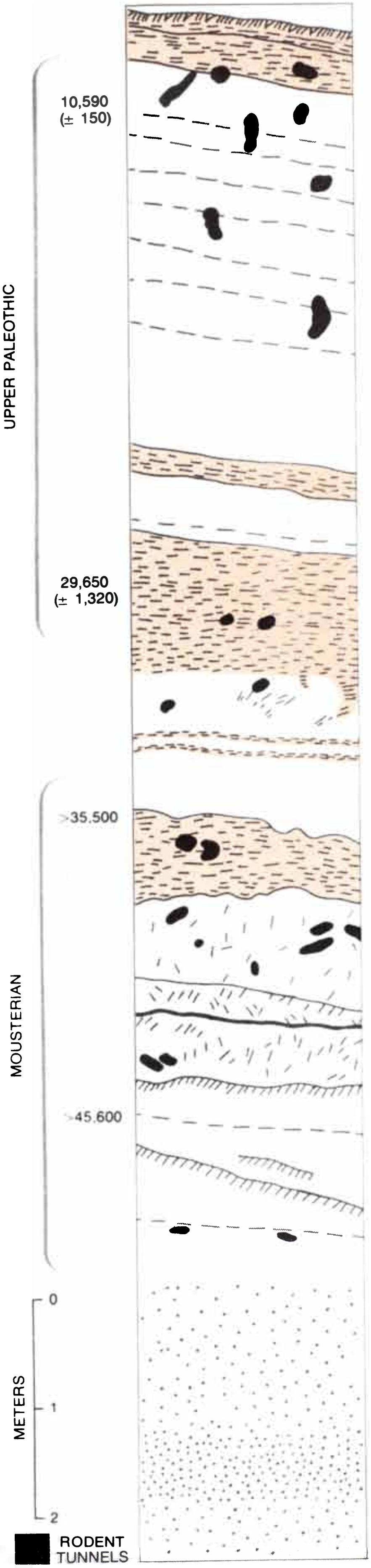
Who were the hunters? The artifacts that have been found at Ukrainian Last Glacial sites have been assigned to two successive cultural units. The earlier of the two is known as the Mousterian; it is also sometimes called the Middle Paleolithic. (Cultural materials belonging to the preceding Lower Paleolithic have not yet been found in the Ukraine.) The later of the two cultural units is known as the Upper Paleolithic. These successive units characterize the Last Glacial all over Europe; the Upper Paleolithic supplants the Mousterian at some time between 45,000 and 35,000 years ago, the exact time depending on the location of the site.

The commonest Mousterian tool is called a sidescraper. (Its actual function, of course, remains a matter of speculation.) These tools are made from large, irregular flakes of stone that were forcibly detached from a "core," usually of flint, and then "retouched" along one edge or more by removing a series of small flakes. Many different kinds of sidescraper have been recognized on the basis of such factors as the number and shape of the retouched edges and the quality and location of the retouching.

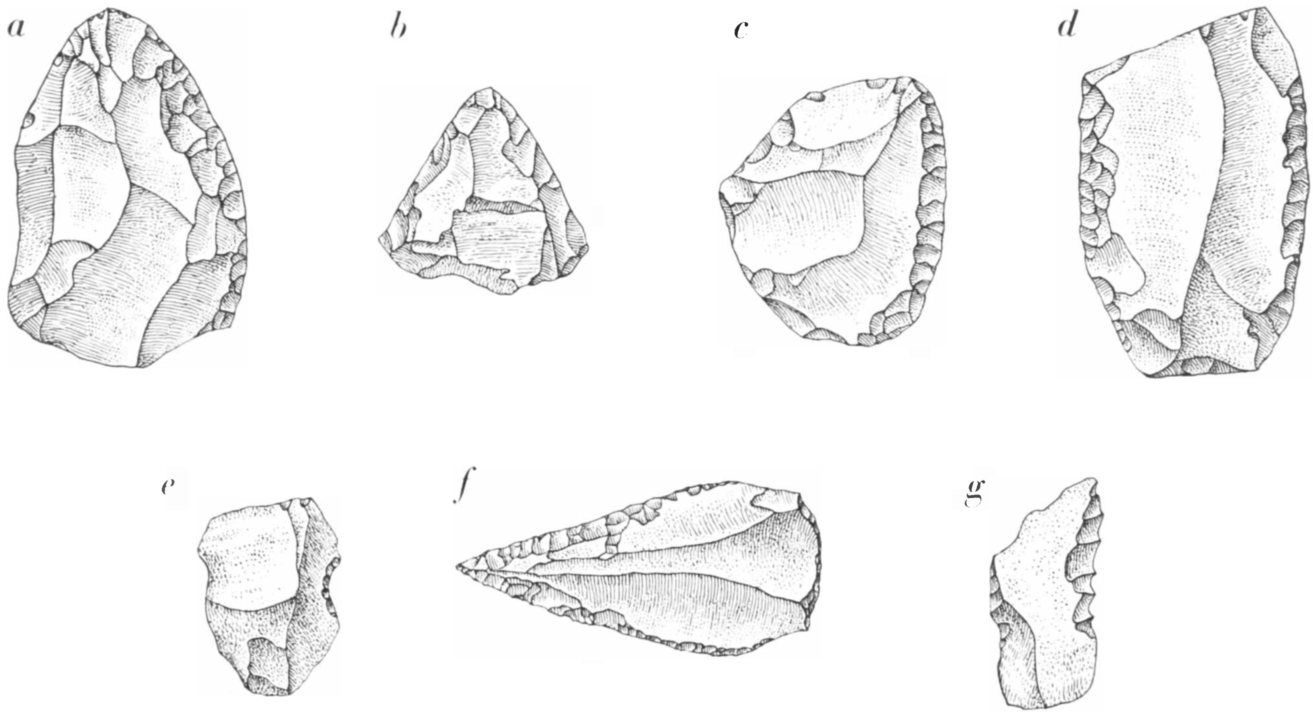
The second most abundant Mousterian implement is a flake with an edge that has been modified so that it has either a single notch or a series of notches on one edge. These are known respectively as notches and denticulates. It has been suggested that they were used as spokeshaves, shredders and saws. Many Mousterian sites also contain roughly leaf-shaped points that are often retouched more or less completely over one face or both [*see top illustration on next page*].

Inventories of the Mousterian stone tools from various Ukrainian sites reveal that they include certain types, for example special varieties of sidescraper, that are not found elsewhere in Europe. The reverse is also true: artifacts such as the small "hand axes" found in some Mousterian assemblages in France are unknown in the Ukraine. Such differences allow us to infer that the Mousterian was not a single cultural unit but rather a complex of subcultures that varied in time and space.

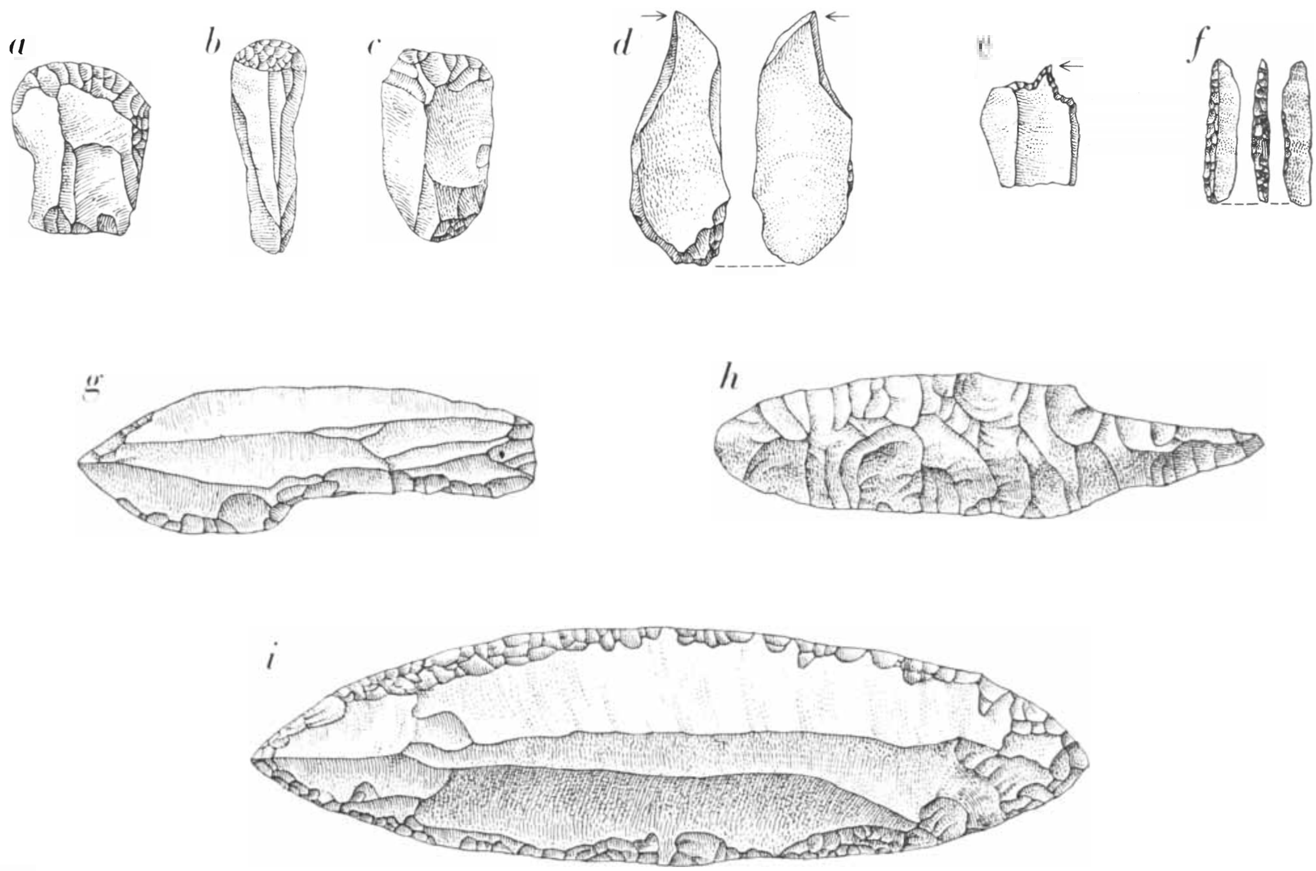
A Mousterian accomplishment that is more fully evident in the European U.S.S.R. than anywhere else in the world is the one implicit in the collections of mammoth bones: the construction of substantial heated dwellings. A. P. Chernysh has excavated a particularly clear-cut example of one of these Mousterian dwellings at Molodova I, a Dniester River site adjacent to Molodova V [*see illustration on page* 97]. The remains consist of an oval of mammoth bones surrounding a floor with the considerable area of some 50 square meters. Within the oval Chernysh and his colleagues uncovered 15 separate hearths, hundreds of animal-bone fragments, some 29,000 pieces of flint and one spot of red pigment. As Chernysh recon-



CROSS SECTION of strata at Molodova V (*right*) shows 19 cultural horizons within the layers of sand and silt accumulated at the base of a slope. The eight lowest horizons contained Mousterian artifacts; carbon-14 analysis suggests they range from more than 35,000 to more than 45,000 years in age. Upper Paleolithic artifacts from the higher horizons range from 10,500 to more than 30,000 years in age. Areas of color show the soils formed during warmer intervals; the stippling shows alluvium of Last Interglacial.



TYPICAL MOUSTERIAN TOOL KIT emphasizes sidescrapers (*a—d*), notches (*e*), points (*f*) and denticulates (*g*). The actual use made of these stone artifacts is not known, but the points may have been hafted to wood shafts and the notches and denticulates could have been used to work wood. The tools illustrated here are from Mousterian sites in western Europe, not from sites in the Ukraine.



UPPER PALEOLITHIC TOOLS, typically made by chipping the desired shape from a narrow flint blade, differ in appearance and in type from Mousterian tools. Shown here are end scrapers (*a—c*), a burin (*d*), a borer (*e*), a backed blade (*f*), a long, leaf-shaped point (*i*) and two shouldered points (*g*, *h*). All are from Russian sites; *h* and *i* are from the lowermost of the Upper Paleolithic horizons at Molodova V. Like the Mousterian tools illustrated at the top of the page, all artifacts are reproduced at two-thirds actual size.

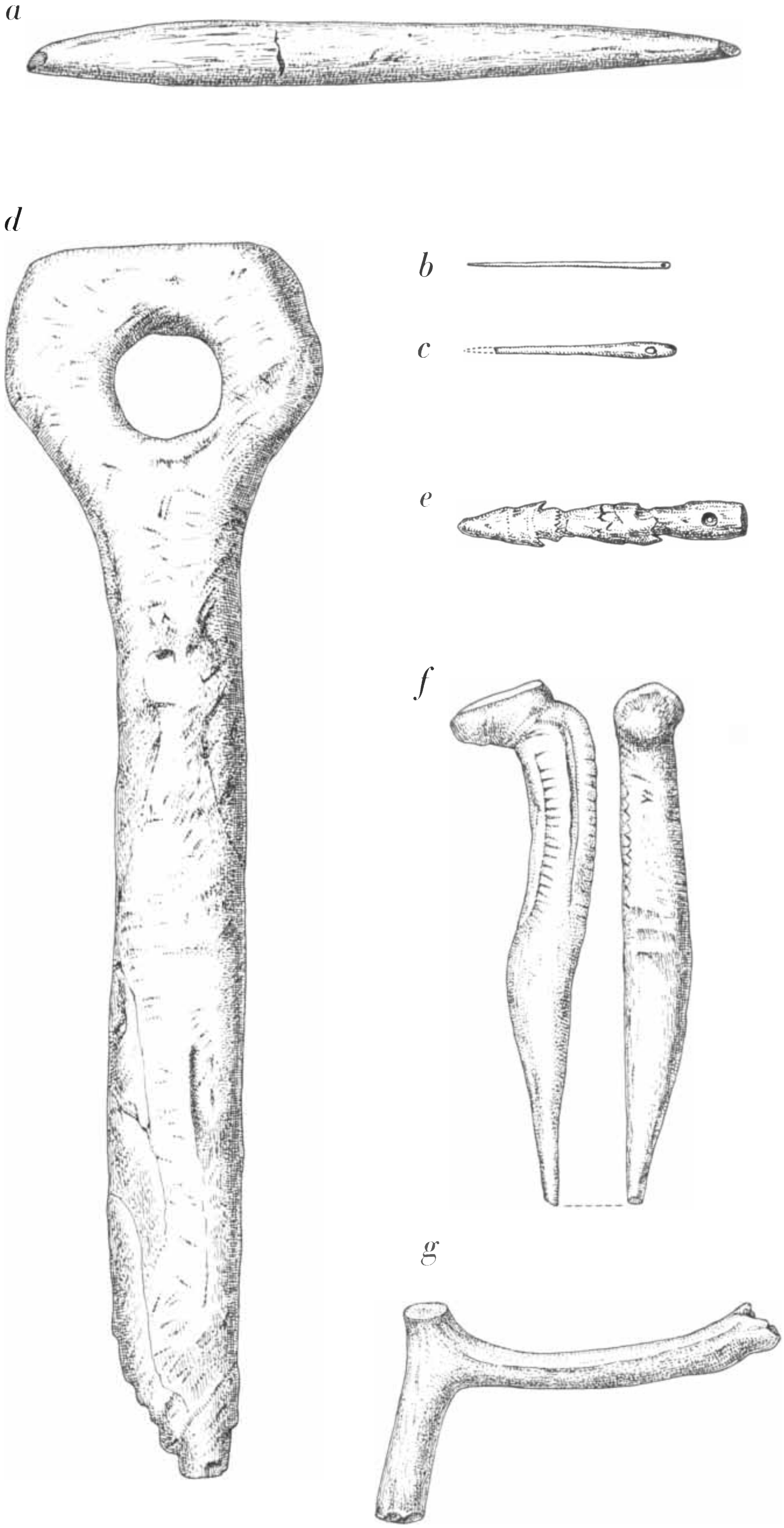
structs the dwelling, it consisted of animal skins stretched over a wood frame; the mammoth bones served as anchors, securing the edges of the skins at ground level and possibly holding the skins in place elsewhere.

In other parts of Europe and in western Asia the people who made Mousterian tools buried their dead, at least on occasion. No such burials have been found in the Ukraine, although they are known in immediately adjacent areas such as the Crimea. In Europe wherever human remains have been found in association with Mousterian artifacts they invariably belong to the anatomically primitive human subspecies we call Neanderthal man (*Homo sapiens neanderthalensis*). As a result it is now generally accepted that Neanderthals were the makers of Mousterian tools.

Tool assemblages of the Upper Paleolithic kind found at Ukrainian sites, like their counterparts elsewhere, differ sharply from Mousterian assemblages. To make their tools Upper Paleolithic hunters preferred flakes that were at least twice as long as they were wide. Archaeologists call them blades. Instead of consisting primarily of sidescrapers, notches and denticulates, Upper Paleolithic assemblages include end scrapers (blades with one or both narrow ends retouched), burins (blades modified to leave a chisellike corner), backed blades (with one edge deliberately dulled) and borers (blades modified to leave a sharp central protrusion). The assemblages also include different kinds of points, some with one side or both narrowed at one end as if to facilitate hafting, and some retouched over both faces [*see bottom illustration on opposite page*]. Stone tools of the same kind are also found at Mousterian sites, but they are rare or less well formed or both.

Upper Paleolithic assemblages also contain numerous artifacts made from bone and antler; these are virtually unknown at Mousterian sites. The shapes of the objects bring to mind projectile points, hide-burnishers, shaft-straighteners, awls and even needles. The putative needles are particularly interesting. If they were indeed used for sewing, it is highly likely that the Upper Paleolithic hunters of the Ukraine wore fitted skin garments. Outside the Ukraine this conclusion recently became all but inescapable. Three ice-age human skeletons have been excavated at Sungir', northeast of Moscow. The skeletons were girdled with strings of beads that must have been sewn on close-fitting clothing.

One trait that the Upper Paleolithic hunters of the Ukraine shared with their

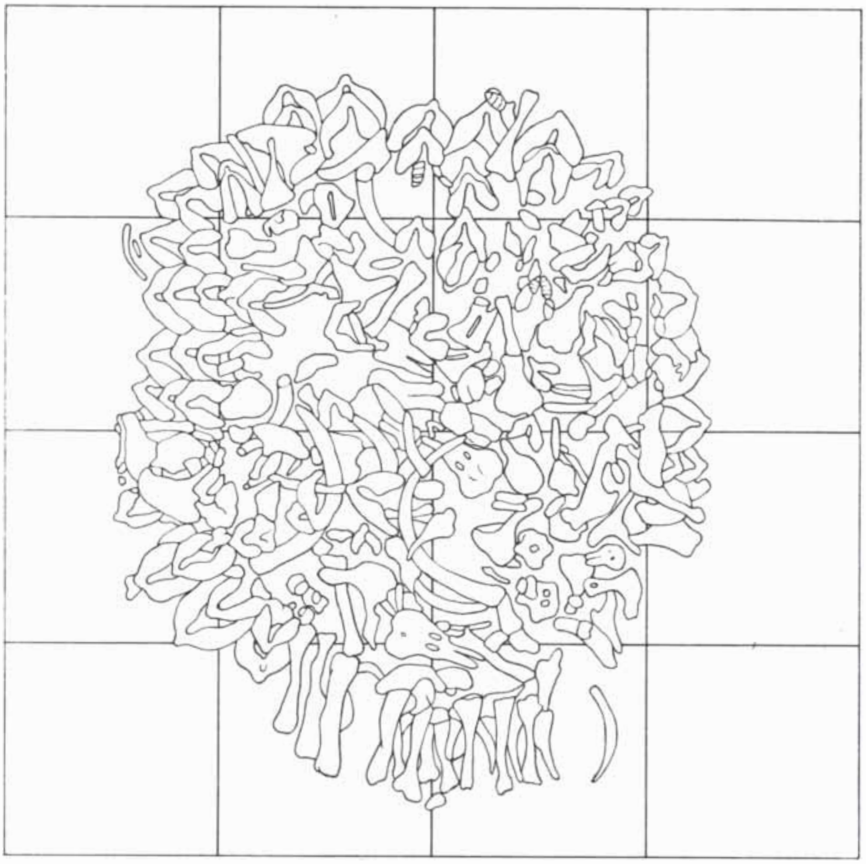


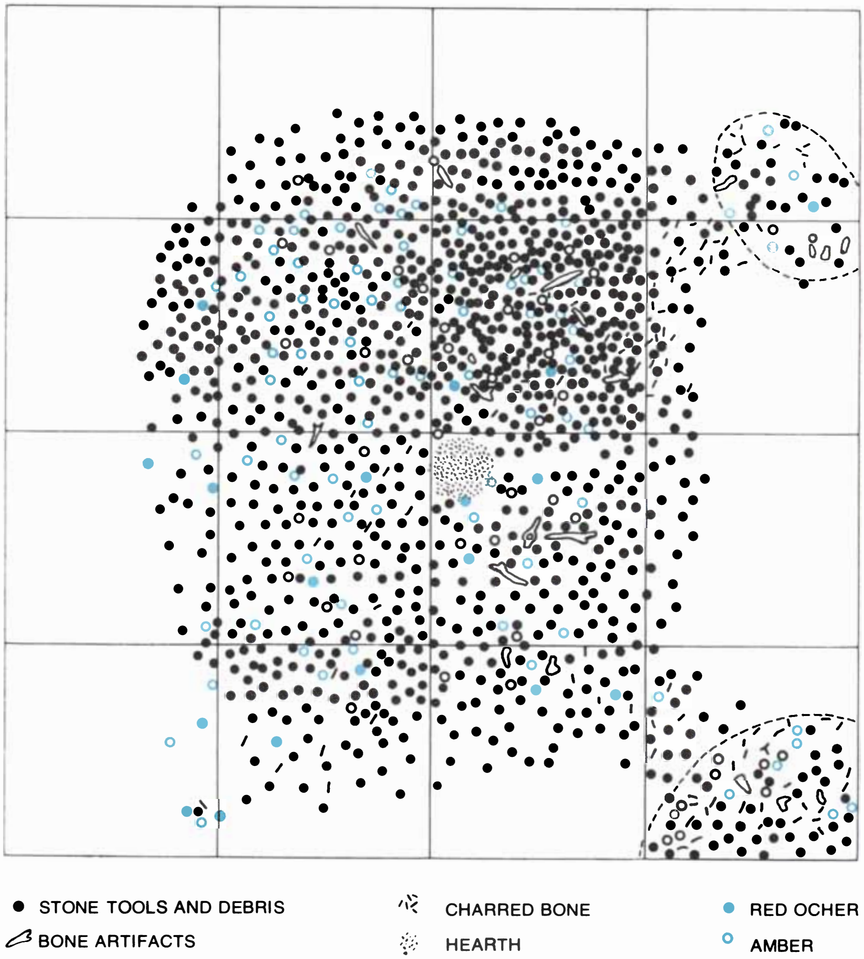
TOOLS MADE FROM BONE are found at Upper Paleolithic sites in the Ukraine but rarely at Mousterian sites there or elsewhere in Europe. Illustrated are a projectile point (*a*), two needles (*b*, *c*), a shaft-straightener (*d*), a harpoon head (*e*), an awl (*f*) and a hammer made from reindeer antler (*g*). The harpoon head is from a horizon near the surface at Molodova V and the shaft-straightener is from an earlier Upper Paleolithic horizon there.

predecessors was the construction of heated dwellings, supported by wood frames and weighted with mammoth bones. In one of the Upper Paleolithic levels at Molodova V even the holes that accommodated the shelter's wood uprights are recognizable. The most spectacular traces of these dwellings are at sites in the Dnieper-Desna basin; there large patterned arrangements of mammoth bones sometimes surround or cover man-made hollows in the ground. In some instances reindeer antlers evidently served in lieu of mammoth bones. The dwelling floors are pockmarked with pits of various shapes and sizes that may have been used for storage or the disposal of refuse.

Evidence from elsewhere in Europe shows that, like the Mousterians, Upper Paleolithic peoples often buried their dead. No such graves, however, have been found in the Ukraine. Where burials have been found in other parts of Europe, including sites in areas adjacent to the Ukraine such as Sungir', the bones are those of the anatomically modern human subspecies *Homo sapiens sapiens*. Evidently at some time during the interval between 45,000 and 35,000 years ago the more primitive human subspecies left the Old World stage, so to speak, and the modern subspecies entered the limelight.

What actually happened to the Neanderthals remains more than a little uncertain. The archaeological record shows that, in terms of material culture as represented by tools and other artifacts, the Mousterian industries were replaced by Upper Paleolithic ones. Does this mean that the Neanderthals were literally shouldered offstage by modern man? Such a conclusion is unacceptable to many authorities, who argue that Neanderthals repeatedly evolved into modern men over a broad front at the same time that Mousterian industries evolved into Upper Paleolithic ones. In the past few years, however, carbon-14 dates





SPECTACULAR RUIN of an Upper Paleolithic shelter, discovered at Mezhirich, a site on the Dnieper south of Kiev, included 385 mammoth bones that covered an oval area some three meters in diameter (*top illustration at left*). When the bones were removed, the excavators uncovered 4,600 flint artifacts, along with nodules of flint, bone artifacts and bits of charred bone, bits of red ocher, pieces of amber and a pit, 20 centimeters deep, filled with ash and charcoal. Two other hearth areas lay beyond the perimeter of the shelter (*bottom illustration*).

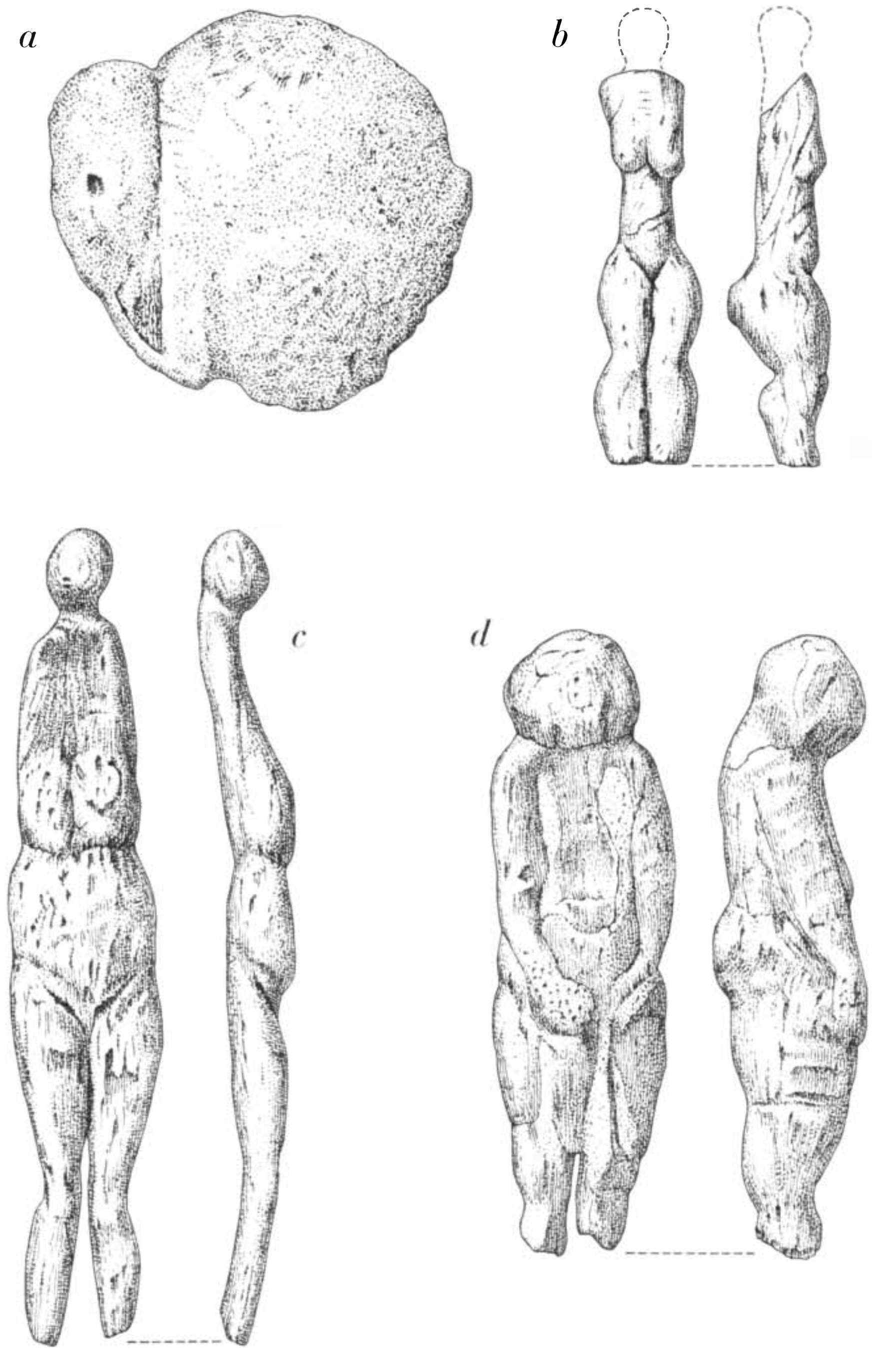
obtained from such sites as Istállóskö Cave in Hungary, taken in conjunction with evidence of where it was that anatomically modern man evolved, have added weight to the hypothesis that modern man and the Upper Paleolithic cultural unit made their first appearance within a geographically narrow region: southwestern Asia and southeastern Europe. On that hypothesis both modern man and Upper Paleolithic material culture, rather than evolving repeatedly in numerous places, radiated outward from this initial heartland.

Regardless of how or where the Upper Paleolithic came into being, the archaeological evidence from the Ukraine and elsewhere in the U.S.S.R. undeniably indicates that it was superior to the Mousterian as an adaptation to Last Glacial climate. This conclusion is supported by the distribution of Mousterian and Upper Paleolithic sites. Systematic reconnaissance in those northern parts of the U.S.S.R. where glaciation has not obliterated all evidence of ice-age man has failed to locate even one Mousterian site north of a latitude of 54 degrees. Upper Paleolithic sites, however, have been found even above the Arctic Circle. The much greater number of Upper Paleolithic sites, particularly ones with "ruins" suggesting a settled or semisettled existence, implies higher overall population numbers and greater population density for Upper Paleolithic peoples.

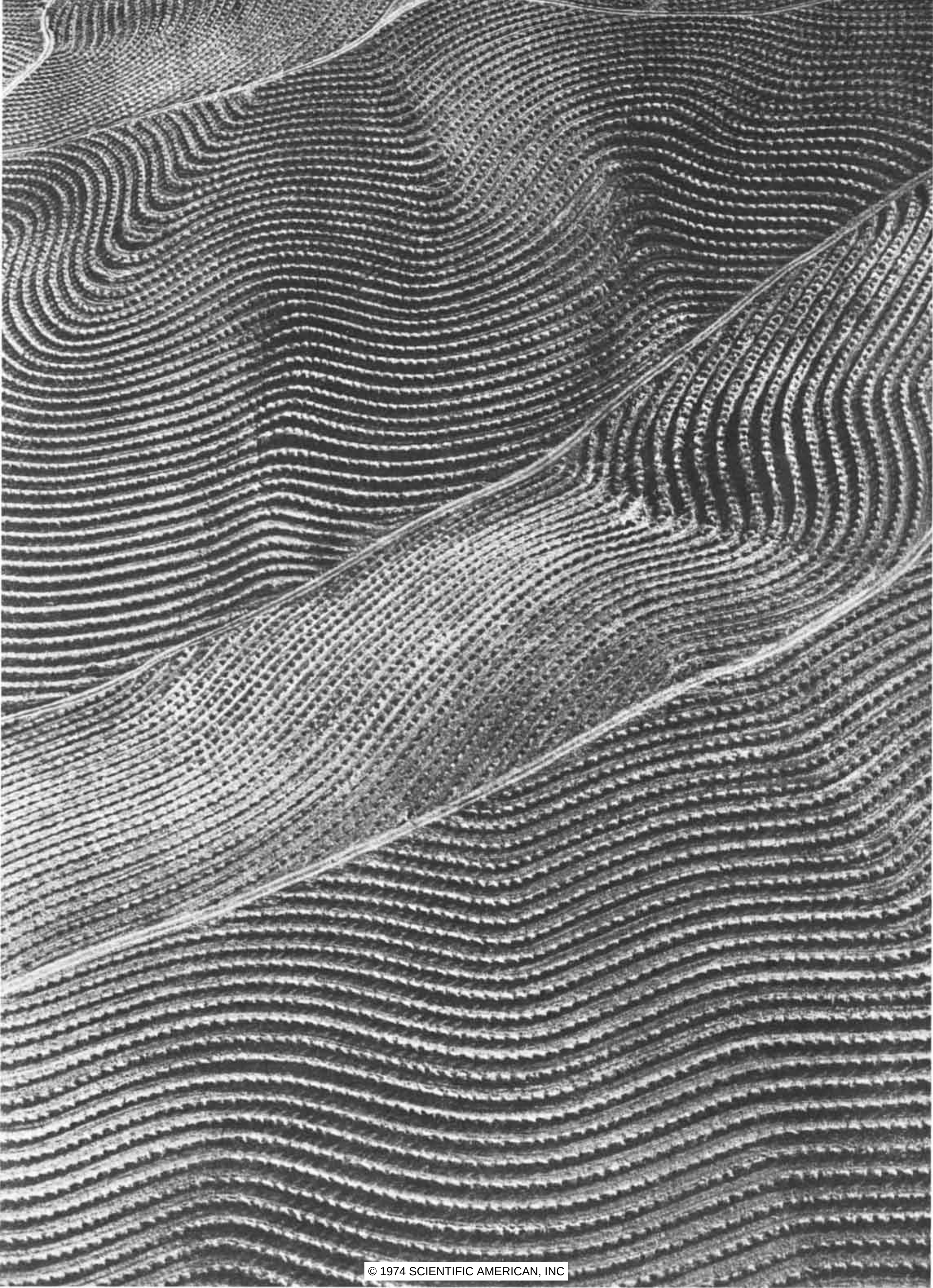
At this point still another type of evidence needs to be considered. Neither here nor elsewhere in Europe have any Mousterian sites yielded undoubted works of art or decoration. Upper Paleolithic sites in the Ukraine and elsewhere, however, contain numerous figurines of humans and animals made of bone and stone, together with bracelets, beads, pierced shells and teeth and hundreds of bone objects engraved with geometric patterns [*see illustration at right*]. Can the absence of such objects from one material culture and their abundance in a succeeding one be a token of biological change, specifically the evolution of a greater intellectual capacity?

If such a hypothesis is accepted, then the greater adaptive success of the Upper Paleolithic is an example of a feedback interaction between biological change on the one hand and material and sociocultural change on the other. Among other developments, new social structures may now have arisen to allow the integration of larger, denser populations. For example, social innovations may have facilitated intergroup cooperation in the hunting of large mammals. If more meat was available, the hunting economy would support more people. More subtly, new social structures may simply have encouraged greater intergroup food-sharing as a means of mitigating temporary local shortages. If the haves in any one year share with the have-nots, the long-term effect of such an act is a greater total number of both haves and have-nots.

However this may have been, it is reasonable to draw demographic inferences from a numerical comparison of Mousterian and Upper Paleolithic sites. When these inferences are taken together with the differences between the two material cultures, particularly with respect to art objects, one is led to conclude that the Upper Paleolithic represents a quantum advance in human cultural evolution.



WORKS OF ART from Upper Paleolithic sites in Russia include "Venus" figurines and a representation of a mammoth. The mammoth (*a*) and two of the figurines (*c, d*) are from Avdeevo, on the Sejm 400 kilometers south of Moscow. The headless figurine (*b*) is from Eliseevichi, on the Sudost' some 350 kilometers southwest of Moscow. Whereas objects like these and many geometrically engraved bones are found at Upper Paleolithic sites in Russia, neither there nor elsewhere in Europe have such artifacts been found in Mousterian sites.



Wines, Grape Vines and Climate

Why are wines so different? An important part of the answer is that the more rigorous (and thus more variable) grape-growing conditions give rise to the more idiosyncratic (and thus more intriguing) wines

by Philip Wagner

Until recently wine was a rather exotic drink for most Americans. Although champagne was traditional at weddings, wine as a regular part of alimentation was mostly limited to families that either were well-to-do and widely traveled or were of Mediterranean or central European origin and had kept the customs of the homeland. Within the past decade the picture has begun to change. Wine has been discovered in the U.S. The increase in wine drinking has been widespread, steady and only partly faddish, and it shows no sign of abating.

The general adoption of wine is so recent that many of the people who drink it are still on uncertain ground in attempting to tell a good wine from a bad one and a great wine from a good one. There is much curiosity but little understanding. To many of the recent converts it is a mystery why the experienced wine drinker calls some wines merely ordinary and yet keeps drinking them with satisfaction but then will become suddenly attentive and respectful in tasting others.

Judgments on wine are always somewhat subjective, in the sense that almost everyone has favorites. Underlying such preferences, however, is a definite and objective base for making distinctions. In what follows I shall try to lay the groundwork for some of the distinctions. Anyone who knows the principles and also has a sound sensory system and a good memory will find that wine can mean much more. Doors are opened. It is the same with music, painting and literature.

Let us consider an experienced wine drinker who does not care much for wines with a sweet finish. That is the subjective part of it. In spite of this preference (or prejudice) he can still recognize one of the great Sauternes or German Trockenbeeren as something very special: a product of the unique conjunction of certain grapes matured under certain conditions and of special and exacting methods of preparation. He also recognizes the differences between that wine and one merely fashioned to resemble it. This is what I mean by an objective distinction.

At the outset I should make clear that in discussing wine I am referring only to the fermented juice of the grape, not to that of strawberries, apples, rhubarb or any other plant. The various kinds of grape all carry the genus name *Vitis,* the Latin word for vine. Embraced by the genus are some dozens of grape species that grow in various parts of the Northern Hemisphere. Most of them are valueless for wine, although a few have other virtues that are important to contemporary viticulture. Wine that meets European standards is made from a single species, the Eurasian grape *V. vinifera,* plus a few hybrids between *V. vinifera* and other species that yield fruit of *vinifera* character.

V. vinifera breaks down into numerous groupings, which are often so different that the question is sometimes raised of how pure a species it really is. If one takes the definition of a species as an absolute rather than as a mere human convenience, then a good deal of miscegenation has taken place during the evolution of wine grapes. For the present, however, it is sufficient to say that the groups and individuals embraced by the name *V. vinifera* are all nominally of the same grab-bag species and that their differences are accounted for by nature's winnowing of the weaklings and man's choice of the best. Taken all together, they account for wines as utterly different as retsina and champagne or port and Beaujolais.

The *V. vinifera* varieties do all have certain common characteristics. They are deciduous. They lack resistance to certain parasites, notably plant lice of the phylloxera type and nematode worms. They are extremely susceptible to various fungus diseases. They require a good deal of heat and sunlight to ripen properly. Their hardiness under winter conditions is limited, so that their survival becomes doubtful where the January isotherm falls much below one degree Celsius (33.8 degrees Fahrenheit). The parts of the world where they will thrive and produce wine are restricted, although the total area is large. Growing wine grapes of the *vinifera* family is as impractical in the Congo and Amazon basins or in Saskatchewan and the Siberian steppes as growing pineapples in Maryland. Climate has the last word.

To illustrate the dominance of climate I should like to draw a meandering line from west to east on a map of Europe [*see illustration on next two pages*]. Europe remains the place where most wine grapes are grown. The line begins in Brittany, north of the Loire River, and proceeds eastward until it turns sharply northeast past Chartres to swing around

TERRACED VINEYARD recently established in Napa, Calif., appears in the aerial photograph on the opposite page. The region has a climate that is a special variant of the two-season Mediterranean climate with its cool, wet winters and dry summers. The Napa Valley benefits greatly from its proximity to the San Francisco Bay and from the influence of cool Pacific winds flowing in through the Golden Gate. It also benefits to some extent from its altitude. These factors have established it as California's premier wine-producing region.

Paris. (There is a small vineyard on Montmartre.) Moving eastward again, the line passes to the north of Reims and the Champagne district, turns slightly southeast toward Nancy in Lorraine and then bends sharply north to follow the Moselle River along the edge of Luxembourg and thence on down the better-known German stretch of the Moselle to Koblenz.

There it turns roughly southeast (it is in fact quite twisty for some distance), takes in a part of Czechoslovakia and then at the Oder River heads abruptly south, almost to the outskirts of Vienna. (Over the Alps, where grapes do not grow, the line is broken.) Just short of Vienna there is another sharp bend, this time to the east. Now it passes north of the plain of Hungary, leaps the Carpathians, crosses the Dnieper and Don well north of the Black Sea, passes through Soviet Georgia north of the Caucasus range, crosses the Caspian Sea and is eventually lost in the wilds of Asian Kazakhstan.

This line describes the northern boundary of culture for the Eurasian grape. It is established by two limiting factors. The western end is under maritime influence, which tempers the winters but in this high latitude keeps the summers much too cool, so that north of the line the vine does not ripen its fruit satisfactorily. Farther east continental influences take over, and by the time the line reaches the Oder winter weather becomes deadly.

One must not view the line as being absolute in the sense that no grapes can be grown north of it. Various early-ripening and hardy crosses with other species will come through and yield wine that is acceptable if nothing more. Such grapes are grown in favored places in southern England and even in Moscow, where the permanent agricultural fair includes a small vineyard. In the Moscow vineyard, however, the vines are ultrahardy Michurin hybrids of no value for wine, and even so they must be buried for the winter in order to survive. The same would be true in the colder parts of temperate North America. In sum, the Eurasian grape can be grown reliably only in a belt of varying width that lies between the line of northern limit and a line embracing the southern littoral of the Mediterranean (and, of course, under analogous conditions in some other parts of the world, notably California, Chile-Argentina, South Africa and southern Australia).

Having given the broad limits, I should like to draw another line. It too runs from west to east across Europe. This second line is of fundamental importance to wine drinkers who are trying to get their bearings.

MEDITERRANEAN CLIMATE

SPAIN
1. SHERRY
2. MALAGA
3. ALICANTE
4. TARRAGONA
5. VALDEPENAS
6. ZARAGOZA
7. BAJA RIOJA

FRANCE
8. ROUSSILON
9. CORBIÈRES
10. MINERVOIS
11. LANGUEDOC
12. CÔTES DE PROVENCE
13. CÔTES DU RHÔNE
14. GAILLAC

ITALY
15. PIEDMONT
16. TUSCANY
17. ALPINE
18. ORVIETO
19. VERDICCHIO
20. VENEZIA
21. NAPLES

YUGOSLAVIA
22. DALMATIAN
23. SERBIAN

U.S.S.R.
24. CRIMEA
25. GEORGIA

TEMPERATE CLIMATE

SPAIN
26. ALTA RIOJA

FRANCE
27. JURANÇON
28. ARMAGNAC
29. BORDELAIS
 a. MEDOC
 b. GRAVES
 c. SAUTERNES
 d. ST. ÉMILION
30. MONBAZILLAC
31. CAHORS
32. AUVERGNE
33. COGNAC
34. VINS DE LA LOIRE
 a. MUSCADET
 b. ANJOU
 c. VOUVRAY
 d. POUILLY
35. CHAMPAGNE
36. BURGUNDY
 a. CHABLIS
 b. NUITS
 c. BEAUNE
 d. MONTRACHET
37. BEAUJOLAIS
38. SAVOIE
39. JURA
40. ALSACE

GERMANY
41. MOSELLE
42. RHINELAND
43. FRANCONIA
 a. BADEN

SWITZERLAND
44. NEUCHATEL
45. LÉMAN
46. VALAIS

AUSTRIA
47. DANUBIAN
48. BURGENLAND

YUGOSLAVIA
49. SLOVENIAN
50. CROATIAN

HUNGARY
51. TOKAY

U.S.S.R.
52. MOLDAVIA
53. UKRAINE

PORTUGAL
54. PORT
55. VINHO VERDE

WINE-PRODUCING REGIONS of Europe and the region around the Mediterranean are shown on a map that is distinguished by two lines running predominantly from west to east. The upper line marks the northern limit of cultivation for wine grapes. The lower line divides the area in which wine grapes are under the influence of the Mediterranean climate from the area in which temperate-

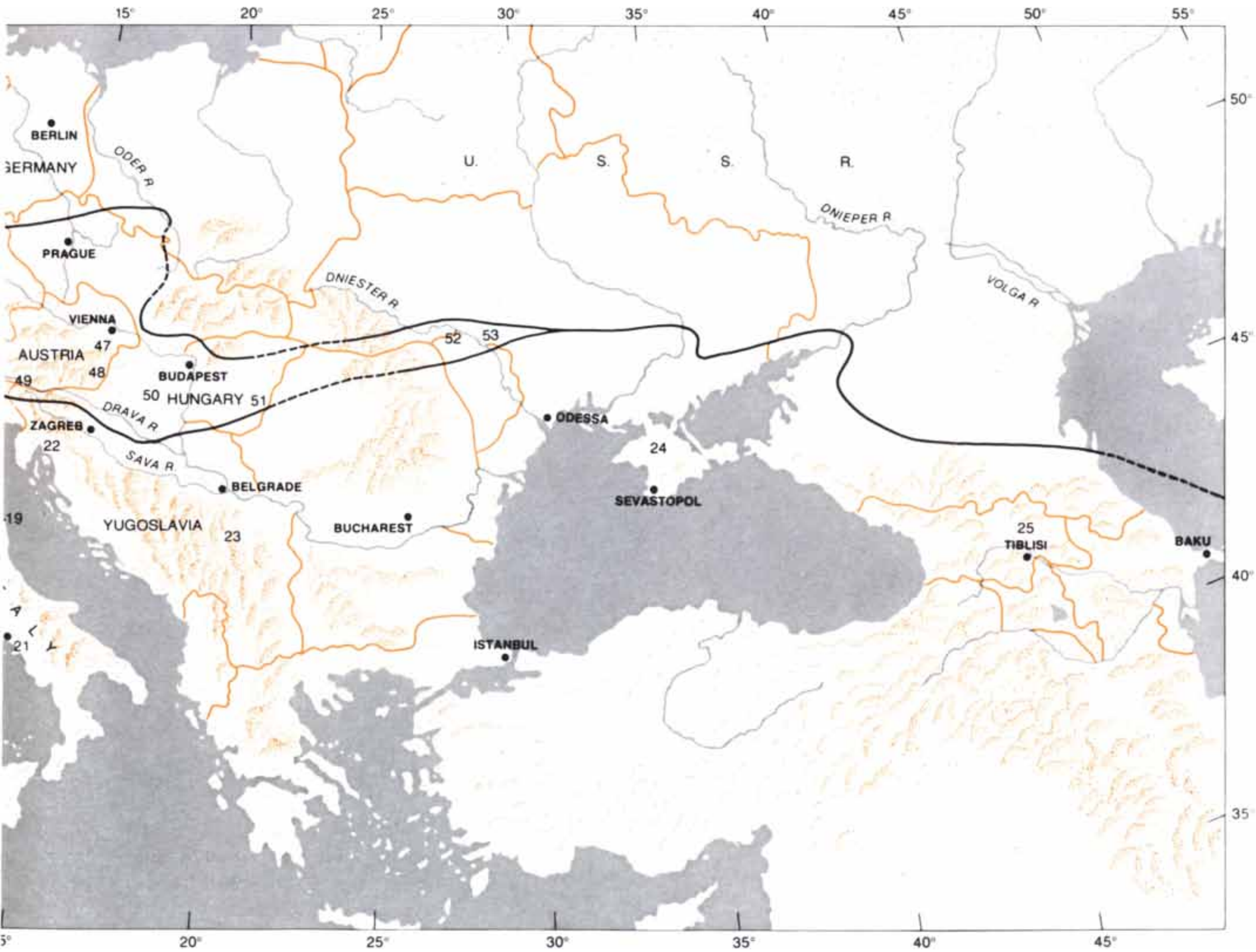
The line divides Europe into two distinct viticultural areas. It can be drawn with considerable precision (more than I have attempted in the accompanying map), although it meanders a great deal and local arguments about the bench marks can be raised all along its course. Oddly enough, this line is never mentioned in the customary reference books, fundamental though it is in the way it separates the wines of Europe into two groups.

The line begins in Portugal south of Oporto and follows the port-producing valley of the Douro River uphill to the east. Once over the spine of the Iberian Peninsula it turns north again, bisecting the Spanish La Rioja district between Haro and Logroño. Making another jump, this time over the Pyrenees, it resumes in the Basque country of southwestern France, heading northeast toward Cahors on the Lot River and bisecting the Garonne River near the point where its tributary the Tarn enters.

Cahors, which has ancient associations with downstream Bordeaux, is a pivotal point. From there the line zigzags up and over the mountainous Centre, swings northeast at the headwaters of the Allier and the Loire, skirting the west side of the Cevennes Mountains, and finally comes out at Lyons on the big bend where the Saône (the Burgundy-Beaujolais river) joins the Rhône from the north. Here it heads up the twisting Rhône valley to Lake Geneva. Beyond that it falls into a confusion best indicated by dashes as it negotiates the Alps. The finger lakes in northern Italy raise a question; probably the line should pass to the north of them, but the point is arguable.

Beyond that the course is clear again. Running well north of the Venetian plain, it proceeds into Yugoslavia, cutting off northern Slovenia around Maribor (the detached corner of what was once an exceptionally fine wine-producing region of Austria), and then moves east along the Drava River. Following the Drava for a while, it skirts the southern and then the eastern border of the plain of Hungary, cuts over the Carpathian Mountains, crosses Moldavia and eventually joins the other (northern limit) line somewhere well above the Black Sea.

This line defines from the vintner's point of view the northern limit of Mediterranean influence. The Mediterranean basin has a two-season climate with mild and rainy winters and hot, dry (practically rainless) summers. It has a characteristic flora, of which the olive and the cork oak, citrus trees and melon plants, the oleander and a host of succulent plants are examples. It also has its own



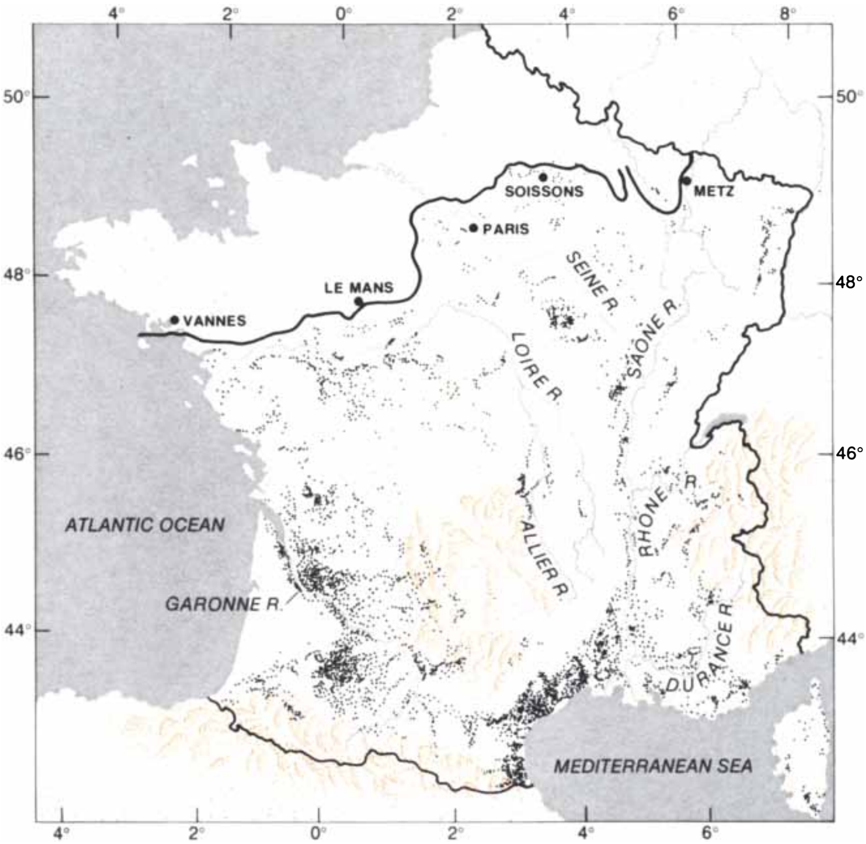
climate conditions are dominant. The numbers indicate notable regions producing Mediterranean wine and temperate-climate wines. The same numbers are related to the key on the left side of the map, where the wine-growing regions are named. A number of cities are identified on the map as points of reference. The countries are the ones that have been cited in the map key.

distinctive group of grape vines, all with pedigrees going back to the dawn of history.

As one might expect, the Mediterranean climate and the common parentage of its vines confer a strong family resemblance on all Mediterranean wines. As a rule they are bland, soft, low in acidity, relatively high in alcohol and not fruity. They have a somewhat coarse and earthy but not disagreeable background in their flavor. Most of them have little or no aroma, but a few (of the Muscat group) are highly aromatic. The differences between the red wines of, say, coastal Spain, Algeria, Tunis, Languedoc, Sicily, Dalmatian Yugoslavia and Greece are minor. The wines are basically much the same. The northern boundary for wines of this style is marked by the second of the lines I have drawn. Again the reader should bear in mind that the line is not absolute and that one can give or take a few miles each way along its course.

North of this line it is a quite different story and a different world. Here, in a narrow and irregular band stretching across Europe, is the four-season temperate climate in all the variations that will support the grape, with cold winters, highly variable temperatures during the growing season and rainfall scattered throughout the four seasons. The soils are different too, because (even allowing for gross differences of geological origin) soil tends to be a function of climate. Soil is "weathered." In this region also the plant life is distinct from that of the Mediterranean, although with a certain amount of overlapping.

And so are the grapes, which look different, taste different and produce different wines. In general the wines are fruitier, lower in alcohol and higher in original acidity than the Mediterranean wines. They also have flavors that are subtler and more delicate. Although the Mediterranean basin has its superior wines (the red Chiantis, certain wines of the Piedmont and of the Italian finger lakes, the best wines of the Rhône valley and, of course, such fabricated specialties as sherry and Marsala), there are relatively few of them. The consensus among experts is that most of the truly fine wines (the burgundies, clarets and champagnes and the wines of Germany and Alsace, not to mention some that are less familiar in the U.S., such as those of Austria and Hungary) are temperate-climate wines.



MEDITERRANEAN ORIGIN		TEMPERATE-CLIMATE ORIGIN	
GRAPE	ACRES	GRAPE	ACRES
CARIGNANE	456,950	GAMAYS	101,270
ARAMON	395,200	SÉMILLON	74,100
UGNI BLANC	118,560	MALBEC	41,990
GRAND NOIR	98,800	GAMAYS TEINTURIERS	39,520
ALICANTE BOUSCHET	86,450	CABERNET FRANC	34,580
GRENACHE	79,040	FOLLE BLANCHE	29,640
CINSAUT	39,520	SAUVIGNON	27,170
TERRETS	37,050	COLOMBARD	27,170
CLAIRETTES	34,580	CHÉNIN BLANC	27,170
MORRASTEL	27,170	MUSCADET	27,170
MAUZACS	16,055	JURANCON	24,700
VALDIGUIÉ	12,350	GROLLEAU	24,700
SYRAH	8,151	PINOTS	19,760

FRENCH WINE REGIONS are indicated by the black dots on this map. The table at bottom lists the most important varieties of grape according to the number of acres planted to vines and according to whether the grapes yield red wine (*color*) or white (*black*).

The general nature of the wines produced on each side of the line is determined, then, by a combination of the prevailing climate and the adapted grapes and, to a lesser extent, the soil. The refinements of wine quality within the two broad groupings are determined by microclimates and microadaptations of the grapes. The Burgundy region offers a good illustration. The great red and white burgundies are grown on the sunny and protected southeastern and southern slopes of the Côte d'Or from certain subvarieties of Pinot Noir and Chardonnay grapes. In the valley of the Saône a few hundred yards to the east the microclimate is different, different grapes are grown and the wines do not qualify as great. The same is true up and over the hills to the west.

Contrary to the general belief, the soil, although it is important, is distinctly secondary to the factors of climate and grape. This relation is evident in the fact that numerous quite different soil series are found in many of the world's most famous vineyards.

The subtlety of the close relation between climate and grape can be illustrated in another way. If the Pinot Noir of Burgundy is transplanted to another part of the same general climatic area, such as the Champagne district or the Mâconnais or the Médoc or the Loire valley, the resulting wine may bear a degree of resemblance, but it will be distinctly different. If the grape is planted across the line in the Mediterranean basin, the specific quality of the Pinot Noir is lost and the result is just another not too distinguished Mediterranean red wine. The same kind of thing happens with other famous temperate-climate grapes such as Sauvignon Blanc, Riesling, Chardonnay and Cabernet. Conversely, if Mediterranean grapes such as Grenache or Carignane are transplanted to the temperate-climate side of the line, they languish and may not even ripen.

This fitting of the right wine grapes to a given climate or microclimate is the proper study of ampelography, which draws on botany to begin with and then on climatology, geology and human history. Let us explore some of these factors.

Going back to the fossil record, traces of *Vitis* (of species now extinct) have been found throughout Europe, including even Scandinavia. By a long, slow process of migration and adaptation two other related groups of species evolved in the Far East and North America. The connection is clear because all species of *Vitis* except one have the same number of chromosomes (38) and interbreed. (The Muscadines of the southeastern U.S. are the exception: they have 40 chromosomes and do not interbreed.) The East Asian and North American groups followed a divergent course and in their pure state are of no interest for wine making.

The evolution of the Eurasian grapes was drastically interrupted by the abrupt changes in climate associated with the glacial period. Each glacial thrust compelled a retreat of the vine southward into one or another of two refuge areas, one embracing the Mediterranean basin and the other south of the Caspian Sea. Each glacial recession was accompanied by a return of wild vines in modified forms. From them a new species (*V. sylvestris,* which is also classified as *V. vinifera ss. sylvestris*) eventually emerged, adapting itself to the many climatic variations and so evolving into a number of distinct groups. It is known from recovered seeds

MEDITERRANEAN ORIGIN		TEMPERATE-CLIMATE ORIGIN	
GRAPE	ACRES	GRAPE	ACRES
THOMPSON SEEDLESS	230,928	FRENCH COLOMBARD	23,878
CARIGNANE	28,779	CHENIN BLANC	14,428
ZINFANDEL	23,786	CABERNET SAUVIGNON	11,486
GRENACHE	16,484	PINOT NOIR	4,985
BARBERA	14,744	CHARDONNAY	4,577
MUSCAT	12,541	WHITE RIESLING	3,128
RUBY CABERNET	11,463	GAMAY BEAUJOLAIS	2,754
PETITE SYRAH	8,236	GAMAY	2,552
ALICANTE BOUSCHET	6,901	SEMILLON	2,489
RUBIRED	6 836	SAUVIGNON BLANC	1,658
MISSION	6,372	SYLVANER	1,424
PALOMINO	6,288	GEWURTZTRAMINER	1,159
SALVADOR	2,677	MERLOT	1,014

WINE REGIONS OF CALIFORNIA are designated by initials that, reading from the top, stand respectively for Sacramento Valley, North Coast, Central Valley, San Joaquin Valley, Salinas River Valley, South Coast and Hot Desert. The table at bottom lists principal varieties of wine grape according to whether they produce red wine (*color*) or white (*black*). Viticultural California has a two-season Mediterranean climate, modified by the fact that the great inner valley is almost mountain-locked, which curtails rainfall and intensifies heat.

STATION	MEDITERRANEAN CLIMATE											
	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.
VALENCIA	**10.4** 30.7	**10.8** 31.1	**13.4** 31.3	**14.5** 33	**18** 28.9	**21.4** 35.4	**24** 8	**24.6** 21.4	**22.6** 49.8	**18** 106.4	**14.2** 57.4	**11.8** 30.9
MARSEILLES	**5.7** 38.9	**6.4** 41.3	**10.2** 42.6	**12.7** 33.8	**16.9** 42.7	**20.8** 34.3	**23.3** 15.1	**22.6** 30.2	**19.8** 56.5	**14.8** 80.6	**10.2** 59	**7.9** 72.6
NAPLES	**6.4** 69	**9.9** 64	**10.3** 38	**13.8** 59	**17.4** 59	**21.4** 21	**23.6** 16	**23.8** 27	**21** 92	**16.9** 118	**13** 129	**9.9** 136
ATHENS	**10.6** 58.8	**11.5** 25.5	**12.3** 32.1	**16.3** 19.1	**20.9** 29	**25.7** 10.8	**28.6** 2.7	**28.9** 6.2	**24.7** 10.6	**19.8** 68.7	**15.6** 64.9	**12.3** 52.4
SALONIKA	**6.1** 42	**7.5** 29.8	**9.4** 32.8	**14.3** 34.5	**19.6** 44.7	**24.8** 30	**27.4** 18.5	**27.4** 16.9	**22.5** 30.6	**16.5** 62.9	**11.7** 61.2	**8.2** 49.7
ODESSA	**−4.1** 15.7	**−1.2** 21.2	**2** 22.3	**9.4** 24.1	**16.1** 33	**20.2** 38.6	**22.1** 41.6	**21.4** 42.3	**17.4** 10.1	**10.4** 28.6	**4.8** 37.4	**−.2** 35
SACRAMENTO (CALIF.)	**8.4** 112.7	**10.4** 78.1	**12.4** 52.6	**15.5** 48.5	**18.4** 14.9	**21.8** 3.3	**24.4** T	**23.4** 2.7	**22.5** 12.8	**18.1** 13.3	**12.2** 41.7	**8.7** 91.7
BAKERSFIELD (CALIF.)	**8.5** 29	**11.1** 29	**13.8** 27	**17.3** 21	**21.3** 6	**25** 2	**29** T	**27.7** T	**24.7** 2	**19.3** 8	**13.3** 12	**9.3** 25
NAPA (CALIF.)	**8.6** 123.4	**10.3** 108.4	**11.9** 84.3	**13.6** 43.9	**16** 22.1	**18.3** 5.3	**19.5** .25	**19.1** .76	**19.2** 4.6	**16.5** 30.9	**12.5** 57.6	**9.4** 124.7
YAKIMA (WASH.)	**−2.6** 30.2	**1.1** 22.1	**5.5** 15.7	**10.3** 11.9	**14.7** 13.7	**18** 20.3	**21.6** 3.3	**21.1** 5	**16.3** 8.9	**10.3** 15.2	**3** 24.4	**−2.7** 28.4

STATION	TEMPERATE CLIMATE											
	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.
REIMS	**1.6** 45.7	**3.3** 38.1	**7.7** 48.3	**10** 48.3	**13.3** 53.3	**16.6** 53.3	**18.8** 66	**18.3** 58.4	**15.5** 43.2	**10** 66	**5.5** 58.4	**2.2** 58.4
DIJON	**1.5** 61.1	**2.2** 44.8	**7.1** 50	**10.2** 36	**14.6** 54.8	**17.5** 97.6	**19.6** 50.8	**18.6** 82.9	**15.8** 75.9	**10.4** 62.5	**5.6** 63	**3.8** 65.9
NANTES	**4.9** 80.5	**5** 73.7	**8.5** 58.4	**10.5** 47.4	**14** 46.8	**16.9** 50.6	**18.5** 52.8	**18.1** 73	**16.1** 88.1	**12.3** 74.9	**7.9** 86.4	**6.8** 108.5
BORDEAUX	**5.2** 101.3	**5.4** 88.1	**9.5** 67	**11.3** 41.3	**14.8** 63.8	**17.6** 76.6	**19.1** 69.6	**19** 75.3	**17** 87.7	**12.6** 86.3	**8.1** 85.2	**6.8** 125.5
FRANKFURT	**−.3** 53.4	**1.9** 50.6	**6.2** 35.1	**11.3** 39	**14.9** 58.5	**17.7** 76.9	**19.7** 67.7	**19** 74.7	**15.7** 43.9	**10.4** 30.9	**5.1** 62.4	**1.5** 58.8
GENEVA	**1.1** 63	**2.2** 56	**6.1** 55	**10** 51	**14.1** 67	**17.8** 89	**19.9** 64	**19.1** 94	**15.8** 99	**10.3** 72	**5.7** 83	**2.1** 59
VIENNA	**−.4** 34.6	**.1** 34.2	**4.4** 46.6	**9.9** 50.7	**14.4** 55.9	**18** 83.5	**19.9** 87.8	**19.2** 72.4	**15.3** 33.9	**10** 41.6	**4.9** 40.2	**2** 48.7
BUDAPEST	**−.2** 41.3	**.6** 38	**5.5** 33.5	**11.4** 39.5	**16.2** 61	**20.2** 94.1	**22** 66.4	**21.3** 56	**17** 39.2	**11.1** 49.8	**5.8** 57	**2.6** 54.4
GENEVA (N.Y.)	**−4.1** 50.2	**−2.8** 63.6	**.3** 77.6	**8.4** 69.8	**13.4** 71	**19.3** 78.4	**21.8** 78.2	**20.8** 69.9	**16.9** 63.1	**11** 89.2	**4.8** 51.7	**1.6** 60
BALTIMORE	**1.5** 86.3	**2** 73.6	**6.1** 96.5	**12.3** 91.4	**18** 101.6	**22.5** 83.8	**24.8** 106.7	**23.8** 132.1	**20** 83.8	**13.8** 81.2	**7.5** 78.7	**2.1** 76.2
SANDUSKY	**−2.1** 68.9	**−.7** 57.9	**2.5** 73	**10.2** 86.1	**15.6** 79.7	**21.5** 88	**23.8** 87.2	**23.1** 97.6	**19.2** 50.9	**13.1** 52	**5.6** 59.4	**0** 56.1
CHARLESTON (S.C.)	**9.4** 73.9	**11.1** 88.3	**13.2** 108	**18** 72.7	**22.4** 97.7	**25.3** 130.5	**26.7** 199.8	**26.5** 134.2	**23.9** 164.9	**18.6** 74.5	**13.3** 43.7	**9.7** 75.7
ST. LOUIS	**−.2** 36.4	**2.5** 64.7	**5.5** 66.9	**13.6** 86.4	**19.1** 83.8	**24.5** 111.6	**26.6** 95.9	**26** 56.5	**21.7** 46.8	**14.8** 63	**6.8** 59.8	**2.1** 42.1
LUBBOCK	**4** 17	**6.2** 15	**9.8** 18	**15.2** 29	**20.1** 81	**25.3** 64	**26.4** 51	**26** 43	**21.9** 60	**16.2** 51	**8.8** 14	**5** 16
ALBUQUERQUE	**1.4** 8.6	**4.8** 10.9	**7.6** 12.2	**13.3** 16.5	**18.4** 20.8	**23.2** 11.7	**25.6** 34.2	**24.7** 29.8	**21.1** 21.8	**14.6** 17.8	**7.1** 5.6	**2.7** 13.2
BISMARCK	**−13.6** 14.3	**−9.6** 12.2	**−4.7** 16.2	**6** 21.5	**13** 60.6	**17.8** 86	**21.8** 48.7	**21.1** 51.5	**14.4** 25.5	**7.9** 15.3	**−2.1** 18.6	**−7.7** 10.7
BOISE	**−.9** 40.4	**1.4** 32.2	**4.5** 28.6	**9.4** 31.3	**13.9** 43.1	**18.2** 22.9	**23.8** 6.3	**21.9** 6.3	**17.6** 8.5	**11.1** 17.5	**3.6** 26	**−.1** 33.2
PORTLAND (ORE.)	**4.9** 211.8	**6.6** 118.1	**7.8** 124.7	**11.6** 66.9	**14.8** 52.9	**17.1** 47.9	**20.2** 6.8	**19.7** 22.6	**18.2** 42.2	**13.8** 97.7	**8.3** 142.9	**6.1** 169.5

CLIMATIC CONDITIONS prevailing in typical wine-producing regions are indicated by figures on mean temperature (*boldface*) in degrees Celsius and mean precipitation (*light*) in millimeters. *T* that appears occasionally among figures on precipitation means trace.

that the fruit of *V. sylvestris* was gathered and used by prehistoric populations.

The wild vines (often called *lambrusques,* from the French for wild vine or wild grape) continued to flourish until well into the middle of the 19th century, much as wild vines are scattered over most of North America today. Then they were virtually obliterated by the phylloxera insect, which had been accidentally introduced from America and to which they had no immunity. Even now occasional surviving specimens are found in parts of the Balkans and the Pyrenees and other isolated places. Finding one is rather like stumbling on an overlooked Neolithic tribe.

The vines that stayed behind in the refuges evolved quite differently. They were winter-tender and, befitting the latitude, were short-day rather than long-day plants. They had smooth rather than leathery foliage, grew stockily rather than rampantly and required a longer growing season.

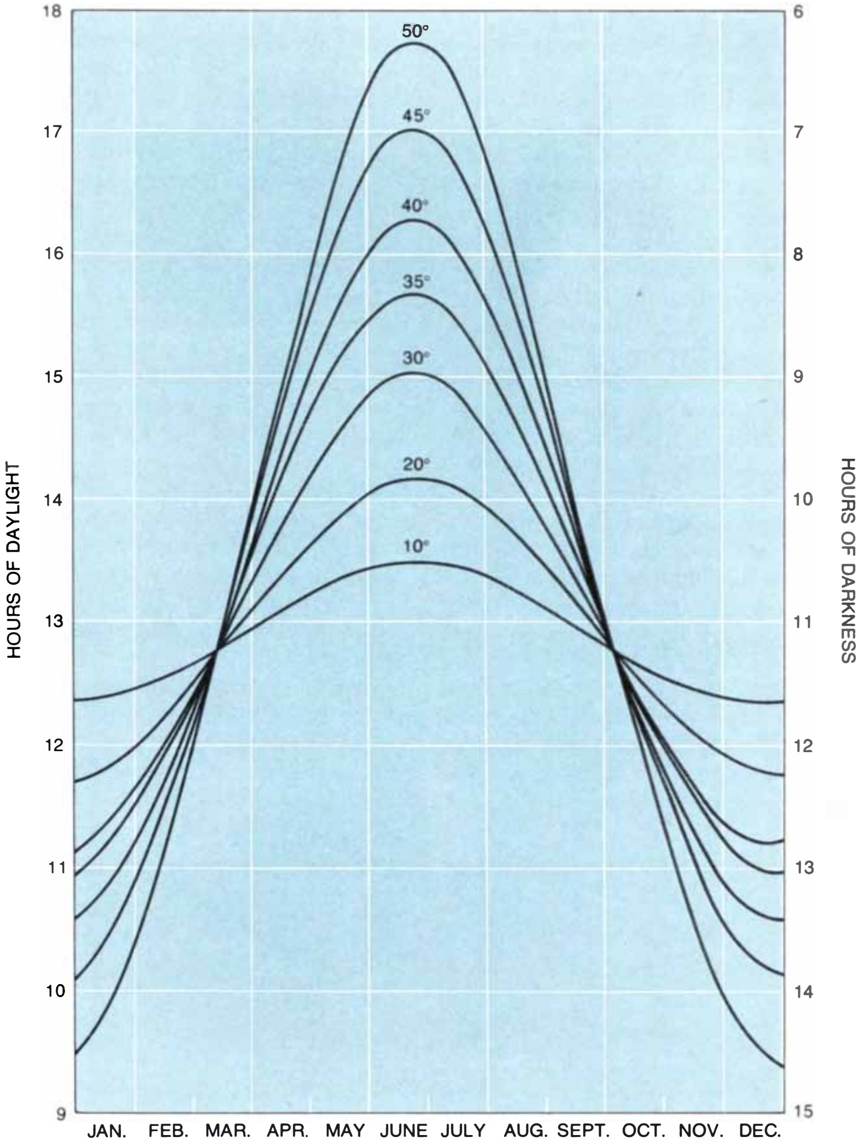
Then a dramatic thing happened. In their wild state grape vines are dioecious, which is to say that there are male and female vines and that to bear fruit the functional female must be pollinated by a male, or staminate, vine. Several thousand years ago in the Caspian refuge, however, one mutant wild vine or more appeared that were monoecious: the flowers combined the pistil, or ovary, of the functionally female vine with the pollen-bearing stamens of the male vine, so that the plants could fertilize themselves. (The story of this development has been largely pieced together by the Russian ampelographer A. M. Negrul and his fellow workers.)

The development was crucially important in several ways. Grapes can be propagated vegetatively, that is, a woody cutting will strike roots and the resulting vine will reproduce the parent vine exactly. In this way any desirable mutant or hybrid can be multiplied uniformly and indefinitely. Thus it was possible to single out, cultivate and multiply a favored monoecious variety and gain all the advantages of uniformity and of regular and reliable self-pollination. It may well be that this is how the first cultivated vines came into being.

A second advantage of the hermaphroditic trait is that it can be introduced by normal cross-pollination into the genetic makeup of nearby wild vines that lack it but might have other desirable characteristics, such as good wine quality or good eating quality. And so the way was cleared. The self-pollinating characteristic was spread by degrees (but not uniformly) through the coexisting population of wild dioecious vines. Man's role was simply to select. Once a desirable new variety was spotted the vegetative reproduction of it was a simple matter.

Selection followed cultural preferences. In some cultures the emphasis was on large grapes for eating or for drying into raisins; in other cultures it was on wine quality. Negrul and others have shown that the large-berried plants, including some that bear seedless grapes, were concentrated in an area south of the Caspian Sea. This group is commonly known today as *Prol.* (from the Latin *proles,* meaning progeny) *orientalis.* The smaller-berried types more suitable for wine congregated in the Black Sea basin; Negrul's name for this group is *Prol. pontica.*

The pathways have been worked out (still incompletely) mainly by the highly refined study of grape seeds, which in ampelography are the equivalent of teeth in human paleontology in the sense that they survive. It is evident from the archaeological remains that wild-grape gathering continued along with the growing of cultivated forms. Evidence has also been found of a specialized wine industry, which by the sixth century B.C. had become the dominant industry of many settlements, particularly around the Black Sea. From places such as Kherson (three miles west of Sevastopol), Myrmekion and Tirithica the remains of everything needed have been exposed: grape knives and presses, fermenting and storage tanks of stone, cement and pottery and treading platforms. Some of them resemble equipment still employed in a few wine districts, and any experienced vintner would recognize all the artifacts immediately.



LENGTH OF DAY at various latitudes in the Northern Hemisphere is charted. The key latitudes for wine are 30 degrees, which is about the southern limit for Mediterranean wines, and 50 degrees, which is the approximate latitude of Mainz and Prague and represents the northern limit for temperate-climate wines. Well-adapted wine-grape vines in 30-degree region are short-day plants, whereas long-day plants are more at home in higher latitudes.

From this viticultural heartland there followed a gradual westward migration, partly in the natural manner and partly by the agency of people who took vine cuttings with them. By incessant cross-pollination of the eastern vines with the wild ones of the more westerly specimens of *V. sylvestris* encountered along the way the Mediterranean basin became inhabited with wine grapes of the kind still dominant there today. The Iberian Peninsula was particularly rich in such useful hybrids, many of which went into a reverse migration eastward throughout the basin. Conspicuous among these more western types (the *Prol. occidentalis*) are the Grenache (Garnacho and many other synonyms), the Carignane (Cariñeno and so on) and the Aramon, which today produce floods of fairly ordinary wine in many parts of the basin.

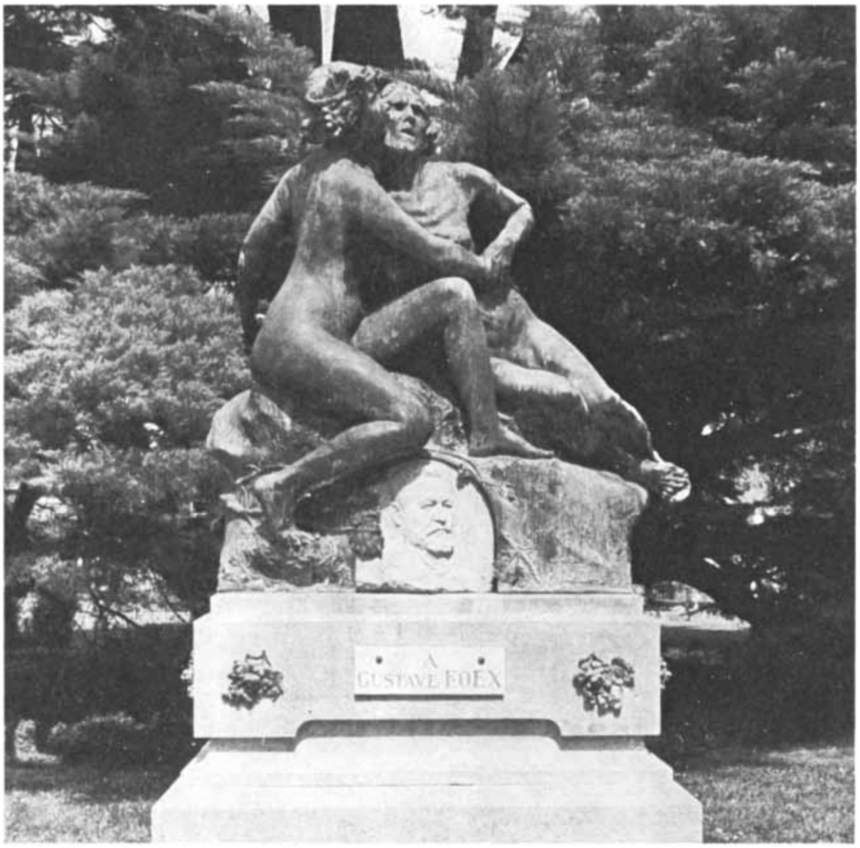
Another chapter in the story was the evolution of the wild *V. sylvestris* in temperate Europe into ecological groupings. The vines of northeastern France differed from the ones in southwestern France and so on. Intruding into temperate Europe, the Mediterranean peoples brought their grape vines along, as usual. For the most part the vines were unsuited to the new climate, but they were quite capable of surviving long enough to cross with the indigenous wild vines. Out of the resulting jumble came wild seedlings combining some of the good qualities of each parent: self-pollinating, long-day plants, well adapted to a temperate climate, showing hybrid vigor and providing a range of new aromas and flavors of particular value for wine.

It is from these vines that the cultivated wine grapes of temperate Europe were selected. The names and synonyms of many of them testify to their wild origin: Plant des bois, Sauvaget, Sauvignon, Savagnin, Sauvagnou, Fer Servadou and Lambrusquet. An etymologist steeped in the country dialects of Europe could add hundreds more.

Further proof of this random origin is found in the immense variation within what the French call a *cépage,* which is not a single and uniform variety that goes back to one first vine but a group of similar but not absolutely identical individuals. Let me mention Pinot Noir again as an example. The name does not stand for an individual but for an assortment sharing the same genetic background. Some of them are called Pinot with a place name added (Pinot d'Aunis, Pinot d'Ambonnay). The names of others introduce some special characteristic along with the place name (Gros Plant dore d'Ay) or the name of the popularizer (Pinot Renault). Dozens of Pinots are found in the Champagne district, dozens more in the Côte d'Or and no fewer than 40 in the German-speaking area of Switzerland alone. The total easily reaches 1,000, all of which are entitled to be called Pinot Noir. So it is with the names of other famous *cépages:* the ubiquitous Chasselas, the Côts, the numerous Carmenets (including Cabernet), the Rieslings—each *cépage* a multitude of cousins and sisters and aunts, the individuals tracing back to some single vine or mutation that crept in unnoticed or appealed to someone once and caught on. If it all seems terribly complicated, it is. A real understanding and appreciation of wines with their gross and subtle differences begins with an insight into these endless adjustments of vine to climate—in lesser degree to soil—and to the men who have grown them and known them intimately generation after generation.

Up to this point the story has been one of natural selection, with man content to pick and choose among the results. The story has had to lean on surmise and inference at various points. Now let us turn to the deliberate intrusion of man into the genetic process over the past century. The work has involved the creation of entirely new varieties by making crosses both within the Eurasian species and between different species.

The devastation of the phylloxera epidemic in Europe during the late 19th century gave the work its real start. The desperate need then was to find phylloxera-resistant rootstocks on which the classic varieties could be grafted and so could survive. Doing so meant resorting to the entire array of American species, which were then hardly sorted out but were and are resistant to phylloxera in varying degrees because of adaptation to that indigenous insect.



ALLEGORICAL STATUE on the ground of the viticultural school at Montpellier in France portrays a young woman supporting an older one. The young woman represents American rootstocks that revitalized the vineyards of France, which are represented by the older woman, after the French vineyards were devastated in 19th century by attacks of plant lice of the phylloxera type. Gustave Foex was a viticulturist who headed the school at the time.

Finding rootstocks that were compatible and also adaptable to the conditions of Europe required combinations and recombinations. The notion of combining one trait of, say, the Appalachian *V. rupestris* with another trait of, say, the Texan *V. Berlandieri* and then crossing the offspring with other varieties up to the fifth and sixth generation became commonplace. Out of this work came the cryptically named rootstocks (AxR 1, E. M. 41, Rupestris St. George, C. 3309, 5BB and so on) on which the vineyards of Europe, California, Australia, South Africa and Chile flourish today.

Along with this work came a divergent and competitive effort, which was to bypass the need for grafting by breeding hybrids that would combine the fruit and wine quality of the European grape with the winter hardiness and resistance to disease of these alien species. Conceivably such hybrids not only would extend the growing of wine grapes into new areas but also would offer advantages over some of the ancient existing varieties.

On both grounds the new effort encountered hostility. The established vintners had no wish for competition from new areas. In terms of quality the producers of the *grand appellations* were well satisfied with what they already had, now that their vines could survive on the new rootstocks.

Further complicating matters, hybridizing yielded rather dreadful results in the beginning, recalling George Bernard Shaw's warning to Ellen Terry that a child of their union might inherit her brains and his beauty rather than the reverse. Hardiness and resistance to disease were improved, but at the cost of quality. The work nevertheless persisted, being done in the main by a little band of dedicated men caught up in a dream.

As one generation of such crosses succeeded another, the results improved. There exists today a range of these hybrids, some of exceedingly complex ancestry, which yield good wine under conditions hostile to *V. vinifera.* They are not great wines; so far no hybrid yields a wine comparable to the great burgundies and clarets and Rhineland wine. One reason is that some genetic linkages bring over both good and bad traits. It is worth remembering, however, that most of the world's wine is from the *vinifera* grapes and that 90 percent of it is far from great. The tendency is to compare the worst hybrids with the best *vinifera,* not the other way around.

Throughout this effort the European wine-producing establishment has stood implacably hostile. In the wine-grape regions of Europe the planting of all but a few of the hybrids is firmly outlawed, and continued private work in hybridization is so hedged about with prohibitions and restraints as to be practically impossible. The attitude is strange in the light of the great gifts that the application of genetics has brought to other branches of agriculture and to animal husbandry.

No doubt viticultural Europe can get along without hybrids. In other areas, particularly the U.S., hybridization is proceeding vigorously. In California the work has taken mainly the direction of producing new *métis* (crosses within species) by crossing Europe's temperate-climate and Mediterranean vines with a view to obtaining better grapes for the fertile but desertlike Central Valley. Some of the results have been remarkable. The leaders in the work have been Harold P. Olmo of the University of California at Davis and Elmer Snyder of the U.S. Department of Agriculture. Even in areas of the West Coast where the Eurasian grape is superbly adapted, the need for resistance to phylloxera and nematodes inspires continuous work with interspecific hybrids as well as intraspecific ones.

Elsewhere, particularly at the experiment station of the Department of Agriculture at Geneva, N.Y., grape breeding concentrates on crosses between species, which is hybridization in the true sense. The reason is that the Eurasian grape is not really at home in North America except in a few areas, most notably in California with its many variants of the Mediterranean climate. The continental extremes of weather prevailing in most other parts of the continent are inhospitable.

Not until the coming of the better Franco-American hybrids with their superior hardiness and resistance to disease was it possible in temperate North America to produce with consistent success wines that meet European norms. Now one finds such success in many places, some of them quite unlikely: on the high plains of Texas, in Utah and the Ozarks and throughout most of the country east of the Mississippi River, including even New Hampshire and Vermont. No one is going to be completely satisfied with the wines that result, because no "ideal" hybrid yet exists. As the best of the present hybrids are planted and new wine-producing regions come into being and the public accepts what they produce, however, the wine drinker can look forward to better and better experiences.

MATHEMATICAL GAMES

Dr. Matrix brings his numerological science to bear on the occult powers of the pyramid

by Martin Gardner

"Does the Great Pyramid of Cheops enshrine a lost science? Was this last remaining of the Seven Wonders of the World... designed by mysterious architects who had a deeper knowledge of the secrets of this universe than those who followed them?"

—PETER TOMPKINS,
Secrets of the Great Pyramid

I was thumbing through one of those sleazy newsstand magazines devoted to the occult when a full-page advertisement caught my eye. It was a photograph of a six-foot-high transparent plastic model of the Great Pyramid of Cheops. Seated inside, clad only in sandals, was a beautiful dark-haired girl with Oriental eyes. She looked exactly like Iva, the half-Japanese daughter of my old friend, the renowned numerologist Dr. Irving Joshua Matrix.

I had seen advertisements before for small models of Cheops' pyramid (in the Edmund Scientific Company catalogue, for instance) but not for models large enough to sit in. Each edge of the structure in the advertisement was mysteriously labeled with a different number from 1 through 10. There was no explanation of the numbering, nor could I find a price for the pyramid. However, $5 would buy an exact scale model six inches high. A booklet came with it, telling how the pyramid's "psi-org energy" would keep razor blades sharp, preserve rosebuds and restore old typewriter ribbons. I would also be informed, said the advertisement, of how to obtain the larger model, which was unconditionally guaranteed to cure my bodily ills, raise my intelligence, strengthen my psi powers and build up my sexual potential. The address was Pyramid Power Laboratories, Post Office Box 123, Pyramid, Nev.

Is there such a town as Pyramid, Nev.? I checked my atlas. Indeed there is. It is on the west shore of Pyramid Lake, about 35 miles north of Reno. I had no trouble obtaining the laboratory's telephone number from the Reno operator. A few minutes later I was talking with Iva herself.

"Come out and see us," she said. "Do you like to fish?"

I told her I did.

"Then bring a rod and reel. If the trout are under 19 inches, you have to throw them back in the lake. The weather's marvelous here in May. Hot days, cool nights. Long time no see."

Before giving an account of my remarkable visit to Pyramid Lake, let me say something about "pyramid power." According to *Time* (October 8, 1973), it all began 70 years ago when a French occultist, impressed by the excellent condition of mummies in the Great Pyramid, asked himself: Is it possible that the pyramid's shape does something peculiar to space and time? He put a dead cat inside a scale model of the pyramid. The body quickly dehydrated and mummified. Fifty years later Karel Drbal, a radio engineer in Prague, discovered that a razor blade, kept inside a six-inch-high model of the pyramid, never gets dull. More than that, a dull blade left inside for several weeks becomes sharp again! Drbal patented his Pyramid Razor Blade Sharpener in 1959 and made a tidy fortune selling little cardboard and Styrofoam models in Czechoslovakia. When Sheila Ostrander and Lynn Schroeder reported this in their 1970 best seller *Psychic Discoveries Behind the Iron Curtain,* it kicked off what *Time* called a "minicraze" in the U.S. and Canada.

Max Toth's Toth Pyramid Company of Bellerose, N.Y., now sells a cardboard razor blade sharpener for $3.50. Evering Associates distributed a similar model in Canada. In Glendale, Calif., G. Patrick Flanagan sells a Cheops Pyramid Tent made of vinyl. You sit inside it to improve your transcendental meditation. Gloria Swanson, reported *Time,* sleeps with a pyramid under her bed because it makes "every cell in my body tingle." James Coburn likes to meditate in his pyramid tent.

Eric McLuhan, eldest son of Marshall McLuhan, has been doing research on pyramid power. A cover story about him in the February 1973 issue of *enRoute* (a magazine published for Air Canada passengers) is based on an interview with Eric when he was teaching "creative electronics" at Fanshawe College in London, Ont. Eric tells of putting one piece of hamburger in the center of his Plexiglas pyramid and another on the floor of the pyramid. Three months later the meat in the center was still fresh. The other was unfit to eat. Eric believes the pyramid's shape alters gravitational and magnetic fields inside and above it. The dull edge of a razor blade is not improved, he says, unless its edge is kept aligned on a north-south magnetic axis. Blue steel blades work better than stainless steel. John Rode, who runs an occult bookstore in Toronto, experimented with larger pyramids. He has dehydrated hundreds of eggs and 10 pounds of porterhouse steak. "It takes," Rode declares, "about 23 days to halt the decaying process in meat.... We've fried the steaks and eaten them. They're delicious."

Al G. Manning, head of the ESP Laboratory in Los Angeles, has an article on pyramid power in the October 1973 issue of *Occult.* He recommends writing on a piece of paper a statement of something you intensely desire to happen. "Feed it with lots of love, and place it tenderly in the properly north-south oriented pyramid. The paper is then left in the pyramid for periods ranging from three to nine days with chanting and feeding of the thoughtform through the north side of the pyramid on at least a daily basis." Manning reports spectacular results.

The 1974 issue of *Other Dimensions,* an annual review of the paranormal, has an informative article on "The Lady Who Lived in a Pyramid." It seems that when Tenny Hale, an Oregon sensitive, sat inside her large wood model of the Great Pyramid, she found that her psychic powers were so enhanced she emerged and typed 100 different prophecies. She also reports that when she waters plants with water stored under her pyramid, the plants grow four times faster than plants watered with ordinary water.

James Mullin, research director of the Southern Parapsychology Foundation in San Diego, is quoted as warning people not to spend too much time in a pyramid. Pyramid energy kills bacteria that cause decay. Since bacteria can be good as well as harmful, says Mullin, prolonged exposure to pyramid power might be a health hazard.

Readers who wish to conduct their own experiments will find that it is easy to make a cardboard model. Simply cut four cardboard triangles and tape their sides together. Each triangle should have its altitude in golden ratio to half its base. Herodotus was the first to suggest that the area of each face of the Great Pyramid is equal to the square of the pyramid's height. Let x be the "apothem" (altitude of a face) and the base equal to 2. The area of each face is x. Applying the Pythagorean theorem, we find that the pyramid's height is $\sqrt{x^2 - 1}$. If the area of the face equals the square of the height, we have the equality $x^2 - x - 1 = 0$. This gives x a value of $(1 + \sqrt{5})/2$, which is phi, the golden ratio, or 1.6180339887.... In other words, if the pyramid's base is 2, its apothem is phi and its height is 1.2720196495..., the square root of phi. There is a surprising bonus. Divide 4 (twice the base) by the square root of phi (height) and you get 3.1446..., a close approximation of pi [*see illustration on this page*].

The difference in slope between a perfect pi pyramid and a perfect phi one is close to one minute of arc: the slope is about 51 degrees 50 minutes for phi and about 51 degrees 51 minutes for pi. Each perfect pyramid gives a good approximation of the other constant. The variation in slope is too minute to be detectable on a small model. The Great Pyramid itself is now so irregular that all one can say for certain is that its slope is close to 52 degrees. No one knows if the Egyptians intended to embody in the pyramid's shape pi or phi, or both or neither.

But back to my trip west. I arrived in Reno by plane on the evening of Tuesday, May 28—a perfect day and a perfect number. I spent the night in a Reno hotel. Early next morning I rented a car, drove east to Sparks, then north on Route 33 to Sutcliffe, on the western rim of Pyramid Lake. I had been there once before. My father, a geologist, had taken me to Pyramid Lake when I was a boy. I remember him telling me how the lake was the shrinking remnant of a vast prehistoric body of water called Lake Lahontan that had covered most of the northwestern part of Nevada during the Pleistocene ice ages. He explained how the lake was fed by waters from the Sierra Nevada, carried mainly by the Truckee River, which flows past Reno. He took me by rowboat to some of the cone-shaped islands that give the lake its name, particularly to Pyramid Rock on the lake's eastern bank. I own my father's copy of the journal kept by John Charles Frémont, the famous explorer and politician (he was once governor of California) who discovered the lake in 1844. My father had marked the passage in which Frémont describes the 600-foot calcareous formation called Pyramid Rock: "We encamped along the shore, opposite a very remarkable rock ... which from the point we viewed it, presented a pretty exact outline of the Great Pyramid of Cheops."

I stopped at Sutcliffe to buy a fishing license. Pyramid Lake is entirely within the huge Pyramid Lake Indian Reservation owned by the Paiute Indians. One is not allowed to fish without a permit obtained from tribal officers. After getting my permit I drove north on the deserted dusty road that winds along the lake's barren western shore.

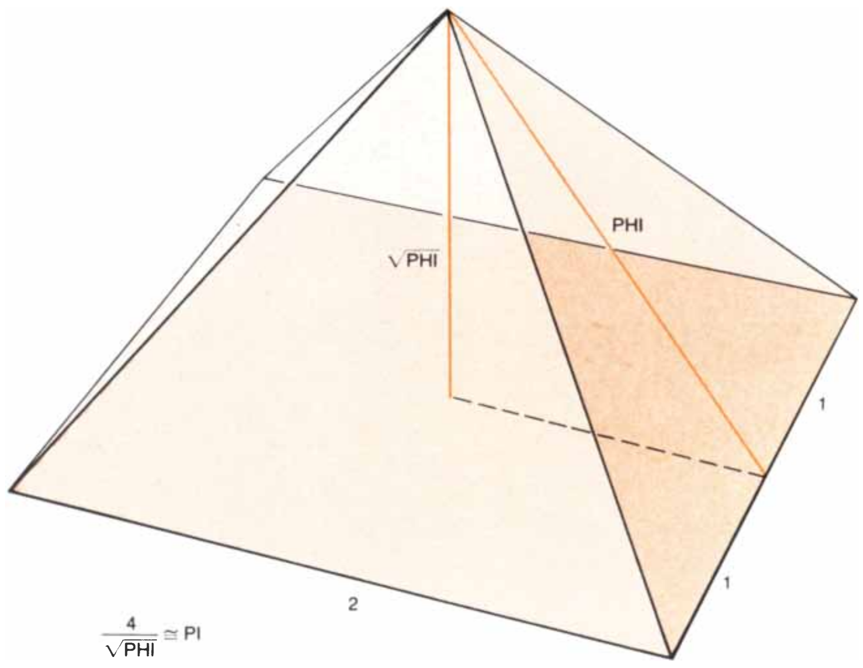
The day was warm and cloudless. Through the right window of my car I could see, beyond the sagebrush, the deep Prussian blue of the lake. Jagged spires and pinnacles along the opposite shore were casting purple shadows over the water, and above the turrets the Nightingale Mountains undulated in soft shades of green and pink. Just before entering Pyramid I turned off on a side road, as Iva had instructed. The road soon ended at a massive factory building of steel and concrete. It had been built in the shape of the Great Pyramid of Cheops. Large blood red numerals gleamed on the structure's edges.

A pudgy Paiute opened the front door for me. He grinned broadly, revealing a mouth that contained only a single front tooth. (His name was Ree, I later learned. Everyone called him One-Tooth Ree.) Down the hall and walking toward me came Iva and her father. Iva was wearing bright orange pants, a beaded Indian headband and a charm bracelet of little silver pyramids that jangled pleasantly as she walked. We embraced. Dr. Matrix stood by, tall and bony, his canny green eyes glittering behind rimless pentagonal spactacles.

They took me on a quick tour of the factory. In one wing about 20 Indians of both sexes were assembling the six-inch pyramid models. In another wing a smaller group of Indians was cutting and packaging the unassembled sides of the larger model. Iva excused herself, and I followed Dr. Matrix up a helical stairway to his office in the factory's apex.

Psi-org, Dr. Matrix explained, leaning back in his desk chair and touching fingertips to fingertips, combines abbreviations for psychic and orgone energy. They are different names for the same force. The psi field, which produces the human aura and is responsible for all psychic powers, is none other than what Wilhelm Reich, Freud's controversial Austrian disciple, called orgone energy.

"I remember Reich's orgone," I said. "It comes from outer space. It makes the



The perfect phi pyramid, after Herodotus stars twinkle, the sky blue and Orson Bean happy."

"Precisely," said Dr. Matrix. "The hundreds of pyramid islands in Pyramid Lake trap the energy and this gives the strong blue color to the water. Reich's great discovery, as you know, was that orgone could be accumulated by building a box with wood on the outside and sheet iron on the inside. The organic material lets orgone through; the metal interior reflects. What I call the 'bluehouse effect' takes over. Abnormally high concentrations of psi-org energy build up inside the box. Reich's basic idea was sound, but he had the shape of his box wrong. The Egyptians knew all about psi-org energy. They used it, you know, to float heavy stones across the desert when they built their pyramids. They were the first to discover that the shape of the Great Pyramid concentrated psi-org. I was the first to discover that if this shape is combined with Reich's principle of laminated substances, the bluehouse effect increases by a factor of 777. I call it the pi-phi-psi pyramid."

"But your pyramids aren't laminated," I said. "They're just single sheets of Plexiglas."

"Wrong," said Dr. Matrix. "Examine them more carefully and you'll see that every side consists of two thin layers of plastic. Each sheet is made with a different and secret formula. The outer layer transmits psi-org, the inner layer reflects it."

"Has this been verified?"

"To the hilt. We've done hundreds of carefully controlled tests under the supervision of Dr. Harald Puton, a very competent Belgian physicist who was active in the Scientology movement in Brussels a few years ago. He found that every form of psi energy is increased by sitting under a pi-phi-psi pyramid. One is more telepathic, more clairvoyant, more precognitive. It is easier to initiate out-of-body experiences. Uri Geller, the Israeli psychic, visited us a few weeks ago and found that when he was inside the pyramid, any metal object he touched melted instantly. Psychic healing is enormously accelerated. Last week One-Tooth brought in his sister. She'd broken her left leg. After one hour in a pyramid her leg was good as new."

Was Dr. Matrix smiling faintly? It is always hard to know where his beliefs end and deception begins. The body's aura, he continued, is more intense inside a pyramid. He took from his desk two Kirlian photographs of a living butterfly. The one made outside the pyramid showed only a faint white aura. The one made inside showed a bright blue aura that extended several inches beyond the butterfly's wings.

"Would you mind explaining the significance of these numbers?" I asked, pointing to the scarlet numerals on a pyramid paperweight.

Dr. Matrix did not mind. Each of the pyramid's eight edges, he said, bears a different number in the set of integers from 1 to 10 inclusive. At each vertex (the four base corners and the apex) the sum of all edges meeting at that vertex is $4^2 = 16$. I picked up the paperweight and turned it in my hands. What he said was true. At each base corner the number on the three meeting edges added up to 16. The four edges meeting at the top also added up to 16.

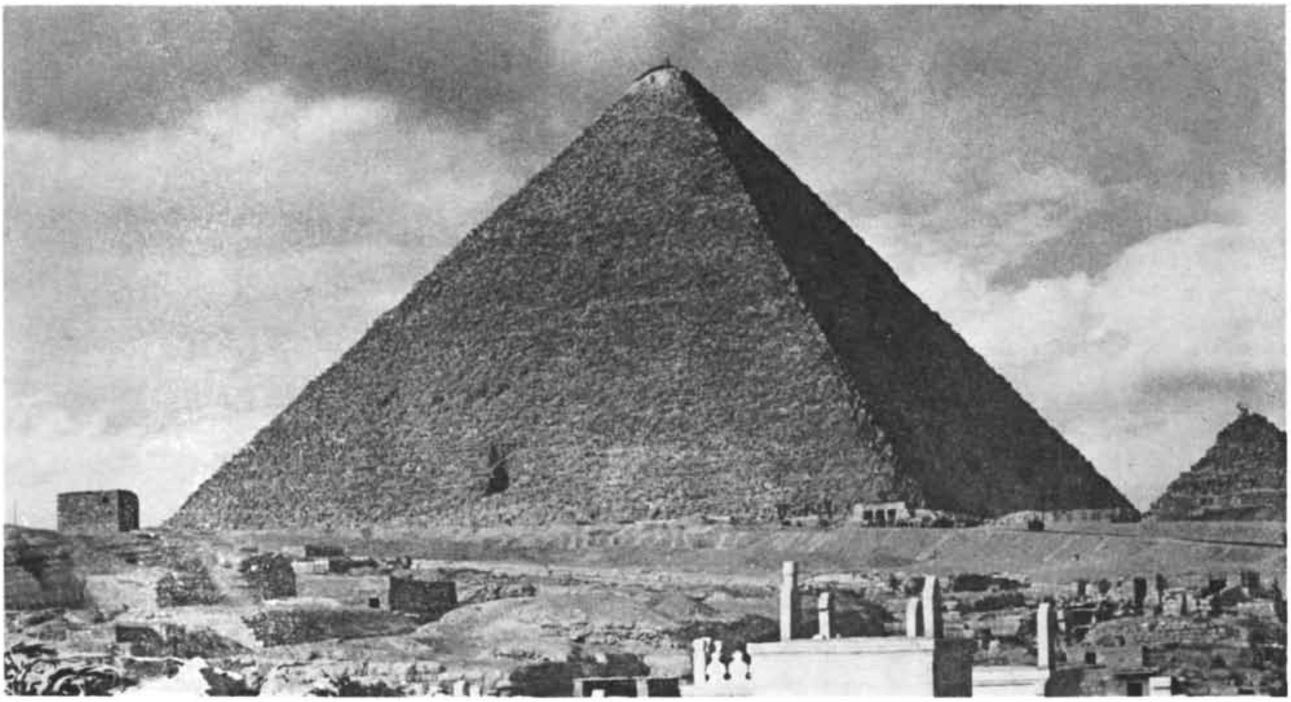
The labeling, Dr. Matrix went on, makes it a magic pi-phi-psi pyramid. The magic increases the pyramid's power. He called my attention to the fact that P, the initial letter of PYRAMID, is the 16th letter of the alphabet. Between the letters PI in PYRAMID is the name of the mother of Jesus spelled backward.

"What about the D?"

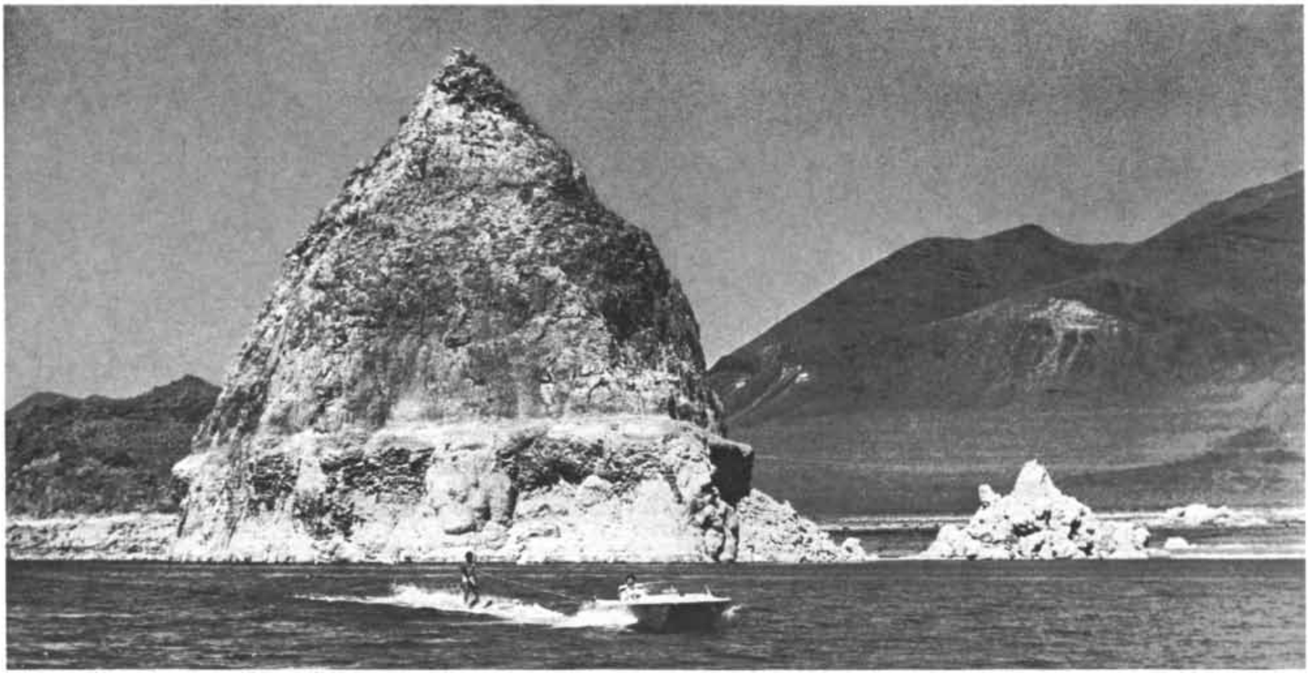
"It's the alphabet's fourth letter. It symbolizes the structure's four sides and the square root of 16, the magic constant. You might ask your readers to see if they can number a pyramid's edges properly to make it magic. There's only one way to do it. Of course we don't count rotations and reflections as being different."

"Excellent," I said, sketching the pyramid on my note pad and numbering the edges in case I forgot the solution. "I'll give the problem but hold the answer until the following month. Anything else you can tell me? Anything numerologically interesting about Nixon and Watergate?"

"*Everything* is numerologically interesting," he said. "But Nixon's future is too painful to discuss. Incidentally, there's a town named Nixon just south of



The Great Pyramid of Cheops



Pyramid Rock in Pyramid Lake, Nev.

the lake. He could do worse than retire there. A few hours a day sitting on top of Pyramid rock would do the President a world of good."

"Any comments on Henry Kissinger? Someone told me recently that if that celebrated Polish opera star Wanda Waleska married Howard Hughes, divorced him and married Henry, she'd be Wanda Hughes Kissinger now."

"Your humor," said Dr. Matrix without smiling, "underwhelms me. You're familiar, of course, with the last verse of Chapter 13 in the last book of the Bible?"

I nodded. "Let him that hath understanding count the number of the beast: for it is the number of a man; and his number *is* Six hundred threescore *and* six."

"Well, I don't want to suggest that Kissinger is the Beast, but here's a little number curiosity your fans may find amusing. Let *A* equal 6, *B* equal 12, *C* equal 18 and so on, taking the multiples of 6 in sequence. Then add the values of KISSINGER. The total is 666."

"Marvelous," I chuckled, scribbling furiously. "Yesterday while I was on the plane it occurred to me that NEVADA, which has six letters, was admitted as the 36th state. And 36 is the square of 6."

"And dice," said Dr. Matrix, "have six sides. Six is the highest number on a die and 36 is the highest number on a roulette wheel—two reasons why the 36th state contains the nation's three largest gambling resorts. Six is the sum of 1, 2, 3, and 36 is the sum of the cubes of 1, 2, 3. Of course, six is also the product of 1, 2, 3. The 36th triangular number is 666, which is a way of saying that 666 is the sum of all the roulette numbers."

"I once suggested somewhere," I said, "that if the Devil plays rotation pool, he plays it on a huge table with 666 balls in triangular formation before the break."

"Quite likely. Six is the Devil's number, all right. It was the sixth hour of the sixth day of Creation when he tempted Eve. But it was the Egyptian sensitives who made the fullest use of the square of the sum of 1, 2, 3."

One-Tooth stuck his head in the doorway. "Did you call?"

Dr. Matrix shook his head and waved Ree away. "As I was saying, Osiris had 36 forms. The human body was divided by the Egyptians into 36 parts, each subject to a different demon, each controlled by one of the 36 parts of the Egyptian zodiac. And I could bore you for hours with a discussion of Nabokov's use of 36 in *The Real Life of Sebastian Knight*. Sebastian's house number is 36. His hospital room number is 36. He dies in 1936 at the age of 36, and so on. But," and Dr. Matrix glanced at his digital wristwatch, "my time is limited."

"Is it necessary to be nude when you sit inside a pyramid?" I asked.

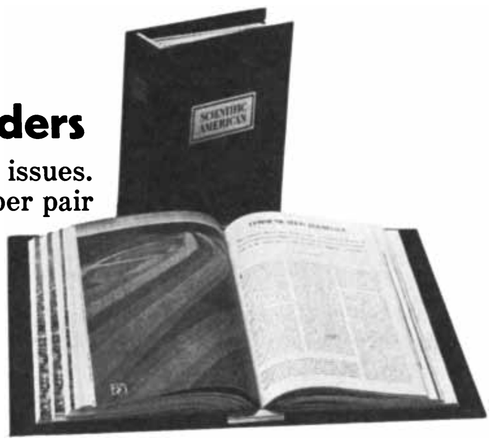
"No, but it helps. We have several opaque models on the beach for visitors too modest to use transparent ones. Last week Judy Clutch, a schoolteacher from Wadsworth, was inside one of them for only five minutes before she was so blue with psi-org that she leaped outside, ran all the way to Pyramid and streaked down the main street until the sheriff caught her and took her to lunch."

I asked about the pyramid's power to prevent and even reverse the decay of organic matter. Dr. Matrix told me he had applied for numerous patents that exploited this aspect of psi-org: a pyramid refrigerator and freezer, a pyramid coffin (no embalming necessary), a pyramid septic tank.

I stayed five days in Reno. Mornings and afternoons I fished in Pyramid Lake for king-size trout and a tasty variety of lakesucker that the Indians call *cui-ui*. Iva joined me evenings for dinner at one of the resort lodges near Sutcliffe. On Saturday I rented a motorboat and we visited Anaho Island, a 250-acre bird sanctuary where we watched the white pelicans go through their spring mating ritual. Iva's large straw hat was shaped like the Great Pyramid. She said it made her think better and corrected her astigmatism.

Several days after I had returned to Manhattan I was distressed to see in *The New York Times* that Dr. Matrix was about to be indicted by the state of Nevada. It seems he had been selling pyramid franchises for several thousand dollars each to a large number of people in other states, who in turn had been trained to sell similar franchises to other distributors, and so on. It was a typical pyramid scheme of the sort that sooner or later is bound to collapse.

I tried to reach Iva, but the laboratory



phone was no longer a working number. Two days later I read in the *Times* that when state troopers went to the factory to arrest Dr. Matrix, they found no one there except One-Tooth Ree. He had just finished burning all the factory records in a large outdoor bonfire. He handed the police a letter from Dr. Matrix. It explained that on June 6 at 6:00 A.M.—the sixth hour of the sixth day of the sixth month—Dr. Matrix and his daughter had entered one of their large pyramids and teleported themselves to a monastery in Tibet.

The discussion of cram, crosscram, Bynum and quadraphage, in February and March, prompted many letters of unusual interest. I had reported that cram, even when played on a 1-by-m board, had not yet been solved, but this proved to be incorrect. The story begins in 1934 when T. R. Dawson, writing in the December issue of his *Fairy Chess Review,* introduced a generalization of kayles in which players alternately remove k adjacent counters from a row. When there is one row, and k equals 2, we have 1-by-m cram. "Dawson's kayles" was first solved by Richard K. Guy and Cedric A. B. Smith in "The G-Values of Various Games" (*Proceedings of the Cambridge Philosophical Society,* Volume 52, July, 1956, pages 514–526). They show that the second player wins if and only if the number of counters is 0, 1, 15, 35, or equal (modulo 34) to 5, 9, 21, 25 or 29.

The same results were independently obtained by Sin Hitotumatu, who published them in a Tokyo journal in 1969. The game is also isomorphic with Regulus, given as being an unsolved game in David L. Silverman's *Your Move* (McGraw-Hill, 1971, pages 185 and 214). R. Banerji, James Hajicek and Robert James Weber were the first of many readers who solved the game with Grundy functions.

Crosscram on the order-4 square was analyzed by Ashok K. Chandra, George Dunsay, Donald Hayes, Abraham Schwartz and John Woolley. The first player wins if and only if he takes two central cells or two at the middle of a side. Chandra, Hayes and Schwartz each proved that the second player wins order-5 crosscram. Hayes and Schwartz found that the first player wins reverse crosscram on both boards.

Woolley pointed out that my description of how to trap a king on an order-8 board with two quads in the game of quadraphage was incomplete. If the king's first move is to head straight for a double corner (say the move is northwest in the illustration on page 107 of the February issue), quads must be placed on the cell two squares west of the king and on the cell immediately above that square, not on two white border cells as the text says.

In March I gave a method of trapping a rook or a bishop with two quads per move and a queen with four quads. Ashok K. Chandra, Robert Holmes, Ned Horvath, Eli Shapiro, Charles Whitmer and Thomas R. Wyant III all supplied proofs that a rook or a bishop can be trapped with one quad per move and a queen with three. Assume a rook is limited to n cells per move. The trapping strategy, on the minimum board of side $8n^2 + 3$, is to use the first $4n$ moves to place (regardless of how the rook moves) n quads at the top of each top corner and at the bottom of each bottom corner. All four corners can be sealed in this way before the rook can attack a corner cell. Since the rook can then attack only one border cell at a time, single quads suffice to complete the entrapment.

A bishop on a sawtooth board is equivalent to a rook on a regular board (as explained in March); as a result the same strategy will trap a bishop with one quad per move on a sawtooth-bordered board of side $8n^2 + 3$. A similar strategy traps a queen with three quads per move on a board of side $2n \lceil 8n/3 \rceil + 3$, where n is the queen's maximum move and the brackets indicate rounding up to the nearest integer. Regardless of how the queen first moves, $2n$ quads are placed on both sides of each corner, leaving the queen unable to attack more than three border cells on all subsequent moves. Chandra also showed that if the queen's maximum move is 2, it can be trapped by two quads per move on a board of side 67 or possibly smaller.

E. N. Adams, Robert Holmes and Thomas R. Wyant III found that the knight could be trapped with three quads per move. Adams proved this could be done on the go board (19 by 19) and possibly on a board as small as 16 by 16. The most surprising letter was from Jerry Butters, who is working for his doctorate in economics at the University of Chicago. Butters found an elegant procedure for trapping a knight with just two quads per move. His 13-page proof requires a board of side 4,500. This board can be reduced considerably, but at present there are only conjectures on the minimum-size board. I have urged Butters to write up his proof for a mathematics journal and shall report here if he does so.

THE AMATEUR SCIENTIST

An unusual kind of gas laser that puts out pulses in the ultraviolet

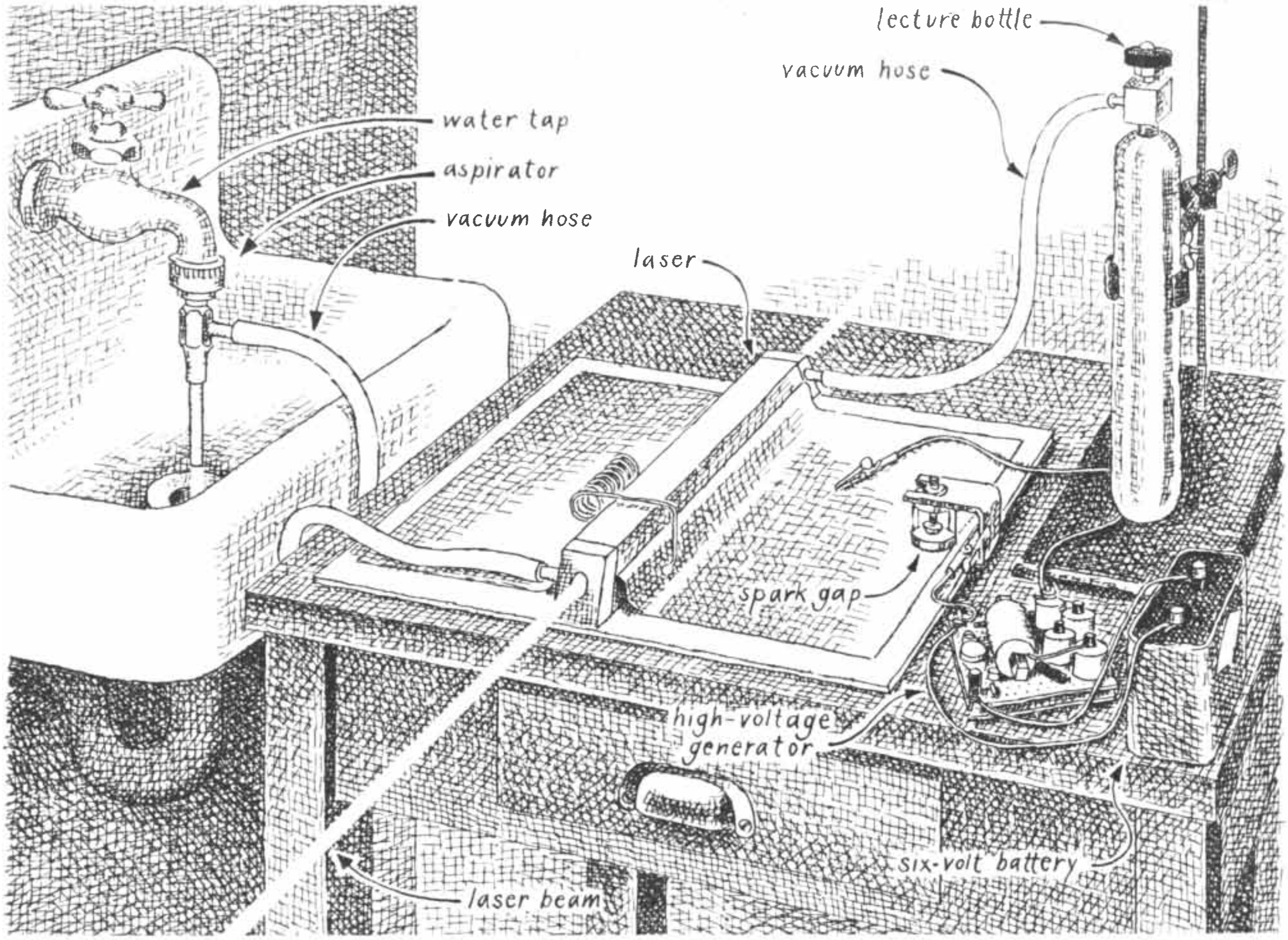
Conducted by C. L. Stong

A recently developed laser that operates on a six-volt dry battery emits 10 pulses of ultraviolet radiation per minute, each pulse about the size and shape of a broomstick. The pulses range in power from 50 to 100 kilowatts. They strike obstructions end on at the speed of light, with consequences that vary with the nature of the target materials. For example, the pulses bounce off clouds just as radar signals do.

With the echoes amateurs can measure distances to reflecting targets miles away; the accuracy is a matter of a few feet. With targets that absorb radiation the effects of impacts range from the emission of fluorescent light to the initiation of chemical reactions, including photochemical reactions. Indeed, as a source of radiation for making photographs the laser is about 10,000 times faster than the high-speed strobe lamps ordinarily used by amateurs.

Many parts of the laser can be assembled with materials that accumulate in the scrap box of anyone whose hobby is electronics. A version of the apparatus that is particularly easy for amateurs to build has been developed and patented by James G. Small, a graduate student of physics at the Massachusetts Institute of Technology. Small explains the principles on which the apparatus is based and the details of its construction and operation.

"It has been known for some time that a high-current electric discharge in nitrogen gas that is flowing at a relatively low pressure can generate a pulse of coherent radiation, which is to say a laser gain, at the wavelength of 3,371 angstroms. The wavelength lies in the ultra-



James G. Small's nitrogen laser

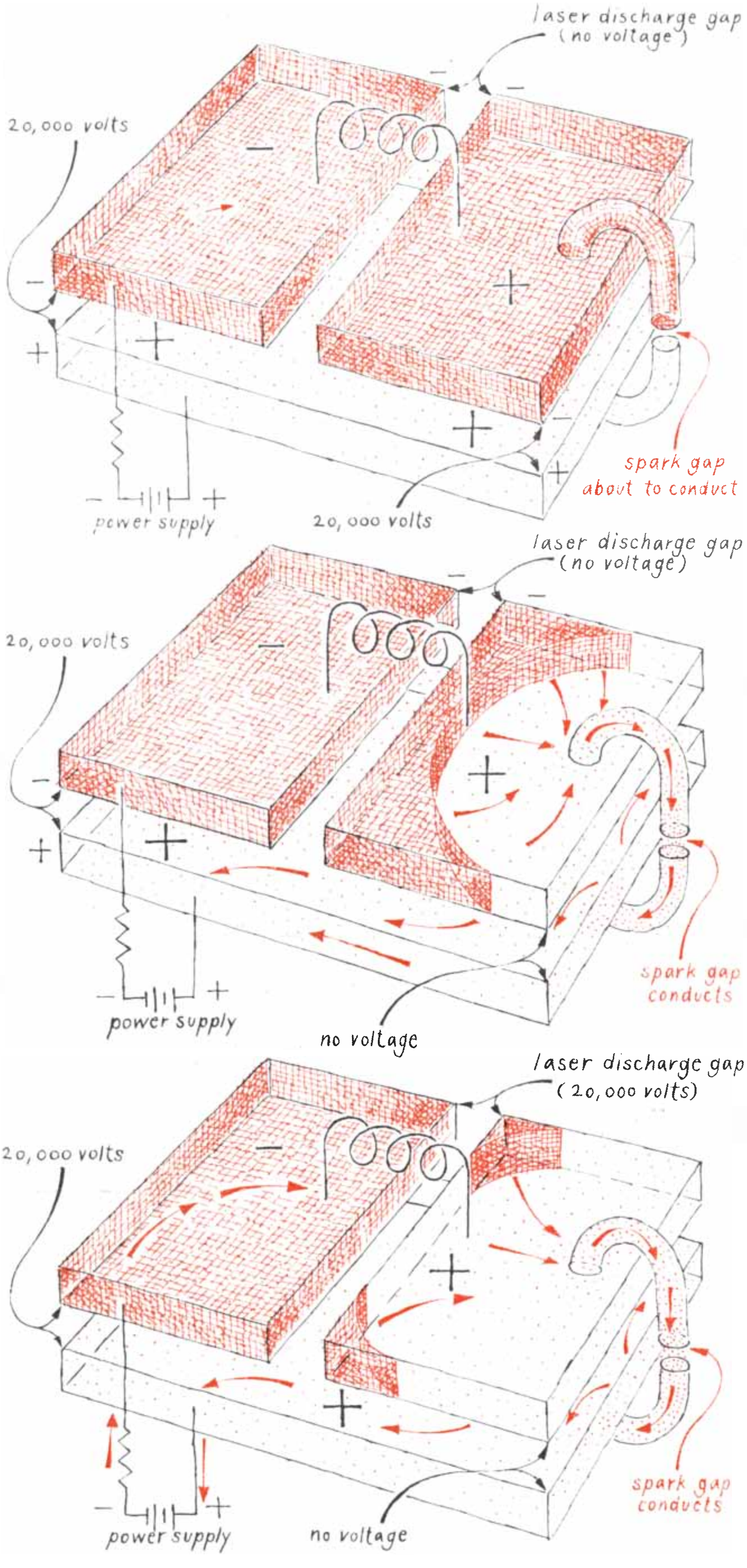
violet region of the electromagnetic spectrum where lenses and windows of most kinds of glass are transparent. The laser action begins when a molecule of nitrogen at room temperature absorbs energy by colliding with an electron that moves in the discharge. The encounter leaves the molecule in an unstable state. Usually it spontaneously falls to a state of lower energy by emitting a photon of radiation at 3,371 angstroms.

"The emitted photon may encounter another excited molecule of nitrogen and merely by its proximity stimulate the molecule to emit an identical photon. In this case the two particles of radiation join forces and proceed in the same direction, with their waves in lockstep. The resulting pulse of radiation contains twice the energy of each photon. This is laser action.

"The action will continue as long as the growing pulse encounters more excited molecules of nitrogen along its path than it does absorbing molecules. The process soon stops, however, because when a large number of molecules are suddenly excited, they will begin to randomly cascade to lower states of energy. Unfortunately in the case of nitrogen the molecules on the average linger at that lower level longer than at the upper one before moving on to still lower states. The number of molecules at the lower laser level builds up rapidly, eventually exceeding the number at the upper level and terminating the amplification. In fact, the gas quickly becomes strongly absorbing to 3,371-angstrom emission. The laser turns itself off even though there are still excited molecules left. Nitrogen lasers are therefore said to be self-terminating. The turnoff time is rather fast, usually less than 10 nanoseconds (billionths of a second), and it is responsible for the extremely short output pulse useful for radar and very-high-speed photography.

"The trick of inducing laser action in nitrogen lies in constructing a mechanism that will almost instantaneously send a huge current of electrons at a high voltage laterally through a column of the gas at a pressure of about 100 torr. An appropriate switching mechanism, which can handle tens of thousands of amperes within nanoseconds, turns out to be quite simple both in principle and in construction. It was invented by Alan Dower Blumlein, a British electronics engineer.

"Essentially the device consists of two adjacent metal plates separated from a third plate of equal total area by a thin sheet of plastic insulation [*see illustration at right*]. In effect the assem-



The Blumlein switching phenomenon

bly behaves as an adjacent pair of interconnected capacitors. The space between the capacitors serves as the gap across which electric current can be discharged through nitrogen.

"The capacitors are interconnected electrically by a coil of copper wire. The capacitors can be charged by applying a potential difference between the interconnected plates and the mutual plate. Both capacitors charge to the same potential and the same polarity. No potential difference exists across the gap between them.

"The switching action devised by Blumlein develops when one of the capacitors abruptly discharges. The action is initiated by the breakdown of a spark gap that connects one adjacent plate to the mutual plate. Current moves through the spark gap when the charge accumulating on the capacitor assembly exceeds a predetermined voltage.

"As the charge rushes through the spark gap a steep difference of potential appears within the plate across a narrow boundary that separates the charged and discharged regions of the metal. The boundary has the form of a circular wave front that recedes from the spark gap at nearly the velocity of light. Shortly after the onset of conduction at the spark gap the voltage wave arrives at the center of the discharge gap between the plates. At this instant a potential difference appears across the center of the gap. Thereafter the advancing voltage spreads to the edges of the gap. In an apparatus 12 inches wide and 18 inches long a potential appears across the full length of the gap in less than .2 nanosecond, rising to its maximum value in about a nanosecond.

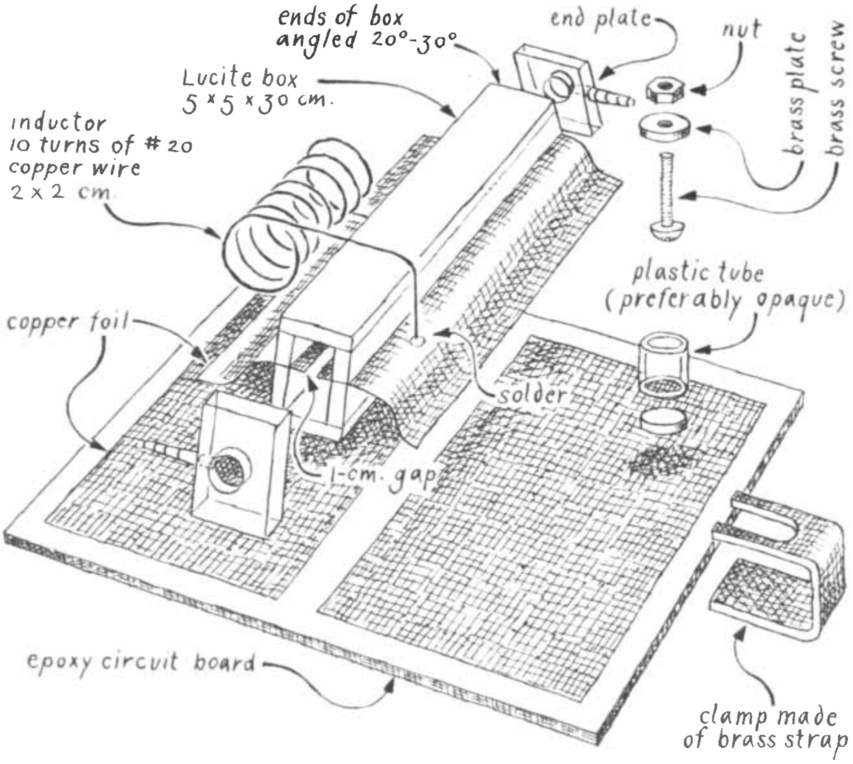
"If the discharge gap is enclosed by a container of nitrogen gas at low pressure and the capacitors are charged to 20,000 volts, the resulting discharge will raise an enormous number of the nitrogen molecules to the excited energy state. Laser action follows during the next five to 10 nanoseconds. The coil of copper wire through which the capacitor assembly is charged responds sluggishly to changes in current. For changes that occur within nanoseconds the coil acts as an open circuit.

"A practical nitrogen laser includes the switching apparatus, a power supply and a source of nitrogen gas, preferably of the grade used by welders [*see illustration on page 122*]. The compressor from an old refrigerator can be connected backward to operate as a vacuum pump for reducing the pressure of the nitrogen. An inexpensive water aspirator of the kind employed by chemists for vacuum filtering will work equally well.

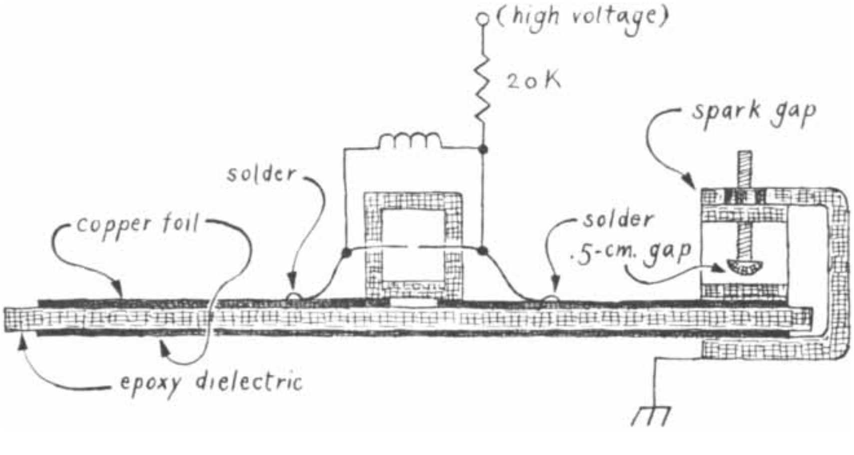
"The capacitors of my laser were made from Type G-10 epoxy circuit board, which is clad with copper on both sides. This commercially available material serves widely in the electronics industry for interconnecting electronic apparatus. The unwanted metal is removed by etching to leave a network of conducting strips.

"My capacitors were formed on a single piece of circuit board that measures 30 × 45 × .04 centimeters (12 × 18 × .015 inches). Copper was etched from a two-centimeter (3/4-inch) margin around the edge on both the top and bottom sides of the board. An additional strip five centimeters (two inches) wide was removed straight across the top side to form the two adjacent plates of the capacitor, each of which measures 18 × 26 centimeters.

"Copper surfaces that are to be preserved can be masked with waterproof tape or with the enamel spray paint that is called 'resist' by the photoengravers who developed this etching technique. The unwanted metal is preferably etched with ferric chloride, although dilute nitric acid can be used. In the U.S. all materials required for building the



Exploded view of the laser



The apparatus in elevation

laser, including the etching chemicals, are available from North Country Scientific (R.F.D. 1, Plymouth, N.H. 03264).

"The capacitors can also be fabricated from a sandwich of metal foils or plates bonded to a thin insulating sheet such as acetate or Mylar and even certain paints made with plastic resins. The insulating material should be as thin as possible but should not conduct or break down when the maximum potential difference is applied across the plates. Epoxy circuit board can be operated reliably at a potential of 1,000 volts per .001 inch of thickness.

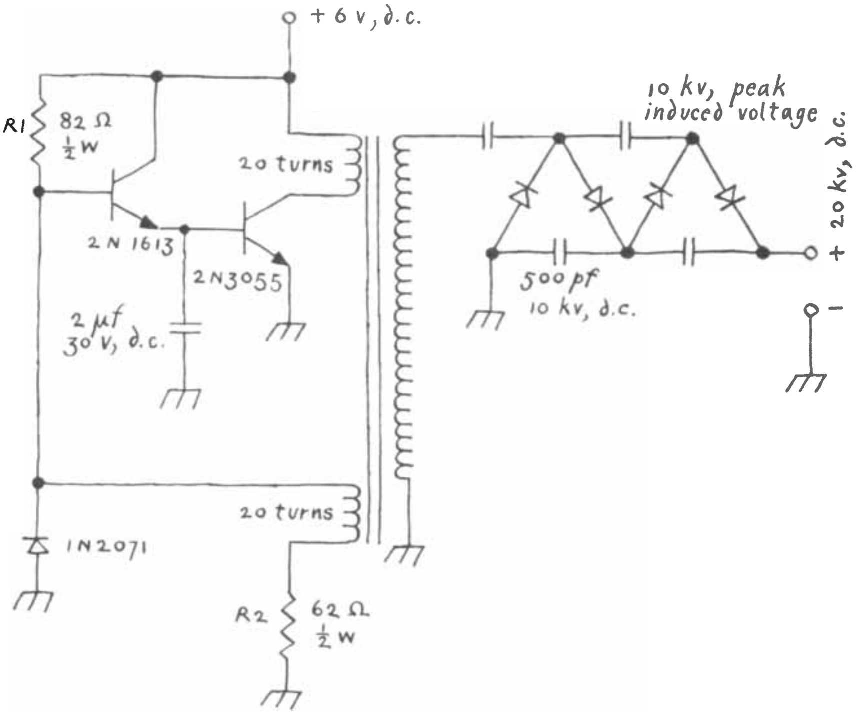
"The housing for the discharge consists of a gastight box five centimeters square and 30 centimeters long (outside dimensions). It can be made by cementing together strips of clear plastic, such as Lucite, .63 centimeter (1/4 inch) thick. An effective cement is clear silicone seal, a rubberlike material that is available at most hardware stores.

"The ends of the box are cut at an angle of from 20 to 30 degrees to prevent the reflection of radiation from the surfaces of glass windows back into the plasma, where it would be further amplified. The end plates are made of plastic that is thicker than the plastic on the sides. The added thickness makes it possible to drill holes in the edges for hose connections. These holes join other holes that pass through the center of the plates and serve as outlets for the radiation.

"The output ports are closed by windows made of glass microscope slides. The windows can be sealed to the plastic with paraffin wax of the kind used in home canning. The wax can be melted and applied safely with a heat lamp. The windows could be cemented in place with silicone seal, but wax has the advantage of reversibility (an advantage that will be appreciated when the glass becomes coated with grease from the vacuum pump or collects other foreign substances).

"The discharge electrodes of copper foil are cemented between two strips of plastic, which form each side of the box, with the edges of the foil opposed and separated by about a centimeter [*see bottom illustration on opposite page*]. The portion of the foil that extends outside the box has the form of a lazy S. The outer edge must be well soldered throughout its length to the foil of the epoxy board because the joint must conduct thousands of amperes.

"The spark gap must transmit maximum current in minimum time. It must be made of wide strap metal instead of wire to minimize its inductance. I made the clamps of the assembly with strap brass because of that metal's good conductivity, but aluminum strap could be substituted. An open air gap will work satisfactorily, but it is noisy and generates bright flashes of light that are rich in ultraviolet emission. I recommend that the gap be enclosed by an opaque tube of plastic or some other material that absorbs at least 90 percent of the emitted radiation.



Circuitry of the transistor power supply

"The laser must be charged to a potential of from 10,000 to 20,000 volts. A power supply that delivers .001 ampere of direct current will charge the capacitors to the ionizing potential of the spark gap several times per second. An alternating-current supply, such as a current-limited, neon-sign transformer rated at an output of 20 milliamperes, would cause the gap to conduct at twice the line frequency (120 discharges per second). Operation at this rate, however, would necessitate some means for cooling the gas if the maximum peak power is to be obtained. Warm nitrogen at low pressure does not 'lase' well. The gas should be allowed to cool to approximately room temperature after each pulse.

"The output is at a maximum when the discharge occurs in relatively pure nitrogen. The gas need not flow continuously. The discharge housing can be flushed with gas, pumped down to 100 torr and sealed off. The pulse-repetition rate of the sealed unit must be less than one pulse per second.

"Pulses at this rate can be generated with a battery-operated power supply [*see illustration above*]. This device consists of a low-voltage transistor oscillator, a step-up transformer and a voltage-multiplying rectifier. The transformer is made from an automobile ignition coil. Carefully cut the top from the housing of the coil. Remove the windings, along with the iron core. The innermost winding in most coils will consist of approximately 25,000 turns of fine wire. Carefully unwrap from the fine wire the outer, heavier-gauge primary winding and in its place wind two 20-turn coils of 20-gauge enameled wire, which should be closely spaced.

"Connect one of the coils in series with the collector of the 2N3055 transistor or an equivalent transistor such as the 2N3236, 2N5039, 2N6271 or HST9203. This transistor needs no heat sink, as it is operated conservatively. The 2N1613 is a low-powered transistor. It is connected in the Darlington configuration to speed the switching time of the larger transistor. Other transistors that will work nicely for this purpose include the 2N696, 2N2222 and 2N3642.

"The circuit is grossly overdesigned. Experimenters outside the U.S. will find that almost any silicon NPN transistor with a beta of 30 or more and a power dissipation of more than half a watt can be substituted for the 2N1613. For the larger transistor try any NPN of the silicon type with a collector current rating of five amperes or more that is housed in a TO-3 case. Surplus dealers in the U.S. have been selling large transistors of this kind at four for $1. When transistors are substituted, the values of resistors R_1 and R_2 may have to be altered experimentally within a factor of three to ensure reliable oscillation. The frequency of oscillation should be a few kilohertz.

"The power supply will operate from a six-volt battery at from 1.5 to three amperes and will deliver 20,000 volts at current sufficient to charge the laser to ionizing potential in from three to eight seconds. The electrical efficiency is quite low (about 1 percent), primarily because of the inefficient but simple transformer. Do not omit the two-microfarad capacitor that connects from the chassis to the junction between the emitter of the first transistor and the base of the second transistor. Without this capacitor the unit may tend to oscillate at a frequency too high for efficient transformer operation, depending on the loading of the high-voltage output.

"The reader may wonder why no mention has been made of laser mirrors. The optical gain of the rapid discharge is so large that emission becomes superradiant, which means that the unit will lase without an optical cavity. Radiation that is spontaneously emitted from molecules at one end of the laser can be amplified so strongly by the time it reaches the other end that the laser approaches saturation, meaning that it reaches the limit of its amplifying ability. Output is emitted from both ends of the column of excited gas, but a mirror at one end will more than double the power at the other end.

"The pulses can be detected easily by holding a piece of bleached cloth in front of either window. The cloth will fluoresce brightly. Almost any bleached cloth will function as a fluorescent screen. In general anything that glows in 'black light' works well: shirts washed in detergents containing fabric brighteners, 'psychedelic' posters, some kinds of paper such as white business cards, some types of clear mineral grease, the radiant dial of a watch or a clock and, of course, any of the dyes for dye lasers (fluorescein, relatively concentrated solutions of rhodamine 6G and so on).

"If the ultraviolet pulses are focused by a cylindrical lens to a line on the surface of the dye rather than to a point, the dye will often lase superradiantly in visible light along the direction of the line. No optical mirrors will be needed. Indeed, the nitrogen laser makes an ideal 'pump' for the dye laser. When it is employed as a pump, it opens up most of the visible spectrum to new experimental investigation [see "The Amateur Scientist," Scientific American, February, 1970].

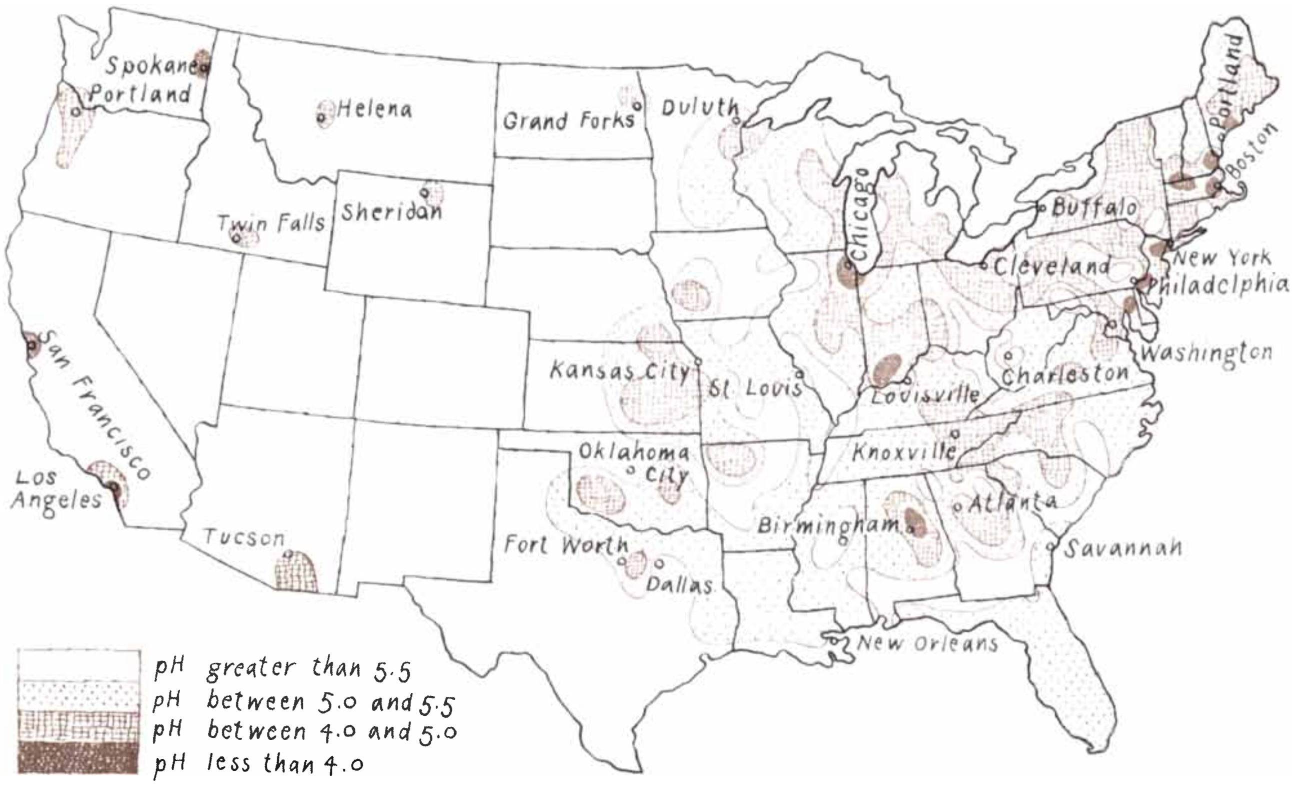
"Rumor has it that some types of Day-Glo plastic will lase when pumped with a sufficiently intense ultraviolet pulse. I have seen motion pictures of this effect but have not done the experiment myself. The smooth surfaces of the plastic appear to function as the cavity of the optical laser.

"The ultraviolet laser can readily be scaled to higher powers. A discharge path one meter long can develop an output pulse of almost a million watts, although there is a trick to it. Because the laser turns itself off so quickly, radiation does not have time to travel the full length of the column before the gain automatically drops to zero.

"This problem can be solved with a traveling-wave discharge. Move the spark gap to a corner of the capacitor. The voltage wave will then arrive first at the end of the discharge channel nearest the spark gap and will race down the channel in step with the growing pulse of emission!

"As I have mentioned, all these lasers work best when the discharge channel is filled with flowing nitrogen at low pressure. Helium can be added with almost no effect other than raising the total pressure. With a sufficiently high percentage of helium the laser will work at atmospheric pressure, thereby eliminating the vacuum pump.

"During the development of this apparatus I have enjoyed the cooperation of Norman Kunit and other members of



Acidity of rainfall in the U.S.

Ali Javan's laser group at M.I.T. I want to thank them for their spirited encouragement. I am particularly grateful to our chemist, Ray Mariella, Jr., for his suggestion of the aspirator vacuum pump.

"Finally, the experimenter must keep in mind that this is a high-powered apparatus and therefore a hazardous one. The ultraviolet emission from the laser and from the unshielded spark gap can harm the eyes. Avoid looking directly into the beam of this laser or any other one, just as you would avoid looking directly at the sun or at the arc of an electric-welding rig. Do not touch the capacitors until they have been completely discharged. Indeed, before touching any part of the apparatus make it a habit to short-circuit the spark gap with a yoke of wire supported at one end of an insulating rod about a foot long. (It is a good idea to cover the exposed high-voltage surfaces with sheets of Lucite.) Also keep in mind the fact that coherent energy, like sunlight, can be hazardous both in the direct beam and on a bounce as a specular reflection from a mirror or a smooth metal surface. Never project direct or reflected pulses into places where there may be people."

The accompanying map [*opposite page*] charts the acidity of all rain that fell in the U.S. during the two-week interval from March 15 to March 31, 1973. It shows that heavy pollution acidifies substantially all rain that falls east of the Mississippi. More important, it dramatizes the kind and quality of data that can be gathered simultaneously throughout a vast area by enlisting the enthusiastic cooperation of grammar school students (16,000 of them in this case).

The study was conceived and organized by Aaron E. Klein, who is editor of *Current Science*, a weekly publication that reaches a million grammar school children. A safe, simple and inexpensive technique whereby youngsters can measure the acidity of freshly fallen rain within an accuracy of *p*H .2 was suggested by Walter Scott Houston, a staff member of *Current Science*. The study was directed locally by teachers. The results were forwarded to Frederic Godshall of the National Oceanic and Atmospheric Administration for processing by computer. The study appears to have set some kind of record for number of amateurs simultaneously involved in a scientific project. Certainly it demonstrates the worth of a national resource that ought not to be left to languish for want of appropriate projects.

Early civilization... astronomy... the brain revolution... particle physics... electrodynamics... history of science... evolution... math...

to keep up with everything that's happening in every scientific area that interests you— get set now for a great summer reading spree!

Take any 3 books all for only $3.95

(values to $51.95) as a member of The

Library of Science

(if you will join now and agree to accept only 3 more books—at handsome discounts— within the next 12 months)

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

The Library of Science 2-40A
Riverside, New Jersey 08075

Please accept my application for membership and send me the three volumes indicated, billing me only $3 95 for all three. I agree to purchase at least three additional Selections or Alternates during the first year I am a member at special members prices Savings range up to 30% and occasionally even more My membership is cancelable any time after I buy these three books A shipping and handling charge is added to all shipments **Send no money. Members are billed when books arrive.**

3 books for $3.95

[ ] [ ] [ ]

Some expensive books count as 2 selections

Name ____________________

Address ____________________

City ____________ State ________ Zip ________

Book selections purchased for professional purposes may be a tax-deductible expense (Offer good in Continental U S and Canada only Prices slightly higher in Canada.)

56530. ISLAND AT THE CENTER OF THE WORLD. *Father Sebastian Englert* Easter Island's mysterious giant statues, the islanders' puzzling past, their lost language — new light is shed on them all here. Lavishly illustrated in full color **$15.00**

73270. RACE. *John R. Baker* Provocative exploration of biology, paleontology, history, the ancestry of man, his ascent to civilization, and the development of intelligence and cognitive ability A major work on an explosive subject. **$15.00**

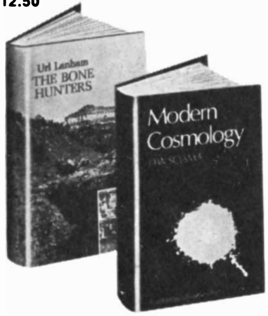
63061. MOONS AND PLANETS. *William K. Hartmann.* "A masterpiece of factual information. Deserves the widest circulation among astronomers, geologists, meteorologists, biologists, philosophers, and laymen alike because it has no equal in linking these disciplines together." —Clyde W Tombaugh in *Mercury.* **$12.95**

36520. THE BONE HUNTERS. *Url Lanham.* Absorbing history of the nineteenth-century adventurer-scientists who startled the scientific world, shattered many near-sacred illusions, and laid the foundations of modern paleontology. Filled with rare old photographs. **$12.95**

33470. THE ANATOMY OF HUMAN DESTRUCTIVENESS. *Erich Fromm.* The first viable and scientifically justifiable theory of human destructiveness. The latest findings in ethology, neurophysiology, behaviorism, paleontology, anthropology, and psychoanalysis — all are integrated into its structure. **$10.95**

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

52080. A GUIDE TO THE MEDICINAL PLANTS OF THE UNITED STATES. *Arnold and Connie Krochmal.* The definitive work on medicinal plants and their uses. Will show you how to identify them, collect them, and even raise them. Profusely illustrated. **$12.50**



(Retail prices shown)

62801. MODERN COSMOLOGY. *D. W Sciama.* "From the standpoint of readability, orderliness, continuity evenness of level, and its ability to hold one's attention, the book is three standard deviations above the mean It is almost unbelievably up to date."—*Physics Today* **$9.95**

36810. THE BRAIN REVOLUTION. *Marilyn Ferguson.* Exciting new survey reveals the incredible discoveries that are being made in the field of brain research — discoveries that you can use to improve your life and health. **$9.95**

87160. WHY JOHNNY CAN'T ADD. *Morris Kline.* The "New Math"— brilliant success or dismal failure? Leading mathematician maintains that it is time to re-examine the mathematics curriculum — and to change it. A very persuasive book. **$6.95**

48630. THE FIRST GREAT CIVILIZATIONS. *Jacquetta Hawkes.* A distinguished English archaeologist has written an extraordinary history of the great civilizations of Mesopotamia, the Indus Valley, and Egypt. Big and lavishly illustrated. **$12.50**

75730. SCIENCE AND SOCIETY IN THE 16TH AND 17TH CENTURIES. *Alan G. R. Smith.* A lucid and succinct account of the scientific advances of the sixteenth and seventeenth centuries which produced a revolutionary reshaping of man's view of the universe and his own place in it. Profusely illustrated in color and black and white. **$8.95**

87540. WORKING WITH NATURE. *John W Brainerd.* Offers perceptive instructions for modifying the environment anywhere. For nonspecialists and professionals — whether you have a window box, a back yard, or an entire wilderness to work with **$15.00**

50110. GENETICS OF THE EVOLUTIONARY PROCESS. *Theodosius Dobzhansky.* The definitive work on the modern synthetic theory of evolution by the world's foremost geneticist. Recent discoveries in molecular biology are incorporated to explain the unity and diversity of life on earth. **$12.50**

38320. CLASSICAL MECHANICS: A Modern Perspective. *Vernon D Barger and Martin G. Olsson.* Provides a thorough understanding of the principles of mechanics by exploring topics of widespread popular interest drag-racing, archery, space missions, sky-diving, etc. Not only builds theoretical and mathematical proficiency, but is fascinating to read. **$11.95**

61210. MATHEMATICAL THOUGHT FROM ANCIENT TO MODERN TIMES. *Morris Kline.* An enormous (over 1,200 pages) and hugely entertaining and enlightening history, from the ancient beginnings of mathematics to concepts that acquired significance *after* 1700. Rich in fact and anecdote. *Counts as 2 of your 3 books* **$35.00**

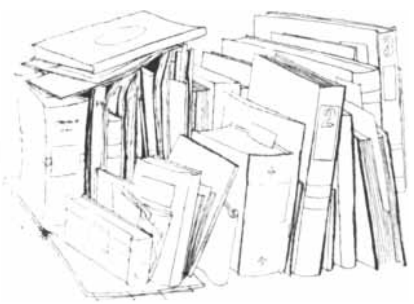
49410. FUNDAMENTALS OF ELEMENTARY PARTICLE PHYSICS. *Michael J. Longo* The author, himself a high-energy experimentalist, offers a concise, but serious account of this exciting new field. Well illustrated with diagrams and photographs. **$9.95**

48700. FOOD IN HISTORY. *Reay Tannahill.* Highly-acclaimed, pioneering work drawing on archaeology, anthropology, biology, economics, medicine, and other disciplines to show how diet and eating habits have shaped man's development **$15.00**

76120. THE SCIENCE OF LIFE. *Paul A. Weiss.* Brilliant research biologist and educator has distilled the knowledge he has gleaned over a lifetime devoted to investigating the meaning, mystery, and mechanisms of life to develop a sound and sensible philosophy **$7.95**

34560. ASTROPHYSICAL CONCEPTS. *Martin Harwit.* The author, the Chairman of the Department of Astronomy at Cornell University, offers a challenging new presentation of the fundamentals of this important science, which bridges chemistry, physics, biology, mathematics, astronomy, meteorology, and the earth sciences. **$14.95**

53720. HOW TO BUILD AND USE ELECTRONIC DEVICES. *Stuart A. Hoenig and F Leland Payne.* In simple, step-by-step fashion shows where to buy op-amps, how to use them to create anything from a 12-volt battery charger to an analog computer that can solve actual problems. A lucid self-teaching text in electrical principles, circuit theory, and simple instrumentation. **$9.95**



BOOKS

The social study of science and scientists, communication satellites and earthworms

by Philip Morrison

THE SOCIOLOGY OF SCIENCE: THEORETICAL AND EMPIRICAL INVESTIGATIONS, by Robert K. Merton. Edited and with an introduction by Norman W. Storer. The University of Chicago Press ($12.50). SOCIAL STRATIFICATION IN SCIENCE, by Jonathan R. Cole and Stephen Cole. The University of Chicago Press ($12.50). ORIGINALITY AND COMPETITION IN SCIENCE: A STUDY OF THE BRITISH HIGH ENERGY PHYSICS COMMUNITY, by Jerry Gaston. The University of Chicago Press ($10.95). The men and women the world over who work as scientists number roughly one million. There are only a few thousand Hopi, a few tens of thousands of Bushmen, and yet their ways of life are certainly better analyzed. These three books present what is clearly a growing trend: systematic study by social science of science and its scientists.

The books are related but distinct. Professor Merton has been one of the founders of the study of scientists as science; the volume reviewed here presents 22 of his papers in the field from 1935 to 1972, in a careful and extensive editorial context. The essays not only exhibit a diverse and penetrating analysis and a deal of historical and contemporary examples, with concrete numerical data, but also make genuinely good reading because of the wit, the liveliness and the rich learning with which Merton writes. The two smaller books are up-to-date, rather statistical studies, all-but-direct tests of some generalizations deriving from Merton's work and therefore narrower. Their flavor derives a tang of reality from the sharp questions they have asked of living people, questions the reader himself might put or answer.

Scientists are fairly easy to study, and perhaps they are worth it. It is easy to study them because they are not very many, they are listed, dated and documented in many ways (degrees, dictionaries, academies) and they pour out a steady stream of signed papers stored in libraries all over the world. (Indeed, it is that stream and its exponential growth, doubling every dozen years, that can be held to define modern science, since the stream first rose in the days of Galileo and his Lynxes.) The British high-energy physicists reported on, for example, are only 220 in all, and 92 percent responded—most of them right into Gaston's tape recorder. Even Oraibi village, where the old joke has it that the nuclear family includes a mother, a father, children and an anthropologist, has hardly been more completely recorded. The Coles, for their part, took dozens of samples, mainly among U.S. physicists. The Merton papers include no such questionnaire studies, but they do present much quantitative matter drawn from the more public record.

The fabric of an entire domain of sociology has been woven in the Merton essays, which often take the form of comment on, or analysis of, the important work of others: a Pareto, a Sorokin or a Kuhn. To follow all the interwoven threads is more than a review can undertake; let a few examples stand for the rest.

The key topic of the two shorter books is an important theme of Merton's work. It is defined explicitly in one paper of 1942. There the four "institutional imperatives" are laid out that constitute the "ethos of modern science": test all claims without prejudice, share all results in the common heritage, seek no material gain and be skeptical of every old belief, as of every new result. Set against this purity and coolness, however—against the dispassionate work of the anonymous referee, the content of the impersonal paper and the scientist as "a little child before the facts"—is the everyday life of laboratory and committee room, the jealous struggle for job and grant, the contest for priority and recognition, the need for self-esteem among those who best know the worth of brilliant colleagues.

The novelist and the biographer—not to overlook Honest Jim Watson's artful self-crossed hybrid—have shown us that remarkable tension. Now the sociologists can define it, and in a way measure it. The Coles strike directly at the first commandment, universalism. They make it clear that science—mainly but not exclusively represented in the U.S. by physics—is "dominated by a small, talented elite." It is those clever people who have the prizes, the prestigious appointments, the visibility to their colleagues. They and their science are performance-oriented. Neither well-known physicians nor high-income businessmen meet production standards like those fulfilled by a successful scientist. The prolific producers of the most consequential work hold what are measurably the best jobs and the widest reputations. The community is a working meritocracy; exploitation of the little-rewarded, suppression of the obscure, substantial discrimination of any kind cannot be found within the structure. The obvious absence of women and of black scientists in the U.S. must be ascribed mainly to the iniquitous conditions of entrance; women in science do receive fewer promotions and perhaps less pay than their due, but not by much. Recognition and honors accrue to women more or less justly—once they are inside the gates.

"For unto every one that hath shall be given . . . ," runs the Gospel of Matthew. So it is in science. We all do praise famous men. The effect is real; indeed, those who have been recognized may publish even after their work has been anticipated; famous names steal the show from novice collaborators. (One scholarly book of 1928 was repeatedly cited by the name of the senior author alone; he was the 65-year-old president of the American Sociological Society, she a woman in her twenties—who went on to marry her coauthor and 25 years later succeeded him in the president's chair!) This phenomenon is tied to two others: the prevalence of multiple independent discoveries and the struggle for priority.

Merton has long studied these fascinating human dramas, which he describes richly and amusingly. He shows us, after a point-by-point study, that Lord Kelvin, for example, was involved in at least 32 multiple discoveries (based on 400 of his 661 papers—"the rest have still to be studied"), with grand competitors such as Stokes, Green and Poincaré, or worthy lesser names such as Hankel, Varley and Pfaff. And so in a way every great man—leave aside the heroes at the "rugged peak of this system" who give a name to an entire epoch, such as Newton, Darwin and Freud—can in principle be replaced by a small corps of good ones. If Ernst Chladni, the "father of modern acoustics," or Adolphe Brongniart, the "father of modern palaeobotany," had not existed, we would still have had the disciplines.

This concern for named recognition is, of course, the personal consequence of the pure social commandments. Priority struggle is nothing new. It goes back to Galileo and Newton, when it was less polite; it embroiled Henry Cavendish, a man of the most modest and self-effacing temperament, as it engulfs the notable egoists. If a man forgoes combat for his clear due, his colleagues and fellows fight it out for him. They see a moral imperative arising from the social need; only that one element of property remains in the world of science.

Now that experiment in high-energy physics involves teams of large numbers of workers, the romantic identification of great names may be beginning to fade. Perhaps the Chinese style—or is it the Du Pont style?—will take over, and corporate identity ("the CERN PS group") will come to replace the names of individuals of brilliance. The star system is still valuable for its charismatic appeal, its aid to validation of the new and its signaling effects. It will not suppress novelty while youth remains so clearly an attribute of contributors, at least in physics. Older referees accept measurably more physics papers by young authors than young referees do: the gatekeepers of publication are not gerontocrats, although they may be too competitive.

All these examinations from within cannot explain science as a whole. Little is said here about the internal logic that flows from the context of each discipline. Gaston does make it plain that within high-energy physics today the phenomenologists are the best-recognized among their colleagues, since they supply the main cohesive force. It is they who build models of testable results, out of what the theorists supply, for experimenters to examine. This single branch of physics includes five groups of specialists ranged along a "continuum of proximity to physical reality"; few other disciplines are as yet so specialized. Eventually this kind of job division must be taken into account everywhere, along with issues related to the imperatives and rewards within the social structure of science.

What makes science run? We do not yet know, in spite of the ripeness and the energy of this Mertonian sociology. Whence its large problems, its clever people, its generous support, its impact? How permanent a feature of our world will it be? Exponential growth cannot persist forever, on paper or even on magnetic tape, let alone in posts and in budgets. Merton has looked outside a bit too. In the years from 1661 to 1687 the Royal Society pursued about 200 separate investigations each year. About 10 percent of them Merton counts as being related to military technology, directly or indirectly, from ballistics to the strength of metals. Today that fraction is perhaps larger. Yet the inner logic of the subject still determines most of the scientist's work, in a context and on a scale necessarily fixed by the still greater forces of history.

When insulin is first synthesized on the plains of the Yellow River and the meson is predicted by a thinker who read not Milton but Chuang-tze, science is ecumenical. It still has some way to go: women remain chiefly outside its walls, and not all countries can yet support real novelty. Fashion will change, perhaps even the title rights of name and person will be stripped away by complex growth, but the main injunctions of the 17th-century founders will persist, one thinks, as long as the activity itself remains. The skeptical study of the world itself will perhaps be less universal in human society than the spirit in which it has been successfully done.

COMMUNICATIONS SATELLITE SYSTEMS, COMMUNICATIONS SATELLITE TECHNOLOGY, edited by P. L. Bargellini. Volumes 32 and 33 of "Progress in Astronautics and Aeronautics." Martin Summerfield, Series Editor. The MIT Press (each volume $20). "Among all organizations throwing money into space, only INTELSAT has shown the capability of getting it back with interest." So the preface (repeated in both of these companion volumes) goes; allowing for some license to neglect the benefits of other working systems to mappers, weathermen, the military and the U.S.S.R., the aphorism is nearly just. One map makes it clear: there is the world of Mercator's projection; along the line of the Equator stand 40-odd dots and circles, marking the ground points over which hang geo-

stationary satellites, all existing or planned within the next few years. Even the Canadian Eskimo towns can view an equatorial satellite if they are willing to look at an elevation a few degrees above the horizon. Only the Russian Orbita scheme, with a number of relay satellites in highly elliptical 12-hour orbits, still presents a working alternative to that unique central ring in space, neatly girdling our planet at a distance of some three diameters, within which a single satellite can be steadily viewed by a stationary antenna. The ring is a limited Newtonian resource shared by the entire world.

These two volumes between them present 42 papers on the concepts, design and components of present and future synchronous communication satellites. The papers are the work of specialists from three continents (no Russian contributors) presented at a conference held in Washington in 1972, augmented by three or four papers later commissioned to round out the work. The volumes are informal, without index, photo-offset from a variety of typed manuscripts, but they display this rapidly growing and powerful technology with authenticity. Bristling acronyms and complicated diagrams make little allowance for a general scientific reader, and yet such a reader is likely to find at least the systems volume, with its more conceptual approach, accessible in full. The companion technology volume is much more specialized, although even in it a few nuggets repay the browser.

Within less than a decade the bulk of transoceanic telephony (and all transoceanic television) has become satellite-borne. By mid-1973 the Canadian Anik satellites, the first geosynchronous satellites specifically designed for national domestic communications, were operating. In the next five years the number of channels aloft will increase by at least an order of magnitude. The current INTELSAT IV system now provides about 5,000 two-way transoceanic telephone channels. With the coming of thin-line networks to ships and distant islands, aircraft traffic control and educational television (all considered here), growth seems to have just begun. The quiet whispering gallery that the ionosphere and the earth's surface provided for short-wave radio propagation is now clogged and clamorous; wire and cable are hampering and antique; the line-of-sight landlubbing microwave relay networks are too expensive. What can the engineers do in orbit?

Plenty of ideas and even a good deal of tested hardware are already available to meet the demand. Let us fish out a few. It is clear that the satellite can use spotlight-narrow beams rather than the present floodlights and thus illumine its ground correspondents more selectively. There is a remarkable new engineering, an orbital station-keeping art, that has to fight not only the tides of sun and moon and the asphericity of Mother Earth to win pointing precision but also the slow, steady effects of solar-radiation pressure; if these effects were ignored, the sunlight alone would eventually spin the spacecraft wildly. One can arrange to dump unwanted angular momentums into flywheels carried aboard the craft. Keeping place in orbit requires the actual emission of momentum, which is now released in the form of gas or rocket flames; in the future this dynamical effluent may be in the form of ionized mercury vapor, pulses of plasma sparked out of a block of inert stuff (even out of Teflon) or a colloid of charged glycerine droplets.

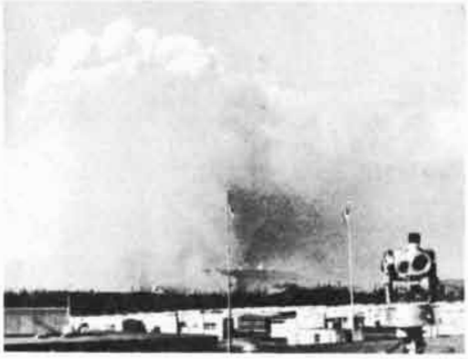
Rapid cyclical switching of interconnections among all the spot beams of frequency channels, pair by pair, each held open for a fraction of the time, allows for one kind of automatic switchboard. As an alternative, timed digital pulses, with special addresses orchestrating all spot beams at once (with the timing corrected every few seconds by noting the moment of actual response, to take care of drift in spatial position and in the clocks), offer a flexible system, easily reprogrammed to meet changes in need. Each earth station would have a time slot 100 microseconds long in one practical scheme, repeating its signal bursts once every millisecond or so.

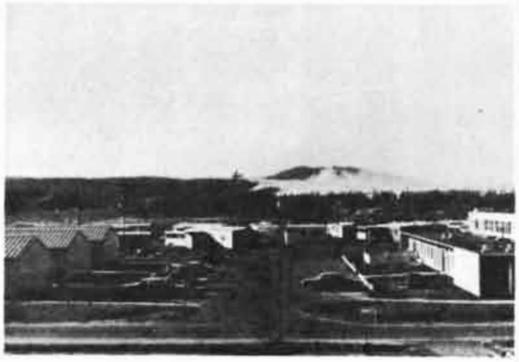
The microwave wavelengths now being used are two or three inches. Newly authorized bands include wavelengths about half as long, which provide about four times the present bandwidth, so that exploiting the two polarization directions would meet the order-of-magnitude gain that is now directly foreseen. Raindrops scatter these proposed bands badly, however, and so earth stations will probably come to be built in pairs, linked by cable or microwave, spaced far enough apart so that no single heavy rainstorm is likely to blanket both of them.

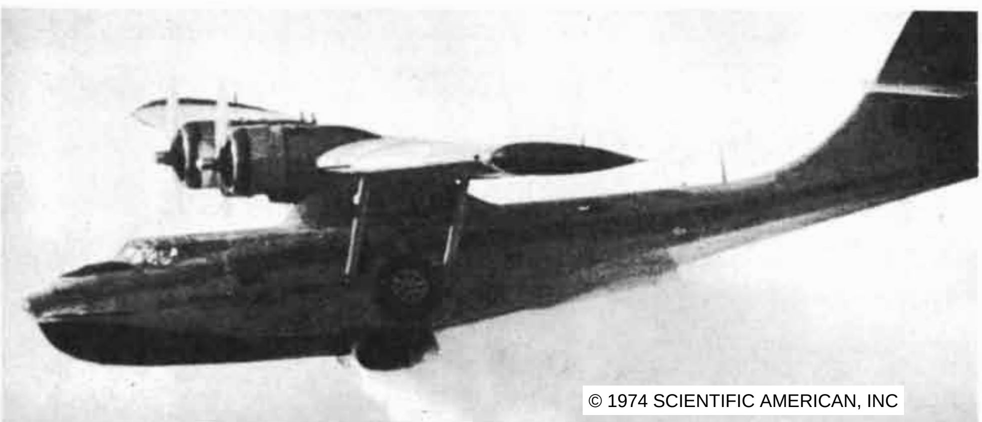
Aside from all these future plans, the government of India will soon try out a particularly bold scheme. The experimental NASA synchronous satellite ATS-F, stationed over the horn of Africa, would be lent to India to receive television transmissions from Indian earth stations and relay them directly to 2,500 community television receivers, each located in a village near a power line. These television sets would be almost conventional but would have a special front end of Indian design and manufacture and would be fed by a 10-foot stationary dish antenna. The installation of these sets is now under way in six states; the first programs should flow late next year in more than half a dozen languages. Who can guess the consequences of television in that archipelago of a million isolated villages if INSAT, the projected Indian-owned successor to the trial with ATS-F, can get under way at full scale? Such a system, described in a very interesting design paper by an Indian author with two American partners, could handle at once a relay scheme for a national urban television network, a direct-reception television and radio system for the villages, heavy telecommunications traffic among the four major cities in India and reliable telephone connections to remote places, from the Andaman Islands to the high Himalayas. INSAT, "a major undertaking, not only technologically but also in the socioeconomic sense," would carry hope into orbit. It would be no petty memorial to its untiring and enthusiastic advocate, the late Vikram Sarabhai of Ahmedabad.

BIOLOGY OF EARTHWORMS, by C. A. Edwards and J. R. Lofty. Halsted Press Division of John Wiley & Sons, Inc. ($8.75). Earthworm research has doubled about every 20 years for the past century. Here two expert soil biologists from the renowned English agricultural research station at Rothamsted cite some 600 papers in their brief and admirably readable undergraduate-level survey of these familiar soil animals. Earthworms are members of a single order of the annelid phylum, with some 1,800 species distributed all over the world (although hardly in deserts, on mountains or in snow and ice).

The form of earthworms is the topic of the first chapter, well illustrated by scanning electron micrographs of the creatures, particularly of the front end and of the genital bristles that help to hold the overlapping, copulating worm pair together. These hermaphroditic animals have male and female apertures, both double, and the pores and bristles play an important part in the classification of species. The sense organs of the earthworms are of two classes: photoreceptors (not recognizable eyes but little organelles with lenslike structures under the skin) and groups of epithelial touch-sensitive cells with sensory hairs that project through the skin. Internal







"strain gage" cells are found in the muscle layers.

Most of the worms leave their eggs in "cocoons" near the soil surface. These cocoons are actually short encircling tubes that form around the worm and harden after copulation; fertilization occurs in the cocoon as it slides off past the head of its backward-moving parent. We know little of the life cycle of even the common species of earthworms, but it seems that cocoons are produced seasonally; the warmer the month, the more plentiful the cocoons. Most species of worms mature in from a few months to a year or two; some individuals have survived in the laboratory for 10 years, but "in the field they are unlikely to attain such ages."

There are three pages on learning in earthworms (but not one sentence about their genetics!). Earthworms are very sensitive to touch and respond to electrical, acoustical and vibratory stimuli by surfacing. In one experiment with electric current they became "U-shaped with both ends directed towards the cathode." The worms are largely water, some 85 percent or more by weight; their covering layers cannot prevent water loss, so that drying out is a serious hazard even though they can recover from a loss of three-fourths of their body weight.

Acid soils repel the worms, which can detect *p*H more readily than they can salt or sugar. Seawater is deadly to them, but even so some species are widely distributed (perhaps by continental drift?). European species have traveled along with human colonizers, to dominate the local species in Chile, New Zealand, parts of India, Australia and southwestern Africa. In the U.S. too most species are European settlers. Glacial regions generally have no endemic earthworms.

At least as far back as Charles Darwin earthworms have been viewed as benefactors of the soil. In Cambridgeshire a grassy orchard may house a couple of hundred grams of living earthworms per square meter, most of them in the top few inches of soil, where most of the cocoons are found. The worms go deeper in cold or dry periods, becoming quiescent or even rolling into a ball covered with mucus. They eat their way through the soil, passing through their gut the enormous amounts of soil first admired by Darwin. His data seem fair enough today; he estimated that earthworms turned over a layer five millimeters deep every year in English pastures. Their work output may be even 10 times as much in some tropical soils. It is the initial breaking up of the tougher plant residues such as leaves, stems and roots that the earthworms perform. It is reckoned that in a deciduous woodland of the Temperate Zone they "consume the annual leaf fall in about three months."

Earthworms are a delicacy among the Maori. No recipes are given, but the last chapter of this unusual and authoritative text is a compact guide to culture, field experiments, observation cages, learning studies and other simple means of investigating these familiar, important and somehow mysterious downstairs neighbors and partners of agricultural man.

CUT MY COTE, by Dorothy K. Burnham. Royal Ontario Museum, Toronto ($2). "I shall cut my cote after my cloth" is a proverb that was already old in 1542. In this original and handsome little catalogue of an exhibition of the textile department of a famous museum the author makes clear how true the adage is. The weavers achieved mastery long before the tailors, and so garments were first made with straight cuts, little waste and much use of selvage. The loom width became a habit: long after materials made on very different looms offered a real choice of cloth widths, the cut of the garment remembers the throw of an old shuttle.

The weavers of classical antiquity stood at a vertical loom and worked down. Their wide, warp-weighted cloth made the chiton, which was used without any cutting or sewing by men and women alike in different lengths. The Roman citizen's toga was a piece of woolen with the lower corners cut off to aid the draping. The poncho of American weavers was woven in the specific size to make the garment, the last weft threads forced in by hand to make a cloth finished on all four sides.

Once the warp-weighted loom died out in Europe, the cloth (exceptions are remarkably special) took a width dependent on the reach of the weaver, who sat at the horizontal loom throwing the shuttle from hand to hand. The comfortable limit is about 50 inches; often the cloth is much narrower. The study of undergarments, of baby clothes and of church vestments reveals the oldest forms, often covered up or modified by a millennium or two of everyday life and fashion.

Many readers will want more of this remarkable narrative. The attractive drawings and photographs of this 34-page booklet are a promise, we can hope, of a general volume to come; so far only a few technical monographs can be cited. Costume history is coming at last to merge into the wide stream of the history of technology.

INDEX OF ADVERTISERS

JUNE 1974

ALLIED CHEMICAL CORPORATION 49
Agency: J. S. Lanza & Associates

CALIFORNIA COMPUTER PRODUCTS, INC. 57
Agency: Dailey & Associates Advertising

CANON USA, INC. 127
Agency: Dentsu Corporation of America

CARDAVON PRESS 127
Agency: Hopkins & Associates

CHRYSLER CORPORATION, DODGE DIVISION . . . . Inside Front Cover
Agency: Batten, Barton, Durstine & Osborn, Inc.

DATA GENERAL 14, 15
Agency: Scali, McCabe, Sloves, Inc.

DELTA AIR LINES, INC. 17
Agency: Burke Dowling Adams, Inc.

EASTMAN KODAK COMPANY 47
Agency: Rumrill-Hoyt, Inc.

EDMUND SCIENTIFIC COMPANY 133
Agency: Chittick Advertising

GENERAL MOTORS CORPORATION, BUICK MOTOR DIVISION/OPEL 80, 81
Agency: McCann-Erickson, Inc.

GENERAL MOTORS CORPORATION, GMC TRUCK & COACH DIVISION . . . . Back Cover
Agency: McCann-Erickson, Inc.

HEWLETT-PACKARD 10, 11
Agency: Corporate Marketing Communications

HUGHES AIRCRAFT COMPANY 13
Agency: Foote, Cone & Belding Advertising, Inc.

INTERNATIONAL TELEPHONE AND TELEGRAPH CORPORATION 9
Agency: Needham, Harper & Steers, Inc.

KONICA 115
Agency: Kameny Associates, Inc.

LIBRARY OF SCIENCE 128
Agency: Henderson & Roll, Inc.

LITTLE, BROWN & COMPANY 133
Agency: Franklin Spier Incorporated

LUFTHANSA GERMAN AIRLINES (Int.) . . . . Inside Front Cover
Agency: H. K. McCann Company mbH

MCDONNELL DOUGLAS CORP. 53
Agency: J. Walter Thompson Company

MICROZOO 127
Agency: Lewellen, Inc.

MINOLTA CORPORATION 51
Agency: E. T. Howard Company, Inc.

NATIONAL CAMERA, INC. 4
Agency: Langley Advertising Agency

NATIONAL CAR RENTAL SYSTEM 2
Agency: Richard K. Manoff, Inc.

NEW YORK STATE DEPARTMENT OF COMMERCE 115
Agency: Rumrill-Hoyt, Inc.

NEWSWEEK BOOKS 54, 55

PARKER PEN COMPANY, THE . . . . Inside Back Cover
Agency: J. Walter Thompson Company

POLAROID CORPORATION 67, 68, 69
Agency: Doyle Dane Bernbach, Inc.

PRINCE HOTEL, THE 5
Agency: Base Hamilton Edwards Ltd.

QUESTAR CORPORATION 134

SCANDINAVIAN AIRLINES SYSTEM (Int.) 9
Agency: Sky Market Promotion AB

SHEAFFER WORLD-WIDE, A TEXTRON COMPANY 95
Agency: Sperry-Boom, Inc.

SIERRA CLUB BULLETIN 121
Agency: Burger, Felix & Wood

SIMULATIONS PUBLICATIONS, INC. 131
Agency: The Advertising Workshop, Inc.

SPRINGER-VERLAG NEW YORK 16, 132

TEACHERS COLLEGE PRESS 4

TRW SYSTEMS GROUP 70
Agency: Inter/Media, Inc.

U. S. PIONEER ELECTRONICS CORP. 1
Agency: Philip Stogel Company, Inc.

WESTERN UNION CORPORATION 6, 7
Agency: Ries Cappiello Colwell, Inc.

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

PSYCHIATRISTS AND THE ADVERSARY PROCESS

Durham v. United States, Vol. 94, U.S. Court of Appeals, District of Columbia, page 228, and Vol. 214 Federal Reporter, 2d Series, page 862, 1954.

Jenkins v. United States, Vol. 113, U.S. Court of Appeals, District of Columbia, page 300, and Vol. 307 Federal Reporter, 2d Series, page 637, 1962.

Rouse v. Cameron, Vol. 125, U.S. Court of Appeals, District of Columbia, page 366, and Vol. 373 Federal Reporter, 2d Series, page 451, 1966.

United States v. Brawner, Vol. 153, U.S. Court of Appeals, District of Columbia, page 1, and Vol. 471 Federal Reporter, 2d Series, page 969, 1972.

FUSION POWER BY LASER IMPLOSION

Fusion by Laser. Moshe J. Lubin and Arthur P. Fraas in *Scientific American*, Vol. 224, No. 6, pages 21–33; June, 1971.

Prospects for Rocket Propulsion with Laser-induced Fusion Microexplosions. R. Hyde, L. Wood and J. Nuckolls. AIAA/SAE 8th Propulsion Joint Specialist Conference, AIAA Paper No. 72–1063, American Institute of Aeronautics and Astronautics, 1972.

Laser Compression of Matter to Super-high Densities: Thermonuclear (CTR) Applications. John Nuckolls, Lowell Wood, Albert Thiessen and George Zimmerman in *Nature*, Vol. 239, No. 5368, pages 139–142; September 15, 1972.

Power from Laser-initiated Nuclear Fusion. Keith Boyer in *Astronautics & Aeronautics*, Vol. 11, No. 8, pages 44–49; August, 1973.

POPULATION CYCLES IN RODENTS

Voles, Mice and Lemmings. C. S. Elton. Clarendon Press, 1942.

The Lemming Year. Walter Marsden. Chatto and Windus Ltd., 1964.

Population Cycles in Small Rodents. Charles J. Krebs, Michael S. Gaines, Barry L. Keller, Judith H. Myers and Robert H. Tamarin in *Science*, Vol. 179, No. 4068, pages 35–41; January 5, 1973.

Population Cycles in Small Mammals. C. J. Krebs and J. H. Myers in *Advances in Ecological Research: Vol. 8*. Academic Press, London, 1974.

NEUROTRANSMITTERS

The Uptake and Storage of Noradrenaline in Sympathetic Nerves. L. L. Iversen. Cambridge University Press, 1967.

Biogenic Amines and Emotion. Joseph J. Schildkraut and Seymour S. Kety in *Science*, Vol. 156, No. 3771, pages 21–30; April 7, 1967.

Biochemistry of Catecholamines. Perry B. Molinoff and Julius Axelrod in *Annual Review of Biochemistry*, Vol. 40, pages 465–500; 1971.

Noradrenaline: Fate and Control of Its Biosynthesis. Julius Axelrod in *Science*, Vol. 173, No. 3997, pages 598–606; August 13, 1971.

The Pineal Gland: A Neurochemical Transducer. Julius Axelrod in *Science*, in press.

THE STEADY STATE OF THE EARTH'S CRUST, ATMOSPHERE AND OCEANS

Geologic History of Sea Water: An Attempt To State the Problem. William W. Rubey in *Bulletin of the Geological Society of America*, Vol. 62, No. 9, pages 1111–1148; September, 1951.

Silica-Bicarbonate Balance in the Ocean and Early Diagnosis. F. T. MacKenzie and R. M. Garrels in *Journal of Sedimentary Petrology*, Vol. 36, pages 1075–1084; 1966.

Sedimentological Consequences of a Steady-State Ocean-Atmosphere. Raymond Siever in *Sedimentology*, Vol. 11, No. 1/2, pages 5–29; October, 1968.

Earth. Frank Press and Raymond Siever. W. H. Freeman and Company, 1974.

ROTATING CHEMICAL REACTIONS

Concentration Wave Propagation in Two-dimensional Liquid-Phase Self-oscillating System. A. N. Zaikin and A. M. Zhabotinsky in *Nature*, Vol. 225, pages 535–537; February 7, 1970.

Explanation of Spatial Band Propagation in the Belousov Reaction. Richard J. Field and Richard M. Noyes in *Nature*, Vol. 237, No. 5355, pages 390–392; June 16, 1972.

Autowave Processes in a Distributed Chemical System. A. M. Zhabotinsky and A. N. Zaikin in *Journal of Theoretical Biology*, Vol. 40, pages 45–61; 1973.

Scroll-shaped Waves of Chemical Activity in Three Dimensions. Arthur T. Winfree in *Science*, Vol. 181, pages 937–939; September 7, 1973.

Rotating Solutions to Reaction-Diffusion Equations. Arthur T. Winfree in *SIAM-AMS Proceedings: Vol. 8*, edited by Donald S. Cohen. American Mathematical Society, in press.

ICE-AGE HUNTERS OF THE UKRAINE

The Old Stone Age. F. Bordes. McGraw-Hill Book Company, 1968.

Environment and Archaeology: An Ecological Approach to Prehistory. Karl W. Butzer. Aldine Publishing Company, 1971.

Ice-Age Hunters of the Ukraine. Richard G. Klein. The University of Chicago Press, 1973.

WINES, GRAPE VINES AND CLIMATE

Éléments de Viticulture Générale. J. Branas, G. Bernon and L. Levadoux. Montpellier, 1946.

La Séléction et l'Hybridation chez la Vigne. L. Levadoux. Montpellier, 1951.

Cépages et Vignobles de France. P. Galet. Montpellier, 1956–1962.

American Wines and Wine-making. Philip M. Wagner. Alfred A. Knopf, 1974.

MATHEMATICAL GAMES

The Great Pyramid in *Fads and Fallacies in the Name of Science*. Martin Gardner. Dover Publications, Inc., 1957.

Secrets of the Great Pyramid. Peter Tompkins. Harper & Row, Publishers, 1971.

THE AMATEUR SCIENTIST

A Simple Pulsed Nitrogen 3371 Å Laser with a Modified Blumlein Excitation Method. J. G. Small and R. Ashari in *The Review of Scientific Instruments*, Vol. 43, No. 8, pages 1205–1206; August, 1972.

It's as easy to live with on the road as it is standing still.

Which is why a GMC MotorHome looks the way it does. Aerodynamic. And low—lower than any other motorhome. So it can handle crosswinds. And corners.

That's also why we've provided it with front-wheel drive, a 455-cubic-inch V-8 with automatic transmission, and tandem rear wheels with air suspension. So it can handle highway speeds. And off-highway bumps.

Not that we've ignored living comfort. In fact, 15 floor plans, in four different color-coordinated decors, are available.

If you want specifics, send for our 28-page color catalog. Simply write: GMC MotorHome Headquarters, Drawer Y, Dept. 311, Lansing, Mich. 48909.

Better yet, see your local GMC MotorHome dealer.

The GMC MotorHome. It doesn't look like a box. It doesn't ride like a truck.

GMC

The MotorHome from General Motors.