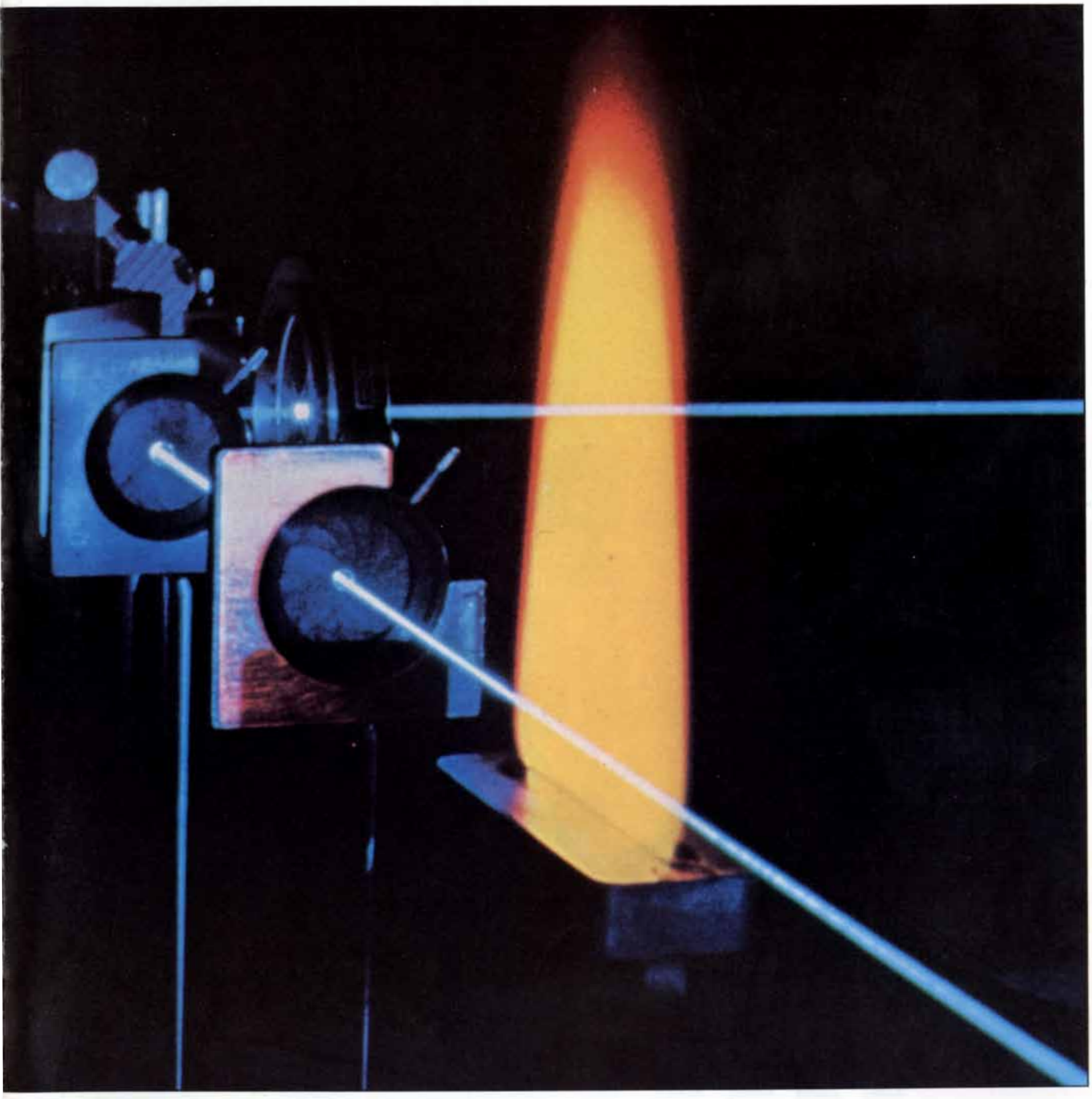
SCIENTIFIC AMERICAN



FLAME CHEMISTRY

$2.00

February 1982

INTRODUCING THE FORERUNNER OF IMITATIONS TO COME.

THE NEW BMW 528e.

It seems the automotive establishment has caught up with automotive enthusiasts in their affection for BMW's.

Lately, affection has become ardor—as witnessed by the number of cars resembling BMW's, comparing themselves to BMW's, and even claiming to perform "like" BMW's.

BMW shares this vision of what a car should be in the eighties, with one major disclaimer: cars should be motivated by foresight, not hindsight. And no better example exists than the new, technologically advanced BMW 528e.

THE 1982 CAR THAT WON'T BE IRRELEVANT BY 1987.

The new 528e is a luxury performance sedan that escapes further categorization largely because its category doesn't exist yet.

It is the sort of car other car makers, or those of them who value performance, will someday have to attempt—one that turns minimum energy into maximum performance.

The 528e accomplishes this through an inspired paradox called the "Eta" engine.

It runs slower than conventional 6-cylinder engines, stretching fuel in the process. Yet at the same time, it develops higher torque (or power) at those slower engine speeds—speeds at which the car is most often driven.

The result is the sort of exhilarating engine response you would expect if fuel efficiency were no obstacle; because in the 528e, it isn't.

Deep within the engine, where the important fuel efficiency decisions are made, another BMW innovation makes sure those decisions are always made in favor of maximum performance.

Digital Motor Electronics is an electronic system that constantly monitors the engine and assures that fuel ignition occurs at the optimum moment. It also cuts off fuel flow to cylinders when they're not needed, allowing the car to run on its own momentum during deceleration, saving fuel for acceleration.

All this makes for a car that performs extraordinarily well because of the fact that it's extraordinarily efficient, not in spite of it. Which helps explain its mileage figures, unsurpassed by any gasoline-powered car in its price class: EPA-estimated [22] mpg, 32 mpg highway.*

ONE OF THE WORLD'S MOST OBEDIENT AUTOMOBILES.

The eighties are already being touted as the era of the personal car, a time when driver regains mastery over automobile.

As trailblazers forge ahead in this direction, they'll find the trail has been previously blazed by the 528e.

Its suspension is a refinement of a design so advanced, Car and Driver termed it "the single most significant breakthrough in front suspension design in this decade."

The resulting agility is complemented by its steering, highly responsive in the manner of all BMW's (a legacy of its racing forebears). On the road, it is one of the most obedient cars in the world.

Inside, it is a model of the "new" science of ergonomics, a recent discovery with some car makers and a tradition at BMW.

The interior is an elegant continuum of driver and automobile created by easily accessible controls, contoured seats and many other functional amenities.

But despite all this, no claim is made for the 528e as the Car of the Future.

It is not a car designed for mass-market accolades, but rather for a finite number of serious drivers—enthusiasts who've watched the performance characteristics of cars diminish in tandem with the world's fuel supply.

For those people, the 528e may well be a deliverance.

You may judge for yourself at your BMW dealer, where the 528e awaits your test drive.

THE ULTIMATE DRIVING MACHINE.

BMW, MUNICH, GERMANY.

*Figures are for comparison purposes only. Actual mileage may vary, depending on speed, weather and trip length; actual highway mileage will most likely be lower.

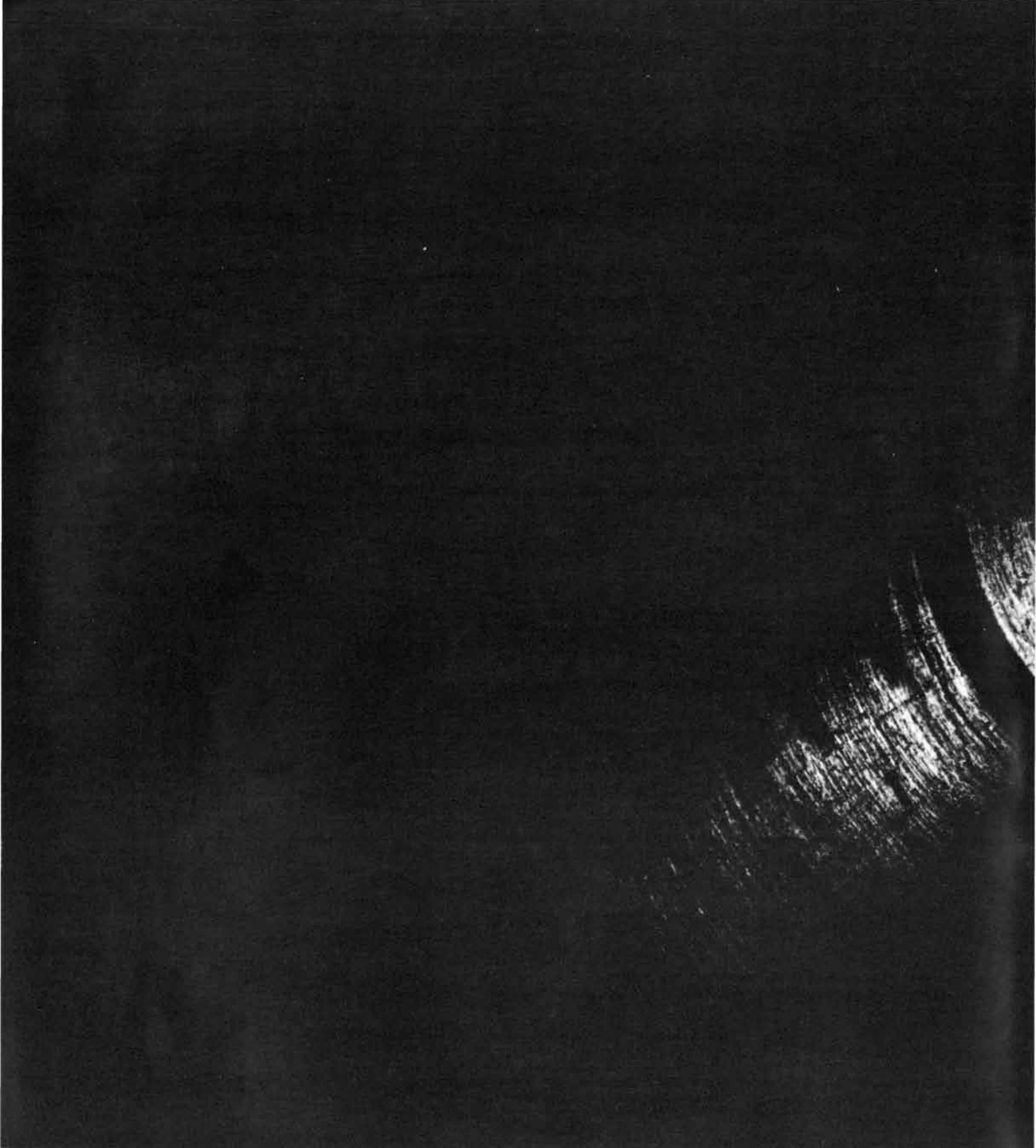
86 PROOF BLENDED SCOTCH WHISKY DISTILLED AND BOTTLED IN SCOTLAND. IMPORTED BY THE BUCKI GHAM CORPORATION, NEW YORK, N.Y.
A LOT OF US HAVE WILD IDEAS. HERE'S TO THOSE WHO GET THEIRS OFF THE GROUND.
For 105 years, attempts have been made to cross the Atlantic by balloon. All of them failed.
Until in 1978, on their second try, Maxie Anderson and his two partners astounded the world. They made it.
The following year, he and his son, Kris, (who holds the world hot air balloon distance record), piloted the first balloon to fly across the North American continent. Now, Maxie is preparing for the ultimate adventure: a 10-day, 20,000 mile flight around the globe.
What the Andersons will try next is anyone's guess. But one thing is certain. Whenever these two daring, unpredictable balloonists get together, they do something very predictable. They pour themselves a glass of their favorite Scotch, Cutty Sark. And they start planning the newest mission impossible.
Maxie and Kris Anderson
Littick
The Scotch with a following of leaders.
CUTTY SARK®

Established 1845 SCIENTIFIC AMERICAN February 1982 Volume 246 Number 2

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), 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: $21 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $27 PER YEAR, ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.



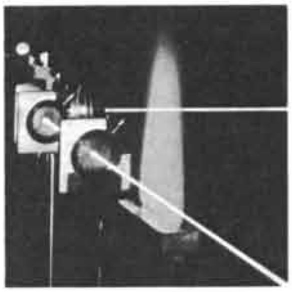
LISTENING CAN IMPROVE YOUR VISION.

Few, we've found, see as far or as clearly as those who listen well.

Good listeners think more broadly– because they hear and understand more facts and points of view.

They make better innovators. Because listeners look at problems with fresh eyes, combine what they learn in more unlikely ways, they're more apt to hit upon truly startling ideas.

Ultimately, good listeners attune



THE COVER

The photograph on the cover shows an experimental setup for studying a flame by means of a laser probe; one of the advantages of this diagnostic method is that it does not disturb the flow pattern of the flame gases. The flame is classified as a laminar premixed flame, that is, a smooth flame in which the fuel (in this case methane) is mixed with air before entering the combustion zone. A blue laser beam is passed through the flame, exciting a green fluorescence (not visible here) from large aromatic (ring-shaped) hydrocarbon molecules in the flame; such molecules may be precursors of soot. The formation of soot is currently one of the most actively studied processes in flame chemistry (see "The Chemistry of Flames," by William C. Gardiner, Jr., page 110). The photograph was made at the Fire Research Center of the National Bureau of Standards.

THE ILLUSTRATIONS

Cover photograph courtesy of Jim Sonnichsen, National Bureau of Standards

Page	Source
41–49	Jerome Kuhl
51	John Reader, © National Geographic Society
52	Richard L. Hay, University of California at Berkeley
53–55	Andrew Tomko
56	Mary D. Leakey
57	Richard L. Hay, University of California at Berkeley
59	Carl-Henrik von Bonsdorff, University of Helsinki
60–61	Bunji Tagawa
62	Kai Simons, Henrik Garoff and Jürgen Kartenbeck, European Molecular Biological Laboratory; Ari Helenius, Yale University
63–66	Bunji Tagawa
99–100	Jet Propulsion Laboratory, California Institute of Technology
101	Walken Graphics
104	Jet Propulsion Laboratory, California Institute of Technology
105–109	Walken Graphics
111	Fritz Goro
112	Fritz Goro (*top*), Ian Worpole (*bottom*)
113–123	Ian Worpole
124	Robin Stevenson, General Motors Research Laboratories
127	Malcolm G. Smith, Royal Observatory, Edinburgh
128–130	Alan D. Iselin
131	Cerro Tololo Inter-American Observatory
132	© 1960 National Geographic Society–Palomar Sky Survey, reproduced by permission of the California Institute of Technology (*top*); Alan D. Iselin (*bottom*)
133–138	Alan D. Iselin
141	Joseph S. Levine, Boston College; Edward F. MacNichol, Jr., Marine Biological Laboratory, Woods Hole, Mass. (*top*); George V. Kelvin (*bottom*)
142–144	George V. Kelvin
145	Joseph S. Levine, Boston College; Edward F. MacNichol, Jr., Marine Biological Laboratory, Woods Hole, Mass. (*top left*); George V. Kelvin (*top right and bottom*)
146–148	George V. Kelvin
151	R. F. Bonifield, courtesy Addison Gallery of American Art, Phillips Academy, Andover, Mass.
154	Gabor Kiss
155	American Institute of Physics
156–157	Gabor Kiss
158	Ralph Morse
162	Jearl Walker
165–168	Michael Goodman

LETTERS

Sirs:

Why is it that violin-plate vibrations are studied on free plates rather than clamped plates? As Carleen Maley Hutchins points out ["The Acoustics of Violin Plates," SCIENTIFIC AMERICAN, October, 1981], the plates in a violin are not free at the periphery but are glued to the quite solid ribs. In spite of the fact that violin makers traditionally whittle their plates on the basis of free-periphery tap tones, it would seem that a scientific analysis of what makes a "good" plate should be done with the periphery clamped.

It is by no means clear, and I doubt that it is so, that two plates with nearly identical free-periphery vibrations would have nearly identical clamped-periphery vibrations. The local stiffness is unimportant at a free boundary but is very important at a clamped boundary. The opposite is true for the mass....

EDGAR PEARLSTEIN

University of Nebraska-Lincoln

Sirs:

Mrs. Hutchins' article was interesting but I could not fail to come to the following conclusion after reading it. From the beginning to the end of the golden era of Cremona the Italian masters produced perhaps 5,000 violins with a great variety of physical dimensions. Most of these instruments possess a distinctive tonal quality recognized as the Cremona sound. If the Cremona sound is based on the plate tunings and other adjustments of the violin, as Mrs. Hutchins advocates, out of the millions of violins of various constructions made in modern times surely at least one would have "hit" on these magical dimensions and resulted in an instrument rivaling a Stradivari in tonal quality.

This would appear to be an extremely meaningful scientific experiment in itself, with significant numbers of violins for comparison. It brings up the possibility that perhaps variations in physical dimensions alone will never yield a violin of Cremona quality. Producing more instruments with minor dimensional variations seems a fruitless avenue of research that is simply duplicating more than 100 years of negative results.

MARY ANN WATSON

Birmingham, Ala.

Sirs:

It was not without considerable thought that the long study of free plates was undertaken as a first step to understanding the situation of the plates assembled into the finished violin. Since the edge coupling of the free plates to the ribs is neither a truly clamped joint nor a hinged joint but a variable combination of the two depending somewhat on how the glue joints are made, there is little chance of getting basic information without introducing unnecessary variables inevitably present in the jointing. There are also many variations in the physical characteristics of each set of ribs and blocks that participate in the vibrations of the assembled plates, not to mention the complicated couplings of the plates with the air in the box, about which we still know very little. Faced with the fact that the early makers and master violin makers to this day are able to judge from certain characteristics of free plates a considerable amount about the tone and playing qualities of each finished instrument, I decided to pursue the mechanisms involved in the free-plate vibrations and their meaning as related to the tone and playing qualities of the finished instruments, with the results reported in my article.

In response to Ms. Watson let me say first of all there are wide variations in the tonal qualities of the "old master" violins, from the fine concert ones used by the virtuoso violinists to some that we have tested that have hardly any tonal output. The superiority of the Cremona instruments has become almost a myth; it is well known among violin dealers and makers that even the instruments of Stradivari that are in existence today have widely varying qualities.

The ones we hear most about are those with the excellent tone and playing qualities that are prized by the artists. Actually every violin made before 1800 in general use today has been considerably altered, because early in the 19th century there was a demand for more power from the strings. To achieve this the neck of the violin was lengthened about two centimeters, the angle of the neck was changed to allow for a higher bridge and the bass bar was increased in size to accommodate the extra force exerted on the top as a result of these changes. Some of the highly arched early instruments, such as those of Jacobus Stainer and the Amatis, have not responded to these changes as well as the flatter arched ones of Stradivari and the Guarneris. With the increased interest in early music many of the highly arched violins are being restored to their original condition with most satisfactory results.

Through the kindness of the late Rembert Wurlitzer I was able to test the free top and back plates of one of the fine concert Stradivarius violins. These tests show the plate modes to be essentially the same as in the best violins reported in my article. I did not report this in the article, because I feel that one such test is not enough evidence for scientific validity.

Concerning the comment that violin makers have had 100 years of negative results, may I say that there are fine violins being made by quite a few makers today that are considered equal in tone and playing qualities to the fine early instruments. Many concert artists use them and keep their valuable "early master" instruments safely locked up. An example of this was one of my violas that was in the Kroll Quartet for more than 10 years. On the jacket of one of their recordings it says, "Mr. Kroll plays the Ernst Stradivarius of 1709" and then adds, "Mr. K. uses a new viola [mine] in these performances, chosen because of its superior recording properties."

The real problem is that the two- or three-century-old instruments are gradually being played out. Wood, like any other material, is subject to fatigue. Fritz Kreisler's Guarnerius, now at the Library of Congress, has lost its tone. The Stradivarius violin of another famous virtuoso was bought from his estate by a fine violinist, who later came to me asking for a violin with some tone and power, since his was played out. In spite of the efforts of the best restorers the old instruments are going into private collections and museums where they are prized for their historical value rather than their acoustic qualities.

CARLEEN MALEY HUTCHINS

Montclair, N.J.

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

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 75 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 5969, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name ____________

New Address ____________

Old Address ____________


GM

The $149.95 personal computer.

Introducing the Sinclair ZX81

If you're ever going to buy a personal computer, now is the time to do it.

The new Sinclair ZX81 is the most powerful, yet easy-to-use computer ever offered for anywhere near the price: only $149.95* completely assembled.

Don't let the price fool you. The ZX81 has just about everything you could ask for in a personal computer.

A breakthrough in personal computers

The ZX81 is a major advance over the original Sinclair ZX80—the world's largest selling personal computer and the first for under $200.

In fact, the ZX81's new 8K Extended BASIC offers features found only on computers costing two or three times as much.

Just look at what you get:

- Continuous display, including moving graphics
- Multi-dimensional string and numerical arrays
- Mathematical and scientific functions accurate to 8 decimal places
- Unique one-touch entry of key words like PRINT, RUN and LIST
- Automatic syntax error detection and easy editing
- Randomize function useful for both games and serious applications
- Built-in interface for ZX Printer
- 1K of memory expandable to 16K

The ZX81 is also very convenient to use. It hooks up to any television set to produce a clear 32-column by 24-line display. And you can use a regular cassette recorder to store and recall programs by name.

*Plus shipping and handling. Price includes connectors for TV and cassette, AC adaptor, and FREE manual.

If you already own a ZX80

The 8K Extended BASIC chip used in the ZX81 is available as a plug-in replacement for your ZX80 for only $39.95, plus shipping and handling—complete with new keyboard overlay and the ZX81 manual.

So in just a few minutes, with no special skills or tools required, you can upgrade your ZX80 to have all the powerful features of the ZX81. (You'll have everything except continuous display, but you can still use the PAUSE and SCROLL commands to get moving graphics.)

With the 8K BASIC chip, your ZX80 will also be equipped to use the ZX Printer and Sinclair software.

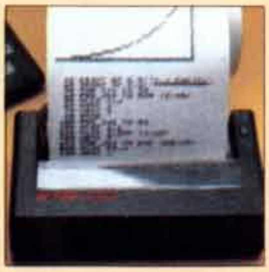
Order at no risk**

We'll give you 10 days to try out the ZX81. If you're not completely satisfied, just return it to Sinclair Research and we'll give you a full refund.

And if you have a problem with your ZX81, send it to Sinclair Research within 90 days and we'll repair or replace it at no charge.

**Does not apply to ZX81 kits.

NEW SOFTWARE: Sinclair has published pre-recorded programs on cassettes for your ZX81, or ZX80 with 8K BASIC. We're constantly coming out with new programs, so we'll send you our latest software catalog with your computer.



ZX PRINTER: The Sinclair ZX Printer will work with your ZX81, or ZX80 with 8K BASIC. It will be available in the near future and will cost less than $100.

16K MEMORY MODULE: Like any powerful, full fledged computer, the ZX81 is expandable. Sinclair's 16K memory module plugs right onto the back of your ZX81 (or ZX80, with or without 8K BASIC). Cost is $99.95, plus shipping and handling.

ZX81 MANUAL: The ZX81 comes with a comprehensive 164-page programming guide and operating manual designed for both beginners and experienced computer users. A $10.95 value, it's yours free with the ZX81.

The $99⁹⁵ personal computer.

Introducing the ZX81 kit

If you really want to save money, and you enjoy building electronic kits, you can order the ZX81 in kit form for the incredible price of just $99.95.* It's the same, full-featured computer, only you put it together yourself. We'll send complete, easy-to-follow instructions on how you can assemble your ZX81 in just a few hours. All you have to supply is the soldering iron.

How to order

Sinclair Research is the world's largest manufacturer of personal computers.

The ZX81 represents the latest technology in microelectronics, and it picks up right where the ZX80 left off. Thousands are selling every week.

We urge you to place your order for the new ZX81 today. The sooner you order, the sooner you can start enjoying your own computer.

To order, simply call our toll free number, and use your MasterCard or VISA.

To order by mail, please use the coupon. And send your check or money order. We regret that we cannot accept purchase orders or C.O.D.'s.

CALL 800-543-3000. Ask for operator #509. In Ohio call 800-582-1364. In Canada call 513-729-4300. Ask for operator #509. Phones open 24 hours a day, 7 days a week. Have your MasterCard or VISA ready.

These numbers are for orders only. For information, you must write to Sinclair Research Ltd., 2 Sinclair Plaza, Nashua, NH 03061.

sinclair

AD CODE 02SA	PRICE†	QTY.	AMOUNT
ZX81	$149.95		
ZX81 Kit	99.95		
8K BASIC chip (for ZX80)	39.95		
16K Memory Module (for ZX81 or ZX80)	99.95		
Shipping and Handling	4.95		$4.95
		TOTAL	

MAIL TO: Sinclair Research Ltd., One Sinclair Plaza, Nashua, NH 03061.

NAME ______________________

ADDRESS ______________________

CITY/STATE/ZIP ______________________

† U.S. Dollars


HE98221
1F9
HE98221
1F3
HE98221
1F3
HE98221
1F4
HE98221
1F4
HE98221
1F3
HE98221
1F3

A Single, Tiny Image from 180,000 Electrical Signals

To shrink the TV camera, start with the pickup tube

Beginning with the transistor in the early 1950s, solid-state components have been instrumental in producing compact electrical equipment by replacing bulky vacuum tubes. But unlike radios and televisions, TV cameras have been largely unaffected by this trend. Built into each conventional camera is a pickup tube that acts like an artificial retina to read images coming in through the camera lens. The intricate mechanism within its glass-enclosed vacuum allows the tube to convert these images into signals indicating shape, light intensity and color. Yet even Hitachi's own SATICON® pickup tube—the smallest of its type in the world—is almost 10 cm long. So the first step toward creating truly miniature TV and video cameras is to shrink the size of this vital component.

Hitachi's solid-state solution

Hitachi's approach to this problem was to replace the pickup tube with totally solid-state circuitry. The result of our research: a metal-oxide semiconductor (MOS) image sensor. This single-chip sensor consists of more than 180,000 photodiodes arranged in a 485× 384 matrix on a rectangular silicon substrate. In area, the diode array is just one tenth as large as a postage stamp. When light from an image strikes these photodiodes, electric current is generated and

SATICON® Conventional Pickup Tube

New MOS Chip

MOS Color Video Camera VK-C1000

picked up line by line at high speed through MOS transistor switches, thus producing video signals. Although this principle has been known for some time, no one had ever been able to put it into practical use. It required the highly precise, micrometric fabrication techniques and new advances in resolution power and color separation that only Hitachi engineers could provide.

Every one of the 180,000 diodes is equipped with a filter

One of the basic challenges was to fit the diodes with mosaic color filters capable of rendering color information as electrical signals. These filters had to be directly mounted piece by piece on each one of the 180,000 photodiodes, a mere 20 μm in size. Difficult as this task is to imagine, much less carry out, the patience and ingenuity of Hitachi researchers made it all possible. And the MOS image sensor is just the first revolutionary step in creating an entirely new generation of miniature TV cameras. Soon there will be home video cameras no bigger than cigarette packs and studio cameras small enough for one-hand operation. What's more, Hitachi engineers are even envisaging capsulated medical cameras so tiny that they can be used for internal diagnosis. The potential benefits of such micro-cameras are virtually limitless.

Technical excellence is embodied in all Hitachi products

The development of MOS image sensors is just one case demonstrating Hitachi's technological strength. You'll find other examples in all of Hitachi's products, from TVs and VTRs to a whole host of high-quality electronic appliances. Our comprehensive technical expertise is your guarantee of convenience, easy operation and high reliability in every product that bears the Hitachi brand.

HITACHI

A World Leader in Technology

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

FEBRUARY, 1932: "The Nobel prize for chemistry has been awarded to Dr. Friedrich Bergius of Germany, who shares it with another German chemist of fame, Dr. Karl Bosch. For 15 years—before, during and after the World War—Dr. Bergius labored on the process now known as 'Berginization.' It converts coal into gasoline or oils. The process is now controlled in America by the Standard Oil Company of New Jersey, which has it under actual though relatively small production at its Bayway, N.J., plant. Owing to the existing world over-production of petroleum this plant is not at present converting coal into gasoline but into lubricating oils. The Bergius process will be at its best when petroleum costs more than $2 a barrel, not the present 50 cents. It will be of great value for the future when petroleum becomes less plentiful."

"The properties of animal fibers have been brilliantly investigated by means of X-ray analysis by W. T. Astbury of the University of Leeds. Until four years ago Astbury was one of Sir William Bragg's assistants at the Royal Institution, where he learned and helped to advance the technique of the analysis of crystals by X rays. Since then he has applied X-ray-crystal analysis to the structure of wool, human hair, llama hair, hedgehog spine, porcupine quill, natural and artificial silk, fingernail and other external animal growths. These fibers all have substantially the same chemical constitution. There are comparatively slight differences in constitution that confer very important practical differences in properties, but to a large extent a human hair, a wool fiber and a ram's horn are made out of the same substance, a protein called keratin. The importance of fundamental knowledge of such materials to manufacturers is plain and great, but Astbury considers that the value of his researches may perhaps prove to be even greater in the pure science of biology, for they will help to reveal the special architecture employed by nature in the construction of living matter."

"Warning against the use of radio-active waters, frequently promoted for the cure of a great variety of diseases and ailments, is contained in a report by Dr. Harrison S. Martland, medical examiner of Essex County, New Jersey. From long study of the effects of radium poisoning on the luminous-dial painters Dr. Martland has concluded that it is dangerous to increase the normal radio-activity of the human body, on the strong presumption that increased amounts of radio-activity over a number of years may cause cancer. Dr. Martland states: 'To drink, over long periods of time, radio-active waters containing radon may allow a small amount of active, long-lived deposits to enter the body, part of which may finally be deposited in the bones and other organs as more or less insoluble salts.' Such deposits of insoluble salts of radium and other radio-active substances were responsible for the development of fatal cancer (sarcoma) in the watch-dial painters. Dr. Martland shows in his report that some of the radio-active waters for sale, when taken according to directions of the promoters, would require to be swallowed each day an amount of radio-active substance equal to that taken by some of the dial painters. These substances are in soluble form in water, but they are changed in the blood to more insoluble carbonates, phosphates and sulfates of radium and mesothorium, and eventually reach the bones. Once deposited there in insoluble form, there is no way known now to eliminate them from the body or to protect it from their fatal bombardment of alpha particles."

SCIENTIFIC AMERICAN

FEBRUARY, 1882: "Let us suppose flying-machines are invented, manufactured and offered for sale. England, for example, has hitherto in great measure been protected from the worst horrors of war by the 'silver streak of sea' interposed between it and the Continent. Large rivers, chains of mountains and fortifications have hitherto frequently enabled nations to make head against a more numerous invader. All such defenses, natural or artificial, would be at once sacrificed if the flying-machine becomes a reality. The nation that has the most numerous army, and that is constantly planning aggression, would be strengthened, while those nations that merely stand on the defensive would be weakened. It must also appear that when war is waged by sending flying-machines to hover over the cities of an enemy, and to let fall shells filled with nitro-glycerine, the advantage will be all on the side of any country that has a scanty and scattered population, with little wealth stored in any one locality. Is it not a solemn duty on the part of inventors to turn their energies in directions where at all events there is no fear of promoting aggressive war? Are there no diseases to extirpate? Is there not the synthesis of plant food and perhaps of human food from inorganic matter to be effected? Are there not new sources of energy to be discovered, that we must needs employ our time and talents in playing into the hands of crime?"

"A bold attempt to develop a new Atlantic seaport is being made in the neighborhood of the historic village of Newport News, on Hampton Roads, at the mouth of the James River in Virginia. A combination of prominent railway men and capitalists have purchased the entire water front for about eight miles and back from the shore about four miles. Streets for a new city have been laid out, and a large force of men are employed in constructing wharves, railway stations, warehouses and other requisites for the terminus of a vast railway system, embracing important central, southern and western connections. The advantages of Newport News for a seaport are very great. Hampton Roads, on which it is situated, is one of the finest harbors in the world, and the only one in the United States that can be safely entered without a pilot. There is no bar at the entrance to Chesapeake Bay, and a dozen navies could ride there abreast. Newport News is but 15 miles from the open sea, while New York is 20, Boston about 50, Philadelphia 100 and Baltimore 160. The Roads are never frozen over nor filled with broken ice, and the water close to the shore will float the largest vessel ever built."

"Some unique experiments have been lately made in France on the strength of the masseter muscle of the crocodile (a muscle passing from the cheek bone to the lower jaw). M. Paul Bert received 10 large crocodiles (*Crocodilus galeatus*) from Saigon. In order to measure the strength of the masseter muscle of the crocodile's jaw the animal was firmly fastened to a table attached to the floor; the lower jaw was fixed immovably by cords to the table; the upper jaw was then attached to a cord, fastened by a screw ring to a beam in the roof. There was a dynamometer placed on this cord, so that when the animal was irritated or given an electric shock, the upper jaw pulled on the cord, and registered the force of its movement on the dynamometer. With a crocodile weighing 120 lb. the force obtained was about 308 lb. avoirdupois. This does not equal the actual strength, for as the dynamometer is necessarily placed at the end of the snout, it is really at the end of a long lever, and the strength must be measured by finding the distance between the jaw muscle and the end of the jaw to show the real force of the jaw muscle, which equals 1,540 lb. The power of the crocodile's jaw was compared with that of an ordinary dog weighing about 44 lb. A force of 72 lb. was obtained. In comparing the jaw force of the two animals it is found that a crocodile is one-third stronger, weight for weight, than a dog."


Nationwide
taste tests prove it!
Windsor Canadian
beats V.O.!
IMPORTED
WINDSOR
Supreme
CANADIAN
Blended Canadian Whisky
V.O.
Five hundred serious Canadian Whisky drinkers coast-to-coast just compared Windsor Canadian to the higher-priced Seagram's V.O.
Windsor was preferred.
So try a sip of Windsor and a sip of V.O. and prove to yourself what the taste tests just proved.
With Windsor, you can't beat the taste.
And you sure can't beat the price.
WINDSOR
ONE CANADIAN STANDS ALONE

THE AUTHORS

ARTHUR C. UPTON ("The Biological Effects of Low-Level Ionizing Radiation") is chairman of the Institute of Environmental Medicine of New York University School of Medicine. He received his undergraduate education, medical degree and graduate training in pathology at the University of Michigan. From 1951 to 1969 he was on the staff of the Oak Ridge National Laboratory. In 1969 he became professor of pathology at the State University of New York at Stony Brook, remaining there until 1977; in that period he was also a visiting pathologist at the Brookhaven National Laboratory. From 1977 until 1979 he was the director of the National Cancer Institute. The biological effects of ionizing radiation have been one of Upton's central research interests since he was at Oak Ridge; his professional interests also include the effects of other environmental agents, the causes of cancer and the biology of aging.

RICHARD L. HAY and MARY D. LEAKEY ("The Fossil Footprints of Laetoli") are respectively a geologist and an anthropologist who have collaborated in the study of the predecessors of man in Tanzania since 1962. Hay's contribution has consisted in working out the stratigraphy, sedimentology and paleoenvironments of the volcanic deposits where the hominid fossils are found. He attended Ursinus College, Northwestern University (B.S. and M.S.) and Princeton University (Ph.D.). He has taught at Louisiana State University and currently works at the University of California at Berkeley. He has studied deposits of volcanic ash in the western U.S., Hawaii, Japan, Kenya and St. Vincent in the West Indies. Among Hay's nonprofessional interests are fishing for striped bass and the finding and polishing of jade. Mary Leakey was privately educated in England. She first visited Olduvai Gorge and Laetoli with her late husband, Louis B. Leakey, in 1935. Most of her career has been devoted to East African paleoanthropology; together with Louis Leakey she has made some of the most important discoveries of early hominid remains in Africa. She has received honorary degrees from, among many other institutions, Yale University, the University of Oxford and the University of Chicago. One of her main interests outside paleoanthropology is game watching.

KAI SIMONS, HENRIK GAROFF and ARI HELENIUS ("How an Animal Virus Gets into and out of Its Host Cell") are cell biologists who are particularly interested in the movement of substances across the cell membrane. Until recently they worked together at the European Molecular Biology Laboratory (EMBL) in Heidelberg. Simons, a senior scientist at the EMBL, obtained his M.D. at the University of Helsinki in 1964. He did postdoctoral work at Rockefeller University; there he was introduced to the Semliki Forest virus, the animal virus discussed in the article in this issue. He returned to the faculty of the University of Helsinki in 1967 and remained there until 1975, when he joined the staff of the EMBL. Garoff and Helenius became members of Simons' research group as graduate students at the University of Helsinki. Both did doctoral research on the Semliki Forest virus; Helenius got his Ph.D. in 1973 and Garoff his M.D. in 1974. In 1975 they accompanied Simons to the EMBL, where their work on the virus continued. Garoff is now a research-group leader at the EMBL; Helenius moved last year to Yale University as associate professor of cell biology.

TOBIAS OWEN ("Titan") is professor of astronomy at the State University of New York at Stony Brook. His bachelor's and master's degrees, both in physics, are from the University of Chicago; he received his Ph.D. (in astronomy) from the University of Arizona in 1965. He has been a member of the imaging science team for *Voyager 1* and *Voyager 2*. "My participation in the Voyager missions," he writes, "provided the totally unexpected opportunity to get close views and detailed information about the objects I had first known as points of light in the night skies of my childhood."

WILLIAM C. GARDINER, JR. ("The Chemistry of Flames"), is professor of chemistry at the University of Texas at Austin. He got his education at Princeton University (A.B.) and Harvard University (Ph.D.). From 1955 to 1957 he was a research associate at the Max Planck Institute for Physical Chemistry in Göttingen. In 1960 he went to the University of Texas, returning to the Max Planck Institute in 1975 and 1976 as a Guggenheim Fellow. His research interests have included shock and detonation waves, combustion, electron-spin-resonance spectroscopy and laser photochemistry.

PATRICK S. OSMER ("Quasars as Probes of the Distant and Early Universe") is director of the Cerro Tololo Inter-American Observatory in Chile. He majored in astronomy at the Case Institute of Technology, obtaining his bachelor's degree in 1965. His Ph.D., also in astronomy, was awarded by the California Institute of Technology in 1970. He began at Cerro Tololo in 1969 as a research associate; he was appointed director last year. The observatory is the site of a four-meter telescope; installed in the 1970's, it is the largest in the Southern Hemisphere. His research interests include the atmospheres of very luminous stars, the Magellanic clouds and stellar sources of X rays. Quasars have been Osmer's major interest since 1974.

JOSEPH S. LEVINE and EDWARD F. MACNICHOL, JR. ("Color Vision in Fishes"), are sensory physiologists with a special interest in aquatic animals. Levine is assistant professor of biology at Boston College and a research associate at the Marine Biological Laboratory in Woods Hole, Mass. He was graduated from Tufts University in 1973 and went on to obtain a master's degree at Boston University in 1976, writing his thesis on the aquaculture of lobsters. He received his Ph.D. in biology from Harvard University in 1980. MacNichol is director of the laboratory of sensory physiology at Woods Hole and professor of physiology at Boston University School of Medicine. After getting an undergraduate degree at Princeton University in 1941 he went to the Radiation Laboratory of the Massachusetts Institute of Technology. He received his Ph.D. from Johns Hopkins University in 1952. He has long been interested in the measurement of electrical activity in individual nerve cells. His initial studies included some of the first intracellular recordings from the photoreceptor cells of the lateral eye of the horseshoe crab. MacNichol left Johns Hopkins in 1968 to become director of the National Institute of Neurological Diseases and Stroke, a position he held until 1973. For part of that time he was also the acting director of the National Eye Institute.

A. D. MOORE ("Henry A. Rowland") is professor emeritus of electrical engineering at the University of Michigan. He was graduated from the Carnegie Institute of Technology in 1915. After a year in the graduate training program at the Westinghouse Electric Company he joined the Michigan faculty; he became professor emeritus in 1964. During World War II he worked for the Naval Ordnance Laboratory, recruiting scientists and engineers for the institution. From 1940 to 1957 he was a member of the city council of Ann Arbor. He has concentrated on graphic means of analyzing magnetic, electric and heat-conduction fields. He has invented several instruments that employ fluids to represent electric and magnetic fields. He has also designed and built electrostatic generators. Since his retirement Moore has utilized the fluid techniques and generators in a lecture program presented at many universities.

METAMAGICAL THEMAS

About two kinds of inquiry: "National Enquirer" and "The Skeptical Inquirer"

by Douglas R. Hofstadter

"Baffled Investigators and Educators Disclose . . . BOY CAN SEE WITH HIS EARS"

"A Cross between Human Beings and Plants . . . SCIENTISTS ON VERGE OF CREATING PLANT PEOPLE . . . Bizarre Creatures Could Do Anything You Want"

"Alien from Space Shares Woman's Mind and Body, Hypnosis Reveals"

—Headlines from *National Enquirer*

Did the child you once were ever wonder why the declarative sentences in comic books always ended with exclamation points? Were all those statements really that startling? Were the characters saying them really that thrilled? Of course not. Those exclamation points were a psychological gimmick put there purely for the sake of appearance, to give the story more pizzazz.

National Enquirer, one of this country's yellowest and purplest journalistic institutions, uses a similar gimmick. Whenever it prints a headline trumpeting the discovery of some bizarre, hitherto unheard-of phenomenon, instead of ending it with an exclamation point it ends it (or begins it) with a reference to "baffled investigators," "bewildered scientists" or similarly stumped savants. It is an ornament put there to make the story seem to have more credibility.

Or is it? What do the editors really want? That the story appear credible or that it appear incredible? It seems they want it both ways: they want the story to sound as outlandish as possible and they want it to have the appearance of authenticity. Their ideal headline should thus embody a contradiction: impossibility coupled with certainty. In short, confirmed nonsense.

What can one make of headlines such as the ones printed above? Or of the fact that this publication is sold by the millions every week in grocery stores, and that people gobble up its stories as voraciously as they do potato chips? Or of the fact that when they are through with it, they can turn to plenty of other junk food for thought: *National Examiner, Star, Globe, Weekly World News.* What do you think?

Your first reaction is probably to chuckle and dismiss such stories as being silly. But how do you know they are silly? Do you also think *that* is a silly question? What do you think about articles that are printed in this magazine? Do you trust them? What is the difference? Is it simply a difference in publishing style? Is the tabloid format with its gaudy pictures and sensationalistic headlines enough to make you distrust *National Enquirer*? But wait a minute. Is that not begging the question? What kind of argument is it when you use the guilty verdict as part of the case for the prosecution? What you need is a way of telling objectively what you mean by "gaudy" or "sensationalistic." That could prove to be difficult.

And what about the obverse of the coin? Is it the rather dignified, traditional format of *Scientific American*—its lack of photographs of celebrities, for example—that convinces you it is to be trusted? If it is, that is a curious way of making decisions about what truth is. It would seem that your concept of truth is closely tied in with your way of evaluating the "style" of a channel of communication, surely quite an intangible notion.

Having said that, I must admit I too rely constantly on quick assessments of style in my attempt to sift the true from the false, the believable from the unbelievable. I could not tell you what criteria I rely on without first thinking about it for a long time and writing many pages. Even then, if I were to publish the definitive guide (*How to Tell the True from the False by Its Style of Publication*), it would have to be published to do any good, and its title, not to mention the style it was published in, would probably attract a few readers but would undoubtedly repel many more.

Well, truth being this elusive, no wonder people are besieged with competing voices in print. When I was younger, I believed once something had been discovered, verified and published it was part of Knowledge: definitive, accepted and irrevocable.

To my surprise, however, I found that the truth has to fight constantly for its life. That an idea has been discovered and printed in a "reputable journal" does not ensure that it will become well known and accepted. In fact, usually it will have to be rephrased and reprinted many different times, often by many different people, before it has any chance of taking hold. This is upsetting to an idealist such as me, someone more disposed to believe in the notion of a monolithic and absolute truth than in the notion of a pluralistic and relative truth. The idea that the truth has to fight for its life is a sad discovery. The idea that the truth will *not* out, unless it is given a lot of help, is upsetting.

A question arises in every society: Is it better to let all the different voices battle it out or to have just a few "official" publications dictate what is the case and what is not? Our society has opted for a plurality of voices, for a "marketplace of ideas," for a complete free-for-all of conflicting theories. But if it is this chaotic, who will ensure that there is law and order? Who will guard truth? The answer is: CSICOP will!

CSICOP? What is CSICOP? Some kind of cop who guards the truth? That's pretty close. CSICOP stands for Committee for the Scientific Investigation of Claims of the Paranormal. It is a rather esoteric title for an organization whose purpose is not so esoteric: to apply common sense to claims of the outlandish, the implausible and the unlikely.

Who are the people who form CSICOP and what do they do together? The organization was the brainchild of Paul Kurtz, professor of philosophy at the State University of New York at Buffalo, who brought it into being because he thought there was a need to counter the rising tide of irrational beliefs and to provide the public with a more balanced treatment of claims of the paranormal by presenting the dissenting scientific viewpoint. Among the early fellows of CSICOP were some of America's most distinguished philosophers (for example Ernest Nagel and W. V. Quine) and other colorful combatants of the occult, such as psychologist Ray Hyman, magician James Randi and someone readers of this column may have heard of: Martin Gardner. In the first few meetings it was decided that the committee's principal function would be to publish a magazine dedicated to the subtle art of debunking. Perhaps "debunking" is not the term they would have chosen, but it fits. The magazine they began to publish in the fall of 1976 was *The Zetetic,* from the Greek for "inquiring skeptic."

As happens with many fledgling movements, a philosophical squabble developed between two factions, one more "relativist" and unjudgmental, the other

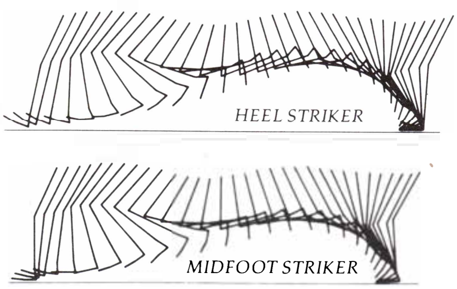
"MY GOD... ANOTHER HEEL STRIKER."

Every day you take one of the body's most intricate structures—some 26 small bones, muscle and connective tissue—and slam it into the ground. With a force three times your body weight.

If that weren't enough, you keep on doing it. About 7,400 times in a typical five mile training run.

Luckily, feet have an uncanny ability to roll with the punches. Each a little bit differently. Some take it on the heel. Others right in the midfoot. Still others in the forefoot.

Obviously, each different footstrike requires a different kind of shoe. But it's a little difficult to design the proper model if you don't fully understand what's happening. And it happens pretty quickly—like in less than a quarter of a second.

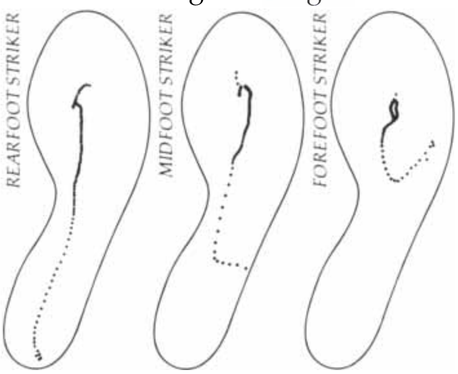


Lower limb kinematics taken from computer-digitized high speed film. Midfoot strikers generally land with greater knee flexion and with the body more directly over the foot.

With high-speed cameras we can break that split-second action into 125 individual pictures. Then digitize the film and feed it into our computers.

We also use force platforms to record what the naked eye never sees—the actual rise and fall of impact forces.

So we know how much cushioning is needed. Where it's needed. And when it ought to be sacrificed for lighter weight.



Center of pressure patterns under the foot while running across force platform at 6:00 mile pace.

Of the 3,000 runners we've surveyed so far, 75% identified themselves as heel strikers; 23% said they were mid- or forefoot strikers; and 2% didn't know for certain.

Most self-appraisals, we've found, are accurate. And since footstrike seldom changes with increased speed (unless it's an all-out sprint), the best advice may be to do what comes naturally.

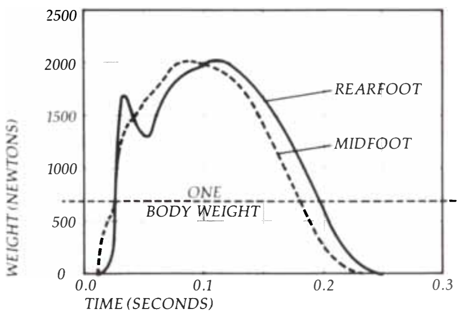
But do it in the proper shoe.

For better midfoot or forefoot cushioning, try the Boston, Atlanta or our new Terra T/C.

For heel strikers, there's the Columbia, Aurora, Tailwind, Tempest, Internationalist, LDV, Daybreak, Yankee and the Intrepid.

But there's more to it than good cushioning. Get yourself to a knowledgeable dealer. The kind who will also ask about your foot type, body weight, and mileage before he starts hauling out the boxes.

If you have trouble finding anyone who takes your shoe selection seriously, drop us a line. We'll send you information that will make it simple. Send a self-addressed stamped envelope to NIKE, Attn. Ellen Borsting, Dept. SA, 3900 S.W. Murray Blvd., Beaverton, OR 97005.



Vertical ground reaction forces for two subjects of same weight, running at same speed. Note sharp impact peak in heel striker which is absent in this midfoot striker and most others.

You may be alone out there. But that's no reason to feel stranded.

NIKE

Beaverton, Oregon

more firmly opposed to nonsense, more willing to go on the offensive and to attack supernatural claims. Strange to say, the open-minded faction was not so open-minded as to accept the opposing point of view, and the rift opened wider. Eventually there was a schism. The relativist faction (one member) started publishing his own journal, *The Zetetic Scholar,* in which science and pseudoscience coexist happily. The larger faction retained the name CSICOP and changed the title of its journal to *The Skeptical Inquirer.*

The purpose of *The Skeptical Inquirer* is simply to combat nonsense. It does so by recourse to common sense, which means it is accessible to anyone who can read English. It does not require any special knowledge or training to read its pages, where nonsensical claims are routinely smashed to smithereens. (Sometimes the claims are as blatantly silly as the headlines at the beginning of this article, sometimes they are much subtler.) All that is required to read this maverick journal is curiosity about the nature of truth: curiosity about how truth defends itself (through its agent CSICOP) against attacks from all quarters by unimaginably imaginative theorizers, speculators, eccentrics, crackpots and out-and-out fakers.

The journal has grown from its original small number of subscribers to roughly 7,500—a David compared with the Goliaths mentioned above, with their circulations in the millions. Its pages are filled with lively and humorous writing—the combat of ideas in its most enjoyable form. The journal is by no means a monolithic voice, an advocate of a single dogma. Rather, it is itself a marketplace of ideas. Even people who wield the tool of common sense with skill may do so with different styles, and sometimes they will disagree.

There is something of a paradox involved in the editorial decisions in such a magazine. After all, what is under debate here is in essence the nature of correct arguments. What should be accepted and what should not? To caricature the situation, imagine the editorial dilemmas that journals with titles such as *Free Press Bulletin, The Open Mind* or *Editorial Policy Newsletter* would encounter. What letters to the editor should be printed? What articles? What policy can be invoked to screen material submitted?

These are not easy questions to answer. They involve a paradox, a tangle in which the ideas being evaluated are also the ideas doing the evaluating. The only recourse is to common sense, that rock-bottom basis of all rationality. Unfortunately we have no foolproof algorithm to uniquely characterize that deepest layer of rationality, nor are we likely to come up with one soon. For now the core of rationality must depend on inscrutables: the simple, the elegant, the intuitive. This paradox has existed throughout intellectual history, but in our information-rich times it seems particularly troublesome.

In spite of such epistemological puzzles, which are connected to its very reason for existence, *The Skeptical Inquirer* is flourishing and provides a refreshing antidote to the jargon-laden journals of science, which often seem curiously irrelevant to the concerns of everyday life. In that one way the *Inquirer* resembles the scandalous tabloids.

The list of topics covered in the 17 issues that have appeared so far is remarkably diverse. Some topics come up only once, others come up regularly and are discussed from various angles and at various depths. Some of the more commonly discussed topics are ESP, telekinesis (using mental power to influence events at a distance), astrology, biorhythms, Bigfoot, the Loch Ness monster, UFO's, creationism, telepathy, remote viewing, clairvoyant detectives who allegedly solve crimes, the Bermuda and other triangles, "thoughtography" (using mental power to create images on film), the supposed extraterrestrial origin of life on the earth, Carlos Castaneda's sorcerer "Don Juan," pyramid power, psychic surgery and faith healing, Scientology, predictions by famous "psychics," spooks and spirits and haunted houses, levitation, palmistry and mind reading, unorthodox anthropological theories, plant perception, perpetual-motion machines, water witching and other kinds of dowsing, bizarre cattle mutilations. And these are by no means the only topics; they are just the regulars.

There are quite a few frequent contributors, such as Randi, who is truly prolific. Among others are aeronautics writer Philip J. Klass, UFO specialist James E. Oberg, writer Isaac Asimov, CSICOP's founder (and current chairman) Kurtz, psychologist James E. Alcock, educator Elmer Kral, anthropologist Laurie Godfrey, science writer Robert Sheaffer, sociologist William Sims Bainbridge and many others. And the magazine's editor, Kendrick Frazier, a free-lance science writer by trade, periodically issues eloquent and mordant commentaries.

There is no better way to impart the flavor of the magazine than to quote a few selections from articles. One of my favorite articles appeared in the second issue (Spring/Summer, 1977). It is by psychologist Ray Hyman (who incidentally, like many other authors in *The Skeptical Inquirer,* is also a talented magician), and is titled "'Cold Reading': How to Convince Strangers that You Know All about Them."

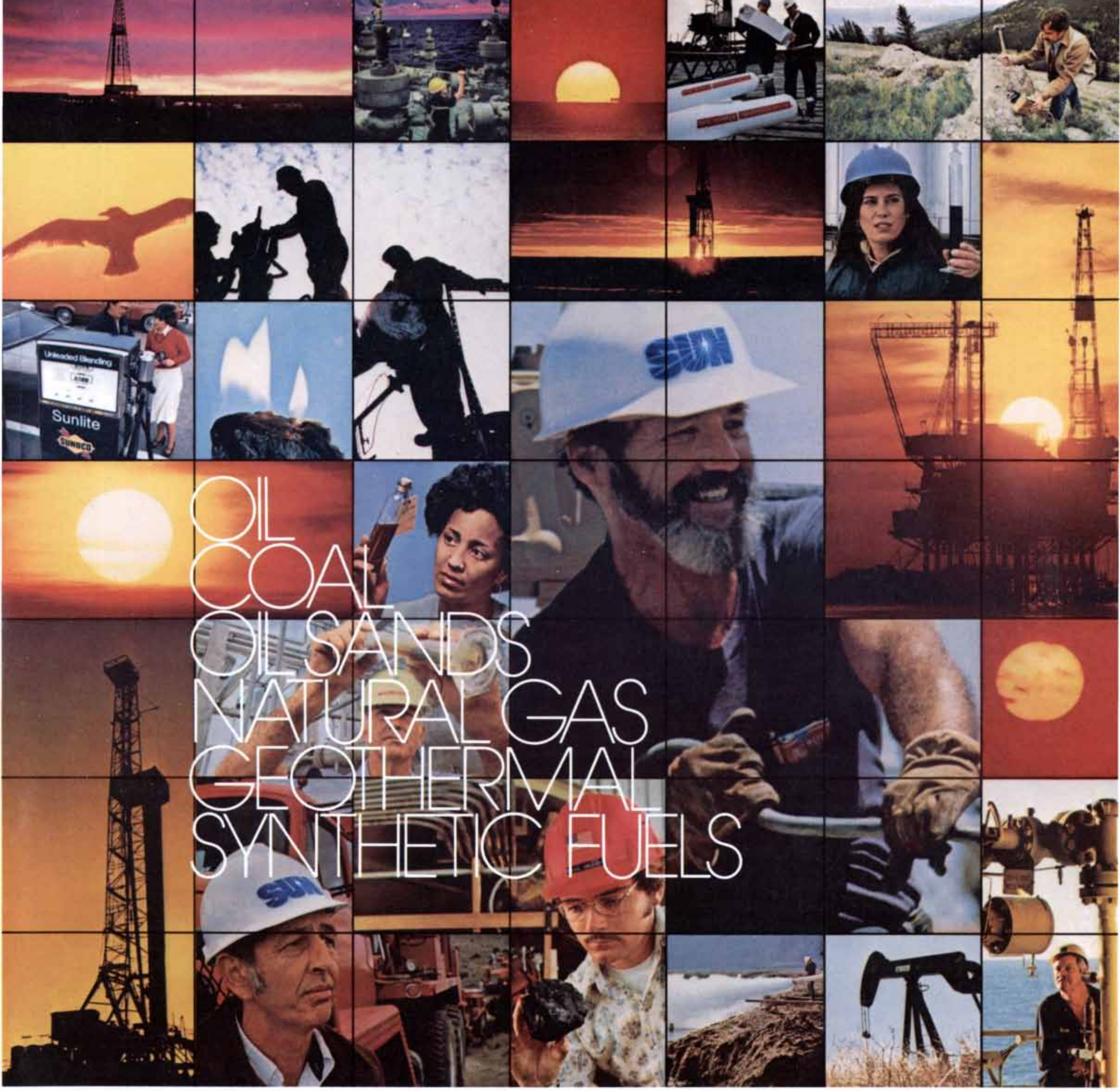
It begins with a discussion of a course Hyman taught about the various ways people are manipulated. Hyman states: "I invited various manipulators to demonstrate their techniques—pitchmen, encyclopedia salesmen, hypnotists, advertising experts, evangelists, confidence men and a variety of individuals who dealt with personal problems. The techniques which we discussed, especially those concerned with helping people with their personal problems, seem to involve the client's tendency to find more meaning in any situation than is actually there. Students readily accepted this explanation when it was pointed out to them. But I did not feel that they fully realized just how pervasive and powerful this human tendency to make sense out of nonsense really is."

Then Hyman describes people's willingness to believe what others tell them about themselves. His "golden rule" is: "To be popular with your fellow man, tell him what he wants to hear. He wants to hear about himself. So tell him about himself. But not what you know to be true about him. Oh, no! Never tell him the truth. Rather, tell him what he would like to be true about himself!" As an example, Hyman cites the following passage (which, by a remarkable coincidence, was written about none other than *you,* dear reader!):

"Some of your aspirations tend to be pretty unrealistic. At times you are extroverted, affable, sociable, while at other times you are introverted, wary and reserved. You have found it unwise to be too frank in revealing yourself to others. You pride yourself on being an independent thinker and do not accept others' opinions without satisfactory proof. You prefer a certain amount of change and variety, and become dissatisfied when hemmed in by restrictions and limitations. At times you have serious doubts as to whether you have made the right decision or done the right thing. Disciplined and controlled on the outside, you tend to be worrisome and insecure on the inside.

"Your sexual adjustment has presented some problems for you. While you have some personality weaknesses, you are generally able to compensate for them. You have a great deal of unused capacity which you have not turned to your advantage. You have a tendency to be critical of yourself. You have a strong need for other people to like you and for them to admire you."

Pretty good fit, eh? Hyman comments: "The statements in this stock spiel were first used in 1948 by Bertram Forer in a classroom demonstration of personal validation. He obtained most of them from a newsstand astrology book. Forer's students, who thought the sketch was uniquely intended for them as a result of a personality test, gave the sketch an average rating of 4.26 on a scale of 0 (poor) to 5 (perfect). As many as 16 out of his 39 students (41 percent) rated it as a perfect fit to their personality. Only five gave it a rating below 4 (the



A lot of energy under one name.

SUN

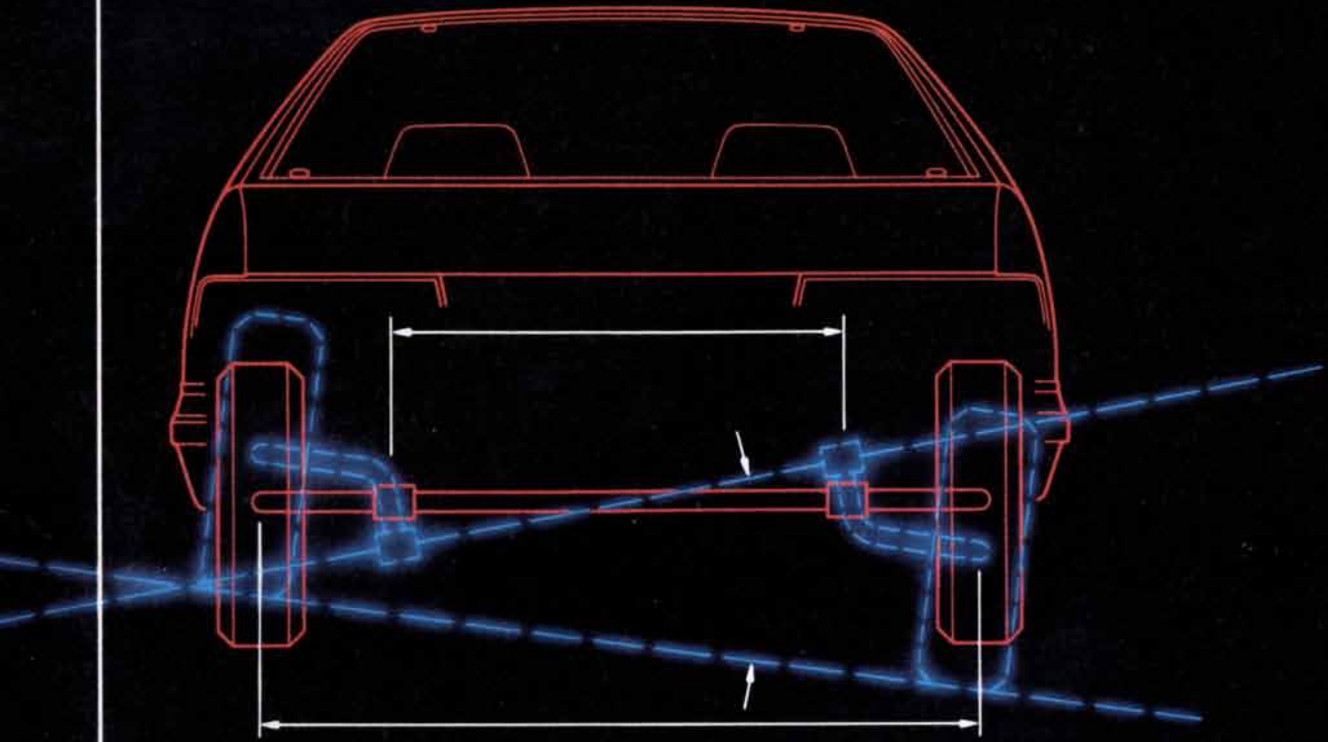
Sun. A lot more than an oil company.

For almost 100 years, Sun has been supplying valuable and needed products to America's industry and consumers.

Today, Sun Company is a broad-based energy company. Sun is working to help secure America's energy future through its commitment to exploration, production and delivery of energy from a variety of sources.

Oil. Coal. Oil sands. Natural gas. Geothermal. And synthetic fuels. A lot of energy under one name. Sun.

Helping to make a new day.



To improve a car's handling, test drive it in a computer.

Maneuverability and responsiveness. Everyone wants a car with those handling characteristics. But what kind of maneuverability, and how much responsiveness? That's the challenge: finding a way to translate subjective tastes into objective engineering guidelines.

At GM, we use sophisticated computer modeling to design and test the properties of vehicle structures, suspensions, steering systems, and tires. These make up the unique "fingerprints" that distinguish each model car from the others.

We pioneered in the development of computer technology needed to do this complex designing back in the early '50s. And we're still a leader in the application of computer-aided vehicle dynamics.

Once a prototype is built, the computer is also used to analyze data on maneuverability collected from gyros and accelerometers inside the car. Measured performance is compared to planned goals at every development stage until the car is ready for mass production.

That's our idea of how to use technology to build cars and trucks. Attention to details where you don't see them, as well as where you do. Appearance and comfort may sometimes sell a car, but today's customers demand real value.

Our goal, as the world's largest automotive manufacturer, is to maintain our lead by using new technology to build cars and trucks that perform better and last longer, with lower maintenance costs, than those built by any competitor —American or foreign.

General Motors

The future of transportation is here.

GM

worst being a rating of 2, meaning 'average'). Almost 30 years later students give the same sketch an almost identical rating as a unique description of themselves."

Hyman gives a 13-point recipe for becoming a cold reader. Among his tips are these: "Use the technique of 'fishing' [getting the subject to tell you about himself or herself, then rephrasing it and feeding it back]; always give the impression that you know more than you are saying; don't be afraid to flatter your subject every chance you get." This deliciously cynical recipe for becoming a character reader is presented in considerable detail, presumably not to convert readers of the article into charlatans and tricksters but to show them how such manipulations are achieved.

Hyman asks: "Why does it work so well? It does not help to say that people are gullible or suggestible. Nor can we dismiss it by implying that some individuals are just not sufficiently discriminating or lack sufficient intelligence to see through it. Indeed, one can argue that it requires a certain degree of intelligence on the part of a client for the reading to work well.... We have to bring our knowledge and expectations to bear in order to comprehend anything in our world. In most ordinary situations this use of context and memory enables us to correctly interpret statements and supply the necessary inferences to do this. But this powerful mechanism can go astray in situations where there is no actual message being conveyed. Instead of picking up random noise we still manage to find meaning in the situation. So the same system that enables us to creatively find meanings and to make new discoveries also makes us extremely vulnerable to exploitation by all sorts of manipulators. In the case of the cold reading the manipulator may be conscious of his deception; but often he too is a victim of personal validation."

(Hyman knows whereof he speaks. Many years ago he was convinced for a time that he himself had genuine powers to read palms, until one day he tried telling people the exact opposite of what their palms told him and found they still swallowed his line as much as ever. Then he began to suspect that the plasticity of the human mind—particularly his own—was doing some strange things.)

At the beginning of each issue of *The Skeptical Inquirer* is a feature called "News and Comment." It covers such things as the latest reports on current sensational claims, recently broadcast television shows for and against the paranormal, lawsuits of one kind or another and so on. One of the most amusing items was the coverage in the Fall 1980 issue of the "Uri Awards," given out by Randi (on April 1, of course) to various deserving souls who had done the most to promote gullibility and irrational beliefs. Each award consists of "a tastefully bent stainless-steel spoon with a very transparent, very flimsy base." Award winners were notified, Randi explained, by telepathy, and they were "free to announce their winning in advance, by precognition, if they so desired." Awards were made in four categories: Academic ("to the scientist who says the dumbest thing about parapsychology"), Funding ("to the funding organization that awards the most money for the dumbest things in parapsychology"), Performance ("to the psychic who, with the least talent, takes in the most people"), and Media ("to the news organization that supports the most outrageous claims of the paranormalists").

The nature of coincidences is a recurrent theme in discussions of the paranormal. I vividly remember a passage in a lovely book by Warren Weaver titled *Lady Luck: The Theory of Probability,* in which he points out that in many situations the most likely outcome may well be a very unlikely event (as it is when you deal hands in bridge, where whatever hand you get is bound to be extraordinarily rare no matter what it is). A similar point is made in the following excerpt from a recent book by David Marks and Richard Kammann titled *The Psychology of the Psychic* (from which various excerpts were reprinted in one issue of *The Skeptical Inquirer*):

"First, we notice and remember matches, especially *oddmatches,* whenever they occur. (Because a psychic anecdote first requires a match, and, second, an oddity between the match and our beliefs, we call these stories *oddmatches.* This is equivalent to the common expression, an 'unexplained coincidence.') Second, we do *not* notice nonmatches. Third, our failure to notice nonevents creates the *short-run illusion* that makes the oddmatch seem improbable. Fourth, we are poor at estimating combinations of events. Fifth, we overlook the *principle of equivalent oddmatches,* that one coincidence is as good as another as far as psychic theory is concerned."

An excellent example of people not noticing nonevents is provided by the failed predictions of famed psychics (such as Jeane Dixon). Most people never go back to see how the events bore out the predictions. *The Skeptical Inquirer,* however, has a tradition of going back and checking. As each year concludes it prints a number of predictions made by various psychics for that year and then evaluates their track record. In the Fall 1980 issue the editors took the predictions of 100 "top psychics," tabulated them, listed the top 12 in order of frequency and left it to the reader to assess the accuracy of psychic visions of the future. The No. 1 prediction for 1979 (made by 86 psychics) was "Longer lives will be had for almost everyone as aging is brought under control." No. 2 (85 psychics) was "There will be a major breakthrough in cancer, which will almost totally wipe out the disease." No. 3 (also 85) was "There will be an astonishing spiritual rebirth and a return to the old values." So it went. No. 6 (81 psychics) was "Contact will be made with aliens from space, who will give us incredible knowledge."

There is something pathetic, even desperate, about these predictions. One can see only too clearly the similarity of the publications that feature these predictions to inane television shows such as "Fantasy Island" and "Star Trek." The common denominator is escape from reality. The point is well made in an article by William Sims Bainbridge in the Fall 1979 issue of *The Skeptical Inquirer.* Perhaps we all have a desire to dilute reality with fantasy, to make reality seem simpler and also more aligned with what we wish it were. Yet at the same time perhaps all of us have the potential capacity and even the desire to sift sense from nonsense, if only we are introduced to the distinction in a sufficiently compelling manner.

But how can this be done? In the "News and Comment" section of the Spring 1980 issue of *The Skeptical Inquirer* there was an item about a lively traveling antipseudoscience lecture act by "Captain Ray of Light," actually Douglas F. Stalker, associate professor of philosophy at the University of Delaware. The article quotes Stalker on his "comical debunking" show (directed at astrology, biorhythms, numerology, UFO's, pyramid power, psychic claims and the like) as follows:

"For years I lectured against them in a serious way, with direct charges at their silly theories. These direct attacks didn't change many minds, and so I decided to take an indirect appoach. If you can't beat them, join them. And so I did, in a manner of speaking. I constructed some plainly preposterous pseudosciences of my own and showed'that they were just like astrology and the others. I also explained how you could construct more of these silly theories. By working from the inside out, more students came to see how pseudo these pseudosciences are.... And that is the audience I try to reach: the upcoming group of citizens. My show reaches them in the right way, too. It leaves a lasting impression; it wins friends and changes minds." (I am pleased to report that Stalker welcomes new bookings. He can be reached at the Department of Philosophy, University of Delaware, Newark, Del. 19711.)

One of the points Stalker makes is that no matter how eloquent a lecture may be it simply does not have the power to convince that experience does. The point has been well brought out in a classic study made by Barry Singer and Victor A. Benassi of the psychology department of California State University at Long Beach. These two investigators set out to determine the effect on first-year psychology students of seemingly paranormal effects created in the classroom by an exotically dressed magician. Their findings were reported in the Winter 1980/81 issue of *The Skeptical Inquirer* in a piece titled "Fooling Some of the People All of the Time."

In two of the classes the performer (Craig Reynolds) was introduced as a graduate student "interested in the psychology of paranormal or psychic abilities [who has] been working on developing a presentation of his psychic abilities." The instructor also explicitly stated, "I'm not convinced personally of Craig's or anyone else's psychic abilities." In two other classes Craig was introduced as a graduate student "interested in the psychology of magic and stage trickery [who has] been working on developing a presentation of his magic act." The authors emphasize that all the stunts Craig performed are "easy amateur tricks that have been practiced for centuries and are even explained in children's books of magic."

After the act the students were asked to report their reactions. Singer and Benassi received two jolts from the reports: "First... in both the 'magic' and the 'psychic' classes, about two-thirds of the students clearly believed Craig was psychic. Only a few students seemed to believe the instructor's description of Craig as a magician in the two classes where he was introduced as such. Secondly, psychic belief was not only prevalent; it was strong and loaded with emotion. A number of students covered their papers with exorcism terms and exhortations against the Devil. In the psychic condition 18 percent of the students explicitly expressed fright and emotional disturbance. Most expressed awe and amazement.

"We were present at two of Craig's performances and witnessed some extreme behavior. By the time Craig was halfway through the 'bending' chant [part of a stunt where he bent a stainless-steel rod], the class was in a terribly excited state. Students sat rigidly in their chairs, eyes glazed and mouths open, chanting together. When the rod bent, they gasped and murmured. After class was dismissed, they typically sat still in their chairs, staring vacantly or shaking their heads, or rushed excitedly up to Craig, asking him how they could develop such powers. We felt we were observing an extraordinarily powerful behavioral effect. If Craig had asked the students at the end of his act to tear off their clothes, throw him money and start a new cult, we believe some would have responded enthusiastically. Obviously, something was going on here that we didn't understand."

After this dramatic presentation the classes were told they had only been seeing tricks. In fact, two more classes were given the same presentation, with the added warning: "In his act Craig will pretend to read minds and demonstrate psychic abilities, but Craig does not really have psychic abilities, and what you'll be seeing are really only tricks." In spite of this forewarning more than half of the students in these classes believed he was psychic. "This says either something about the status of university instructors with their students or something about the strange pathways people take to occult belief," Singer and Benassi observe philosophically.

Now comes something astonishing. "The next question asked was whether magicians could do exactly what Craig did. Virtually all the students agreed that magicians could. They were then asked if they would like to revise their estimate of Craig's psychic abilities in the light of this negative information that they themselves had furnished. Only a few did, reducing the percentage of students believing that Craig had psychic powers to 55 percent.

"Next the students were asked to estimate how many people who performed stunts such as Craig's and claimed to be psychic were actually fakes using magician's tricks. The consensus was that at least three out of four 'psychics' were in fact frauds. After supplying this negative information, they were again asked if they wished to revise their estimate of Craig's psychic abilities. Again, only a few did, reducing the percentage believing that Craig had psychic powers to 52 percent."

Singer and Benassi muse: "What does all this add up to? The results from our paper-and-pencil test suggest that people can stubbornly maintain a belief about someone's psychic powers *when they know better*. It is a logical fallacy to admit that tricksters can perform exactly the same stunts as real psychics and to estimate that most so-called psychics are frauds and at the same time maintain with a fair degree of confidence that any given example (Craig) is psychic. Are we humans really that foolish? Yes."

A few years ago Scot Morris (now a senior editor of *Omni* in charge of its "Games" department) carried out a similar experiment in a first-year psychology class at Southern Illinois University, which he wrote up in the Spring 1980 issue of *The Skeptical Inquirer*. First Morris assessed his students' beliefs in ESP by having them fill out a questionnaire. Then a colleague performed an "ESP demonstration," which Morris calls "frighteningly impressive."

After this performance Morris tried to "deprogram" his students. He had two weapons at his disposal. One is what he calls "dehoaxing." This process, lasting for only three minutes, consisted in revealing how two of the three tricks worked, together with a confession that the remaining one of the baffling stunts was also a trick, "but," said Morris, "I'm not going to say how it was done, because I want you to experience the feeling that, even though you can't explain something, that doesn't make it supernatural." The other weapon was a 50-minute anti-ESP lecture, in which secrets of professional mind readers were revealed, commonsense estimates of probabilities of "oddmatches" were discussed, "scientific" studies of ESP were shown to be questionable for various statistical and logical reasons and some other everyday reasons were adduced to cast ESP's reality into strong doubt.

After the performance only half of the classes were dehoaxed but all of them heard the anti-ESP lecture. The students were then polled about the strength of their belief in various kinds of paranormal phenomena. It turned out that dehoaxed classes had a far lower belief in ESP than classes that had simply heard the anti-ESP lecture. The dehoaxed classes' average level of ESP belief dropped from nearly 6 (moderate belief) to about 2 (strong disbelief), whereas the nondehoaxed classes' average level dropped from 6 to about 4 (slight disbelief). As Morris summarizes this surprising result, "the dehoaxing experience was apparently crucial; a three-minute revelation that they had been fooled was more powerful than an hour-long denunciation of ESP in producing skepticism toward ESP."

One of Morris' original interests in conducting the experiment was "whether the exercise would teach the students skepticism for ESP statements only, or a more general attitude of skepticism, as we had hoped. For example, would their experience also make them more skeptical of astrology, Ouija boards and ghosts?" Morris did find a slight transfer of skepticism, and from it he concluded hopefully that "teaching someone to be skeptical of one belief makes him somewhat more skeptical of similar beliefs, and perhaps slightly more skeptical even of dissimilar beliefs."

This question of transfer of skepticism is to my mind the critical one. It is of little use to learn a lesson if that lesson always remains a lesson about particulars and has no applicability beyond the case in which it was first learned. What, for instance, was the lesson of the Jonestown incident? Simply that you should never follow Jim Jones to Guyana? Or, more generally, that you should be wary of following any guru halfway across the world? Or that you should never follow anyone anywhere? Or that all cults are evil? Or that any belief in any kind of savior, human or divine, is crazy and dangerous? Is it likely that fundamentalist "Moral Majority" Christians in America would see their attitudes reflected in those of fundamentalist Moslems whose fanaticism they abhor, and that they would there-

THE TASTE BEYOND
12-YEAR-OLD SCOTCH
Discover more.
Let your mature taste lead you from the finest premium Scotch on up to the most expensive 12-year-old Scotch in the world. The Glenlivet. The ultimate in Scotch.
Most premium Scotch is blended and depends on several whiskies for taste and smoothness.
Yet the unblended character of The Glenlivet is noble enough to stand alone. In this distinctive Scotch whisky you'll experience an exceptional smoothness and full-bodied richness, unmatched in all other premium Scotch.
Try the taste beyond premium Scotch.
The Glenlivet
Unchanged since 1824.
The GLENLIVET
AGED 12 YEARS
12
BOTTLED IN SCOTLAND
The GLENLIVET
12 YEARS OLD
Unblended all malt
Scotch Whisky
Distilled and Bottled in Scotland
GEORGE & J. G. SMITH LIMITED

fore be led to renounce their own fanaticism? Why not? At what level of generality is a lesson learned?

Stalker's Captain Ray of Light expressed a faith that by debunking his own "miniature" pseudosciences before audiences he could transfer to people a more general critical ability, an ability to think more clearly about paranormal claims. Is this true? There are many believers in some types of paranormal phenomena who ridicule other types. There are people who will scoff at headlines in *National Enquirer* and at the same time believe, say, that through transcendental meditation you can learn to levitate, or that astrological predictions come true, or that UFO's are visitors from other worlds, or that ESP exists. Many people have said: "Most psychics are, unfortunately, frauds, which makes it all the more difficult for the *genuine* ones to be recognized." You even get believers in tricksters such as Uri Geller who say, "I admit he cheats some of the time, maybe even 90 percent of the time, but I still think he has genuine psychic abilities"!

If one is hunting for a signal in a lot of noise and the more one looks the more noise one finds, when is it reasonable to give up and conclude that there is no signal at all? On the other hand, sometimes there just might be a signal. The problem is, one does not want to jump too quickly to a negative generalization, particularly if one's feelings are based merely on some kind of guilt by association. After all, not everything published in *National Enquirer* is false. The subtle art is in sensing just when to shift, in sensing when there is enough evidence. For better or for worse, however, it is a subjective matter, one that few journals heretofore have dealt with.

The Skeptical Inquirer concerns itself with questions ranging from the ridiculous to the sublime, from the trivial to the profound. There are those who would say it is a big waste of time to worry about drivel such as ESP and other so-called paranormal effects. Others, and I am one of them, believe that anyone who is unable or unwilling to think hard about what distinguishes the scientific system of thinking from its many rival systems is not a devotee of truth at all, and furthermore that the spreading of nonsense is a dangerous trend that ought to be checked.

The question arises, in any case, whether *The Skeptical Inquirer* will ever amount to more than a drop in a huge bucket. Surely its editors do not expect that someday it will be sold alongside *National Enquirer* at supermarket checkout counters! And, carrying this to an upside-down extreme, can you imagine a world where a debunking journal such as *The Skeptical Inquirer* sells millions of copies each week at supermarkets, along with its many rivals, while one lone courageous voice of the occult comes out four times a year in tabloid form and is sought out by a mere 7,500 readers? Where the many rival debunking journals are always to be found lying around in laundromats? It sounds like a crazy story fit for the pages of *National Enquirer*! This ludicrous scenario serves to emphasize just what the hardy band at CSICOP is up against.

What good does it do to publish their journal when only a handful of already convinced antioccult fanatics read it anyway? The answer is found in, among other places, the letters column at the back of each issue. Many people write in to say how vital the magazine has been to them, their friends and their students. High school teachers are among the most frequent writers of thank-you notes to the magazine's editors, but I have also seen enthusiastic letters from members of the clergy, radio talk-show hosts and people in many other professions.

I would hope that by now I have aroused enough interest on the part of some readers for them to want to subscribe to *The Skeptical Enquirer.* The subscription rate is $16 per year, and they should write to Box 229, Central Park Station, Buffalo, N.Y. 14215. In the interest of open-mindedness I also give the address and subscription rate of *The Zetetic Scholar* (Department of Sociology, Eastern Michigan University, Ypsilanti, Mich. 48197, $12 per year) and *National Enquirer* (Lantana, Fla. 33464, $13.95 per year).

Certainly one will never be able to empty the vast ocean of irrationality that all of us are surrounded by, but the ambition of *The Skeptical Inquirer* has never been that great; it has been, rather, to be a steady buoy to which one can cling in that tumultuous sea. It has been to promote a healthy brand of skepticism in as many people as it can. As Frazier said in one of his eloquent editorials, "skepticism is not, despite much popular misconception, a point of view. It is, instead, an essential component of intellectual inquiry, a method of determining the facts whatever they may be and wherever they may lead. It is a part of what we call common sense. It is a part of the way science works. All who are interested in the search for knowledge and the advancement of understanding, imperfect as those enterprises may be, should, it seems to me, support critical inquiry, whatever the subject and whatever the outcome."

It is too bad we should have to constantly defend truth against so many onslaughts from people unwilling to think, but on the other hand, sloppy thought seems inevitable. It is just part of human nature. Come to think of it, I seem to remember reading somewhere recently about how your average typical-type person uses only 10 percent of their brain. Talk about sloppy—it's amazing! Even the scientists are stumped!

BOOKS

The art of mapping, engineers' drawings, life in moving fluids and a cave mystery

by Philip Morrison

THE MAPMAKERS, by John Noble Wilford. Alfred A. Knopf ($20). Colonial surveyors knew the country. "Here are no churches, Towers, Houses, or peaked Mountains to be seen from afar, no Means of obtaining the Bearings or Distances of Places, but by the Compass, and actual Mensuration with the Chain." America was no England; it was "every where covered with woods." So wrote Lewis Evans in 1755 of his *General Map of the Middle British Colonies in America.* Ten years later two sophisticated English experts made their straight way westward from a limestone mark they had set in front of Alexander Bryan's fine home on the Brandywine south of Philadelphia. They followed their axmen, who had cleared a rough swath 10 yards wide through the forest. Along the path came the survey party, steadily plying a chain of 66 iron links and a surveyor's level. Every 10 minutes of longitude they checked their position by the stars; by day the bearing was set by magnetic compass and corrected as they went along by multiple star transits caught at night on their new zenith-sector instrument. For many days of the work a single terse journal entry purposefully reads: "Continued the Line." The line stopped in 1767 after a few years of meticulous work; the axmen refused to cross Dunkard Creek south of Pittsburgh because Shawnee war parties were about and the Six Nations escort said: "You go no further." The last post was "233 Miles, 3 Chains and 28 Links from the Post mark'd West" at the beginning. The Mason and Dixon Line was complete.

Our country is still not mapped like tight little England. To this day the Coast and Geodetic Survey is at work; leveling parties are seen along roads and railroads, wearing orange vests to alert oncoming traffic. They still set new bench marks in special circumstances, but there are already half a million of the bronze disks marking points of known height above mean sea level. Most of the effort in geodesy is now horizontal-control work. The mark of the horizontal-control party in the U.S. is the 100-foot portable steel tower; its acrobatic crew can set it up within the day and strike it even faster. It is double, a tower within a tower, the inner one holding the instruments, the outer one standing free, bearing the weight of the observers and the signal lamps. The tower was devised by signalman Jasper S. Bilby with the help of Aermotor, the famous windmill company, in 1927. Churches and houses seen from afar are still rare in our country, and the wood towers of the time of the clearing of our forests had become too costly. The Coast and Geodetic Survey has three or four mobile parties always at work, using 200 Bilby towers.

The teams are on the road the year around, the original lonely group of rough camping men supplanted now by a few dozen trucks, trailers and mobile homes, complete with families, often with a computer link by telephone to headquarters in Rockville, Md. Today they mount on the tower by day and night not only theodolites but also laser Geodimeters; these electronic time-of-flight instruments work very well. The High-Precision Transcontinental Traverse completed in about 1976 had a loop-closure error in a 1,200-kilometer test of only one part in a million. That is a tenfold improvement in the geodetic mapping of the U.S. Our maps rest still on the North American Datum of 1927; the use of satellite triangulation, now fully warranted after direct comparison with the electronic traverse, will one day take over. The next North American Datum, planned for next year, is a recomputation of the geodetic base, incorporating the field results of the traverse.

These tales come from two of the two dozen 15-page chapters in this winning chronicle of mapmakers over time and space. The histories of cartography on which it largely rests "seem to be more interested," the author comments, "in the maps themselves than in the people who went into the field." John Noble Wilford, an experienced journalist, centers instead on narrative and process; he has used the scholars extremely well, and a large number of his sources are in the primary literature. He talks of maps from Han China to Viking on Mars, and he has listened carefully to the stories told by many a mapmaker of today, workers with computer, aircraft or radio interferometer. Technicalities are avoided, although detail abounds; Wilford has a flair for broad explanation as well as for context. There are half a dozen vignettes of his own experiences with a mapmaking team of our time, a wonderful field survey led by Bradford Washburn to the central Grand Canyon. The book opens with a dizzying shot of theodolite and surveyors on untrodden Dana Butte, a pinnacle 10 feet wide high above the canyon floor but still only halfway up to the rim. Wilford climbed to it in silence one hot June day, the resting helicopter that had brought him to the scene the only sign of humanity, although the vista opens wide in every direction.

The first section of the book is a well-made summary of more familiar history, the growth of mapping from antiquity past the marvels of Alexandria up to Mercator. Here the narrative is more or less standard. The second section is fresher; it treats of the growth of scientific cartography from Charles Marie de La Condamine to the coming of aerial photography. Wilford has combed the accounts to offer a variety of adventures and perceptions not often so well described. The latter section, rather less than half of the book, extends the story to the present, mappers now aloft, or again sounding the sea, or attentive to echoes of microwave pulses from hidden rain-forest lands along the Amazon or from the unseen continental surface buried under the miles-thick Antarctic ice. The final section treats of extraterrestrial mapping, from the moon to Mars.

The strangest tale of all is perhaps that of the Victorian survey of India. The truth is more colorful than *Kim,* that wonderful novel in which the survey enters, half as cover for clandestine intelligence, half as a basis for imperial action. The Great Trigonometrical Survey crossed India from Cape Comorin to Kashmir, more ambitious than any national survey in Europe. The theodolite used was captured at sea in 1802 by the French, who sent it on with their compliments to its owner in India, as a contribution to science. (The Enlightenment had not grasped our modern concept of total war.)

Colonel William Lambton flung his big triangles northward undeterred by the practical road-plotting men of John Company. His successor (Lambton died in his survey tent in the field at 70, with only half of his work done) was George Everest. The name rings worldwide, although Everest may never have seen his mountain. That peak, numbered XV, was a little beyond his last triangles. In 1849 six sight lines had reached it as a matter of routine; the height fell out unexpectedly in the computing office of the survey much later. The peak of course had its Tibetan name; the survey was proud that it added no European names

to well-known and much revered spires of rock. XV, however, was too tempting; it gained its present name in the 1850's.

In Kim and Rudyard Kipling's time no European could enter Tibet. The survey had an answer; maps could be made by Indians in disguise, well trained and sent to map the forbidden land surreptitiously. Young Nain Singh was a man of letters, a pandit, a teacher in a mountain valley. He entered the forbidden Chinese protectorate as a pilgrim lama bound for Lhasa. In his trader's box were a few instruments; his rosary was strung not with 108 beads but with 100, every tenth bead a little larger than the others. His prayer wheel held paper slips for note taking. He attended the Great Lama in awe, and at night he climbed quietly to the roof of his hut to take the latitude with a hidden sextant and a mercury horizon, the small sample he had carried poured into his beggar's bowl. The altitude of Lhasa—off by only 100 meters—he found using a thermometer in his boiling teakettle. The pandit explorers became an institution for a generation, just as they did in *Kim*.

The survey extended its triangles finally in 1911 across the Himalayas to join openly with the Russian net. Maps could now fully include the roof of all the world. More than statecraft and Victorian romance, however, came as a prize to the Indian geodesists. In the 1850's the puzzle of an error in the Great Arc, a serious discrepancy between the triangles and direct surface measurement, was solved. It was widely known that the Himalayas deflected the plumb line, but the fact remained that they did so less than was estimated by a factor of three. Mountains are light and float in denser rock; the Great Arc had sounded the roots of the Himalayas. Today small variations of the gravity field of the earth interest the nations because of the trajectories of missiles.

Few books manage as wide a technical topic so lightly, or supply meaning with as little call on mathematical argument. There are only one or two misleading geometrical remarks. The illustrations do not meet the high standard of the text; although the subjects are well chosen, the reproduction is as a rule quite inadequate to the demand. The radar view of the Amazon region and the environs of Paris mapped in the great *Carte de Cassini*—the first national triangulation—are happy exceptions.

THE ART OF THE ENGINEER, by Ken Baynes and Francis Pugh. The Overlook Press, distributed by the Viking Press ($75). PALLADIO'S ARCHITECTURE AND ITS INFLUENCE: A PHOTOGRAPHIC GUIDE, photographs by Joseph C. Farber, text by Henry Hope Reed. Dover Publications, Inc. ($6.95). Diderot's influential *Encyclopédie* portrayed all the industrial processes of its day in fond detail over hundreds of large, beautiful copper-plate engravings. The tradition was not new; mere mechanisms drawn by Leonardo are regarded today as masterworks. Miners and stonemasons, artillerymen and surveyors, with all their tools of art, crowd many well-loved pages and canvases. Here one sees, for example, a fine two-page engraving, heavily annotated and labeled, of "the engine for raising water by fire," a Newcomen engine installed in London in 1725. Art it surely is, engineering too, but it is not yet the art of the engineer.

Such a discrimination is the central theme of this attractive volume, which arises from and amplifies an exhibition held in Cardiff and London a couple of years back. The clarifying text and the chronological survey of sketches, drawings, plans and photographs with which it is supported amount to an original contribution to the history of the modern mind. All those forerunners of the engineer's real art are part of technical education. They are meant to instruct the reader for some purpose other than the consuming need to get on with the specific job of production. The signs are numerous: scale is often unimportant and materials unclear; the craftsman or the operator is given strong visual place; once in a while omissions that have little visual force would have a gross effect on function.

The art of the engineer is aimed not at instruction but at production. It ranges widely in form and substance. The informal and sketchy notebook work of individual design engineers, who tend to highlight the novel and the difficult, is one end of the spectrum. The other end is the presentation drawing, made of some finished product "as fitted." These drawings border on the ceremonial and celebratory, but they also serve practical ends. They are often the basis of modification and repair. Between the two ends of the spectrum are project drawings, overall accounts close to design drawings but made more impersonally to set rules. Then there are engineering drawings as most people think of them, comprehensive sequences of production details at every level, often within a hierarchical order, usually employing some projective technique of spatial description. The essence of the true art of the engineer is that it provides a visual language and a grammar by which the productive work, not only of a craftsman or two but of a large number of people, can be commanded to one defined end. "The ability to do this is at the root of engineering."

So viewed, the real art of the engineer-draftsman's careful work to serious purpose begins with Matthew Boulton and James Watt of Soho in Birmingham, that first factory for the construction of stationary steam engines (1773). Watt himself originated the style; he drew for a decade or more at home with one assistant. Steam engines were still custom jobs; the building was a part of the engine structure, and some parts were to be procured locally. Watt worked so hard at the drafting table that his partner frowned on small engines; they took Watt's time but brought little income. About half a dozen pages of *The Art of the Engineer* exhibit such drawings; many are colored, the colors keying not only materials but also function. One look at the samples and the rise of a new art comes clear.

There is one predecessor. Ships had long been built from drawings. The shipwrights held enormous traditional skill; most of the processes and details were perforce left to the men who did the work. Only rather general forms and outlines appear on paper. Even these go back in Britain only to the time of Elizabeth's navy; the plans we have up to the middle of the 18th century are more treatises than working drawings. The shipwrights were no friends to centralized design.

The scope of the volume cannot of course cover all engineering. Its center is mechanical engineering for the transport industry. This is a strong limitation only well after the middle of the 19th century. Steamships, locomotives and then aircraft are among the most complex engineering products of the time. Their discrete and variable nature encourages draftsmanship. The pinnacle of the art seen here is surely the *Great Eastern*. Isambard Kingdom Brunel's "great ship" was also documented by many photographs. The party watching the aborted launching of 1857 is one prime image (high-hatted, black-coated, grave gentlemen, one with a large roll of drawings). The big lithographs of the vessel that "exhibit the draughtsman's art at a peak of virtuosity" are unusual. Hand-colored and even textural, the meticulous and elaborate sectional drawings span pages. They were probably made over five or 10 years by a few special artists at the builder's yard; in the end they were public-relations works, made to encourage the nervous directors, already skeptical of profitability!

There is clear evidence that the battles between craftsmanship and design were fought even to the days of iron ships. Brunel demanded compliance with his detailed designs, down to the use of wood screws instead of bolts in certain cabin beams. "'I must have them,' he demanded." In shipbuilding and later in aircraft construction there is a long revolution as wood is supplanted by metal and increasingly the craftsman's hand loses its traditional feedback skills to the designer and the machine tool. Within the past 15 years subdivision and skills transfer has hit the drafting offices. This volume ends with drawings of the VC10, a Vickers–British Aircraft Corporation long-haul transport jet of the mid-1960's. True, there are about 50,000 production drawings for the

The Blaupunkt CR-2010. Richer, purer sound than you ever thought possible in a moving vehicle.

Autosound is only as good as the equipment it passes through. Which means that, in the case of the Blaupunkt CR-2010, the sound is exceptional.

The CR-2010 coddles and shapes highs and lows into a sound as full-bodied, as richly-textured, as anything you'll hear from a home stereo.

4 Channel Amplifier

Blaupunkt increased the conventional two channels to four, each with a maximum output of 7.5 watts. Even when hooked up to a front

The essential controls are fully illuminated.

end, two-speaker system, the CR-2010's crisp reproduction will surprise you. Add two rear speakers and the home stereo effect is complete—sound that surges to new heights of clarity and richness while holding its delicate balance through the magic of a built-in front-to-rear fader.

Holds Signal Longer

Drive away from the signal source of your favorite stereo station and what happens? Reception breaks up into a barrage of crackles and hisses. Not with the CR-2010. Thanks to the "Soft MPX" feature your Blaupunkt automatically shifts reception from stereo to mono *before* the hissing sets in.

Higher Volume without Distortion

The CR-2010 has a pre-amp output jack that lets you bypass the built-in amp and plug directly into a high power amplifier. Yet the boost in volume doesn't come at the expense of distorted sound. And at lower volume the clarity is actually enhanced.

Of course, you get Dolby Noise Reduction—but for FM reception as well as tape. A Sendust Alloy tape head reproduces a fuller range of recorded frequencies.

Hear the Full Blaupunkt Line

The CR-2010 retails for only $396.00** and is the latest in a full line of Blaupunkt AM/FM stereo cassette radios priced from $290.00.**

CR-2010 Features

- **4 channel (4 x 7.5W)**
- **Autoreverse Cassette**
- **Sendust Alloy Tape Head**
- **Auto Hi-Cut Filter**
- **Soft Mute**
- **Soft MPX**
- **Night Illuminated Controls**
- **Dolby Noise Reduction Circuit**

*Dolby is a registered trademark of Dolby Laboratories.

Blaupunkts can be installed in virtually any car, import or domestic. For more information, write:

Robert Bosch Sales Corporation
2800 South 25th Avenue
Broadview, IL 60153

Robert Bosch Canada, Ltd.
6811 Century Avenue
Mississauga, Ontario L5N 1R1

**Suggested U.S. retail prices. Blaupunkt is a registered trademark of Blaupunkt Werke GmbH.

BLAUPUNKT

"YOU CAN ALWAYS TRUST THE TASTE OF A BOLLA"

—Franco Bolla

"Why do so many people prefer the Bolla family of great wines?

"Taste. Not just exceptional taste, but exceptional taste, bottle after bottle. So whether you choose the full-flavored Valpolicella, the mellow Bardolino, the crisp, refreshing Soave or the light, dry Trebbiano, you won't be disappointed. No matter how many times you choose it. After all, when you're a Bolla, you don't let your friends down."

FROM THE BOLLA FAMILY OF GREAT WINES.

IMPORTED BY THE JOS. GARNEAU CO., N.Y., N.Y.

plane. Work is changing, and the drawings reflect the work. The assembly of parts gives way to the numerically controlled machining of ingot; unity has passed beyond the control of the man at the bench.

Even in the drafting room the microprocessor and graphics transformations take over from the draftsman the tedious variations, but they also take over much of the visual record of production and its interest and beauty. Comprehension comes only at a more abstract level; indeed, the drawings begin again to take on more of the tinge of technical instruction. A few of the sample sequences of drawings here are delightful pages of geometric forms; one such page of the 1850's shows a set of some 80 patented railway-engine and carriage wheels, as much variants on a circular theme as a collection of daisies and asters. Another few pages show production drawings of the BE2C of the Royal Aircraft Factory in 1914, a wood-and-canvas biplane still plainly from the hands of the carpenter and the seamstress.

The second book, *Palladio's Architecture and Its Influence,* is an album of the wonderful pillars and porches of the Paduan stonemason Andrea di Pietro della Gondola. He soon rose to architect and then to author of his famous *Four Books of Architecture* (also in print in an inexpensive edition by Dover), which spread his peerless sense of Roman form over the world. (His Vicenza patron, Count Giangiorgio Trissino, renamed his architect Palladio in praise of his Athenian wisdom.) We are shown here his best villas, with a couple of churches and an unusual theater, still in service. Palladio's books built, however, beyond his own place and time. There are buildings after the *Four Books* far from Vicenza, in Whitehall and Covent Garden, at the White House, in Monticello and Charlottesville, all evoking the era of Hadrian by dome, pillar and coffer. If inked lines build in iron and aluminum, they also do so abundantly in elegant works of stone and paint and brick.

Twenty-nine Italian buildings are shown here in excellent, well-reproduced photographs, together with 38 more of British and American structures originally drawn by designers who had read the *Four Books* or their commentators. A Palladian revival is now under way; these photographs, and the knowing introduction by one of the effective contemporary advocates of the classical, should give some foundation for the taste. The provincial Venetian of the 16th century was one who abided by the rules of the ancients more than any other, yet "he is the one who has gained the most attention among the devotees of Modern architecture." Two grotesque fireplaces, with monster heads in stucco or stone agape at the hearth, are startling forms in one Vicenza palace interior, done by Bartolomeo Ridolfi. Is it broken symmetries that in the end lie deepest?

LIFE IN MOVING FLUIDS: THE PHYSICAL BIOLOGY OF FLOW, by Steven Vogel. Willard Grant Press, 20 Providence Street, Boston, Mass. 02116 ($18). Not even a strong rinsing flow is as effective for cleaning a surface as any swipe of a dishcloth. Dust accumulates on whirring fan blades, pipes scale up and thicken instead of wearing smooth and erosion by streaming water is wonderfully enhanced by a little sand. In fact, right at the solid surface there is stillness. Fluids simply do not slip over solid surfaces, whether they are rough or smooth, greasy or clean. The boundary layer always holds a velocity gradient. That was the great contribution of the early 20th century to fluid mechanics. Ludwig Prandtl first recognized that even in turbulent flow the thin layer at the boundary remembers that real fluids are viscous; therein was the solution to the paradoxes of classical hydrodynamics. From that success the eddies flowed, both in the air and in the theory, to bring lift, drag and flight under the laws of Leonhard Euler and Isaac Newton.

Life has known it all along, of course. Organisms dwell in fluid flow, in wind or water, beating surf or the stagnant air at the grass roots. Much depends on what passes the boundary layer and how fast: water, heat, edible particulates and the propagules of many life forms. Diffusion is all very well for bacteria, but once organisms are big enough they can walk and swim. Along the scale, however, something must help the tiny spore and pollen particles to gain access to "the free transit system in the sky."

For doubters there is a fine control study. Club-moss spores were allowed to settle on paper models of leaf surfaces. Then they were blown free in the wind tunnel. It took storm winds to lift the tiny spores off surfaces with a roughness comparable to the spore diameter. Any little elevation helps greatly; in a modest breeze even a millimeter is a real winner. Many are the inventions of life. A barley smut manages to stiffen the infected heads so that they do not bend in the wind, and the spores float free. The famous ballistic fungus *Pilobolus* is not the only gun carrier; its optimization is a delicate trade-off. The gun fires a small spore case, bigger than any spore. Small spores would not go as far, but they would fall out more slowly. Puffballs, poppies, sage, lichens and mistletoes have their own little catapults, some with a power assist from raindrops or the wind itself.

Out on the Pacific coast (perhaps related to the sophistication of the aerospace industry?) there is a colonial kind of sand dollar. It lives on edge in the tidal currents. The lift its cambered form develops is substantial, but with its edge buried firmly in the sand the creature does not move at all. The lift is horizontal, plainly of little interest as a force. The streaming circulation measured by the lift, however, brings the passing particles closer to the feeding appendages on the surface. These echinoderms always cluster, all parallel on edge, a multiwinged arrangement of hydrofoils with their spacing adjusted over a factor of 10 to the surge velocity of the local currents.

"Fluid flow is not currently in the mainstream of biology," writes this author in his personal, iconoclastic way. He has nonetheless written a model of a book, halfway between an introductory physics of flow and an overall sketch of many problems of life in the stream. Whether he is analyzing the simple equations he uses or telling how to measure the drag of an obstacle by indirect means, his sense of aptness and physical understanding comes first. "I once measured the lift of a fixed fly wing by first measuring the airspeed at each of a series of points behind it; then for each point I centered the wake of a tiny wire on the airspeed transducer to get the local wind direction. From these data it was a simple matter to calculate the downward component of momentum."

Perhaps a third of the book is that artful an introduction to the physics of flow—supersonics and hot gases apart—as life encounters it. The Reynolds number is a powerful motif, beautifully modulated; the range examined finds this key dimensionless index from 10^{-7} deep in the tiny filter sieves of the plankton up to 10^8 and more for the free-swimming whale. (How whales can filter feed as well as they do has not yet been studied.) Vogel tells of the drag of trees and weeds and insects, of the life in velocity gradients mentioned above, of flow through pipes, of lift in soaring, gliding, flapping and just spinning down quietly in the autumn twilight. Flow through pipes is treated in sketch outline; it is the most complex, if medically the most important: "I just don't want to dwell," Vogel concedes wisely, "on the nonsteady flow of a non-isotropic, non-Newtonian fluid through nonrigid pipes. Something of the order of 10^4 papers appear on the subject of blood flow each year."

For biologists who want to come to the beginnings of a quantitative understanding of a wide variety of adaptations, for general readers who want to see how fluid mechanics works in a varied and often surprising context, for lecturers and students who want to read a text that seems as personal as a confidential chat, this book, full of data, rich in up-to-date and well-appraised references, is a first-class opportunity.

THE HUNTERS OR THE HUNTED? AN INTRODUCTION TO AFRICAN CAVE TAPHONOMY, by C. K. Brain. University of Chicago Press ($35). "This is a detective story, but a rather odd one. The clues are bones, and the aim of the investigation is to establish the causes of death, but the evidence is ancient and no witnesses survive." Just as unusual is the scrupulous fairness of the narrative. The evidence is placed before the reader at length: there are 70 pages of tables and many pages that are mere lists of numbers, catalogues of the bones under study. The argument is nonetheless clean enough and the conclusion firm enough for a verdict, although by the very nature of the topic there can be no certainty.

The scene is the slopes of a river valley in the open limestone country a few dozen miles to the northwest of Johannesburg. Three beautiful caves in dolomite there have been visited for many decades, although mining for the lime-burning kilns has from time to time encroached on the caverns. Here and there, sometimes in the rubble left behind by the miners, there are tons of bone-bearing cave breccias, bones in plenty cemented into a kind of concrete. The admissible evidence is the bones; there are about 20,000 of them from at least 1,300 animals. It is tedious work to sort and clean the clues.

The key events took place at the time of passage from our apelike walkabout ancestors, the australopithecines, to the larger-brained toolmaking primates usually called *Homo erectus,* members of our own genus. It was a long time ago, maybe a million years, but there are no good dates for the material.

Our detective has a bright record. His method is taphonomy, the study by statistical means not only of the bones that fit into one skeleton but also of the entire assembly of bones found together. It was by this approach that Brain cleared up another suspected crime, in a cave not far away. There the pioneer Raymond A. Dart (his picture is the frontispiece) had unearthed what appeared to be evidence of mass murder. It seemed that the australopithecines there had killed a zoo of other animals, including a few of their own kind. Broken baboon skulls abounded; bone bludgeons were present to explain the damage. The antelope remains included big bones but no tail vertebrae. The tails had not been brought back to the cave, apparent evidence of culture.

The inference was circumstantial and hasty. Brain showed before 1970 that the discarded skeletons of goats killed and eaten in Hottentot settlements today show the same bone distribution. The Hottentots do not sort the bones; it is only that the bones that endure dog chewing and other such processes are the denser and larger ones. The verdict: Not guilty. The early hunters did not sort meat or bone or kill their own kind. The baboons and the occasional hominid were all victims of big carnivores. There remains some doubt; the excava-

tions were never made "with due regard to subtle detail."

The bone accumulations in a dozen caves of the wider region have been sampled. The bone wastes of people, big carnivorous cats, hyenas, owls and porcupines have been examined in the field. Porcupines make gnaw marks; owls collect small bones and deposit them under their roost in the cave mouth. Broken flakes of bone are frequently polished by the feet of animals around a watering place; so it goes. It is detailed attention to clues that is required; the witnesses must be cross-examined.

We come next to the fossil animals of long ago. A fine chapter lists and shows in illustration the animals large and small that lived around the caves. The primates were baboons, ape-men, then men; there are sabertoothed cats, false and true. There is a strange antelope resembling a musk ox, and a hunting form of hyena.

The case is prepared. Now the bones are teased out, identified, counted and sorted by their level in the cave. Among the big bones in two of the caves a clear pattern emerges from the data. There is a sharp change between two levels. In the earlier of the two the bones of the primates, baboon and ape-man, are very frequent. In the later level (it is a pity that no absolute dates are available) we see a switch: primate remains are all but absent, and antelope kind provide the bulk of the bones. (In the third cave there is no strong change seen; the situation remains consistent with the earlier level in the other caves.) The pattern makes good sense in only one way.

The primate bones found in the lower level were not those of animals that had entered voluntarily to die. They were carried in, prey to the carnivores of the lair. It is not clear whether the carnivores were leopards or other species. Possibly an unknown specialized predator on primates was one of the old cats or hyenas, whose skeletons we know but whose habits we can only guess at. It may have been the false sabertooth, a big cat that hunted its prey by stealth. Its powerful teeth would have enabled it to eat all parts of a primate skeleton except the skull.

The case is tentatively solved. "At Sterkfontein, the interface between the top of Member 4 and the bottom of Member 5 represents a time interval crucial in the course of human evolution. During this interval the gracile australopithecines disappeared from the Transvaal scene and the first men appeared.... During Member 4 times the cats apparently controlled the Sterkfontein cave, dragging their australopithecine victims into its dark recesses. By Member 5 days, however, the new men not only had evicted the predators, but had taken up residence in the very chamber where their ancestors had been eaten."

The bones give eloquent witness. Were there fire, weapons, speech, clearer thought? We cannot say. But something new tipped the balance for good, and here we are, a species that, unlike the australopithecines, is truly threatened only by its own kind.

ETHNOGRAPHIC ATLAS OF IFUGAO: A STUDY OF ENVIRONMENT, CULTURE, AND SOCIETY IN NORTHERN LUZON, by Harold C. Conklin, with the special assistance of Puggūwon Lupāih and Miklos Pinther, cartographer. Yale University Press, with the cooperation of the American Geographical Society of New York ($75). The spectacular mirrored staircases of ponded fields climb in 60 or 70 steps to the heights of the narrow mountain valleys. This architectural wonder is not some fountained imperial playground, nor is it the recent product of the dozers and draglines of any diesel-powered heavy engineering firm. It represents the means of subsistence of the Ifugao people on the terraced eastern slopes of the steep cordillera of Luzon, where for 1,000 years they have built for their livelihood. Their resources were ingenuity and wit, baskets and wood paddle spades, much shared toil and plenty of water going downhill.

The culture of this society has long been a topic for ethnographers. We see the people in many photographs in this remarkable study too: the men at wet, heavy work in conical sun hats and not much else, the women in traditional handweaves stooping to transplant the rice seedlings. Here one man points out the subtleties of drainage on a clay model of the lands, there a burial rite musters a band of mourners as precise and uniform as any ballet. The novel center of this work, however, is the Ifugaos' use of the land. The author is a Yale ethnographer long caught by Ifugao uniqueness. He is able to produce the facts because he has been aided by aerial photography specially carried out for him over sample districts of the region. For 20 years he has worked to add indispensable ground truth to the distant view, and he presents the evidence in the form of some 50 maps on every scale that fill the big pages of this outsize atlas.

There are about 100,000 of the people now, spread densely across the rain-beaten slopes for about 1,000 square kilometers. On the average an Ifugao family of five maintains in freehold half a hectare of ponded fields, a quarter of a hectare in dry slopes fire-cleared for the upland sweet-potato crop that shares the role of staple with the paddy rice, and a woodlot of somewhat more than a hectare planted for fuel, woodworking and building materials and grove crops sheltered or supported by the forest trees. Pigs and chickens are kept; the rice is often the glutinous variety needed for the beer essential to ritual and good cheer.

The ponds are rich in yield; the best of them rival the best anywhere, although fertilizer is not used. They hold a foot-deep layer of soil behind the ponding walls, a careful culture medium of puddled clay with various additions. The average family again holds about a kilometer of terrace embankment, of which a third is watertight stone wall a couple of meters high. This is a capital resource the replacement of which would occupy the entire Ifugao work force for 30 years. Of course, such an investment is not made; no savings can be so complete. The resource of the terrace structures has grown up "in small increments" over the centuries. One carbon-14 date for a terrace gives the 16th century; a four- or five-century rate of investment seems reasonable, and other dates push early works back more than 1,000 years. Land titles by genealogies in this jealously private yet cooperative world go back as many as 20 generations. There are traditional forms of mortgage, rental and boundary ritual in wide and ingenious variety.

Water is the key. Rainfall is 10 or even 20 feet per year, brought by the reliable wet-season trades. The fields are hydraulic wonders; one marvelous photograph shows a rain-soaked hillside dotted with little waterfalls from the sluices that cunningly connect each pond to one below. From the headworks on the summit down to the drainage far below the entire enterprise is a masterpiece of water control. Less obvious, but persuasive once it is pointed out, is the Ifugao proposition that "in all terrace work, one never lifts what water can move." It is by temporary canals and ditches that heavy loads of rock and soil are fetched from the heights.

Sluicing the water may call for the cleaning and rewalling of kilometers of channels, with "incredible feats of engineering... accomplished every year." The average family must find each year some 400 man-days of field labor. A day's rice costs about an hour's work overall. This economy works, and works well, superbly adapted to its surroundings, good for many millenniums under constant dynamic care. "There is no evidence that this Ifugao pattern ever required or resulted in a complex bureaucratic organization, a widely based form of political integration, or recognition of a centralized authority." There is no sign of a single prime mover, no master plan, no imported cultural package, no new migration. It is a hydraulic society of clever and energetic smallholders that has grown locally in the hills of Luzon. A demographic study over time is plainly the next step.

This atlas seems indulgent, with a sheet size much larger than the big world atlases. True, the maps are rich with data, current and historical, but they would not be cramped at a more modest size.


AMERICA'S TRUCK.
BUILT FORD TOUGH.

The Country Music Foundation, Nashville, Tennessee, is proud to announce its Official Archive Collection

THE GREATEST COUNTRY MUSIC RECORDINGS OF ALL TIME

Unprecedented and unequaled in our time — the complete and definitive collection of America's country and western music.

- ***The best of 75,000 records from the Foundation archives and the vaults of every country music label.***
- ***The first collection to include all the great country artists.***
- ***Featuring all the great hits, the milestone performances, out-of-issue pressings and unreleased recordings.***

For the first time in history, the greatest recorded performances of country music's greatest artists will be brought together in a single, definitive collection. It will include the most important recordings by the leading country artists of today . . . landmark performances by legendary greats . . . and all the diverse and varied musical styles that have enriched America's country music.

This remarkable collection is the official issue of the Country Music Foundation — home of the Country Music Hall of Fame and Museum, and the world's largest library of recorded country music. And it is the first complete collection that the Foundation has ever issued honoring *The Greatest Country Music Recordings of All Time.*

A unique collection that only the Country Music Foundation could assemble

To assemble this collection, the staff of the Country Music Foundation carefully reviewed the Foundation's own archives of 75,000 records. In addition, they enlisted the support of all the country music record companies — whose vaults hold many of the master recordings selected for this collection. And they were able to obtain rare recordings from private collectors and country music artists themselves.

As a result, the Country Music Foundation Official Archive

"Country is the music of the people. Songs of the soil, forsaken and fulfilled love. Story songs whose music is both contemporary and timeless . . . I love it, and I am proud to be part of the first collection to tell the whole country music story."
— Johnny Cash

© 1982 FMRS

Collection is unique both in scope and importance. A collection that would be difficult — or impossible — for any individual to duplicate.

All the great performers

The collection will feature milestone recordings from the careers of country music's most important artists. Such *great contemporaries* as Loretta Lynn, with 'Blue Kentucky Girl' and 'Coal Miner's Daughter.' Kenny Rogers and The First Edition, with 'Ruby, Don't Take Your Love to Town.' Johnny Cash with 'I Walk The Line' and 'Sunday Morning Coming Down.' Dolly Parton, with 'Coat Of Many Colors' and 'My Tennessee Mountain Home.' The "outlaw" music of Willie Nelson. The Nashville sound of Chet Atkins and Eddy Arnold. Country rock, with The Charlie Daniels Band. And country classics by popular music artists Linda Ronstadt, Glen Campbell and Anne Murray.

Also included will be the unforgettable recordings of such long-time *favorites* as Hank Snow, Ernest Tubb and Merle Travis. The *legendary giants:* Hank Williams, Patsy Cline, Jim Reeves, Flatt and Scruggs, and Jimmie Rodgers. And recordings that reflect regional influences and evolving musical styles — *bluegrass, Cajun, country gospel, western swing, honky tonk* and *rockabilly.*

The collection will include such *rare recordings* as Vernon Dalhart's 1924 recording of 'The Prisoner's Song' — country music's first million selling record, and Loretta Lynn's early classic 'Honky Tonk Girl' — now out of issue. And from the Foundation's archives will come *previously unreleased recordings* — studio "takes" never before made generally available.

Records of superior quality

Every step has been taken to ensure the *technical excellence* of the collection. Thus, all of the *early* recordings will first undergo a painstaking restoration process in the Country Music Foundation's newly opened Audio Restoration Laboratory. Here, recordings of classic performances will be electronically "cleaned" groove-by-groove to eliminate extraneous surface noise and preserve the original sound.

To produce the records, the Foundation has appointed The Franklin Mint Record Society — judged by audio experts to be a leader in producing records of superior quality. The vinyl used will be of a special formula containing its own anti-static element. This material, together with the careful process by which the pressing is made, results in a record that is more rigid, durable and resistant to dust. A true *proof-quality record* — providing exceptional tonal fidelity and clear, clean sound when played through any of today's audio systems.

In the Foundation's audio laboratory, the latest electronic techniques are used to restore the original sound quality of early recordings. Then, the records are pressed in a dust-free "clean room," where strict production standards, and audio and visual inspection, assure high quality.

Fascinating musical "program" in each album

Custom hardbound albums have been designed to house and protect all 100 proof-quality records in this collection.

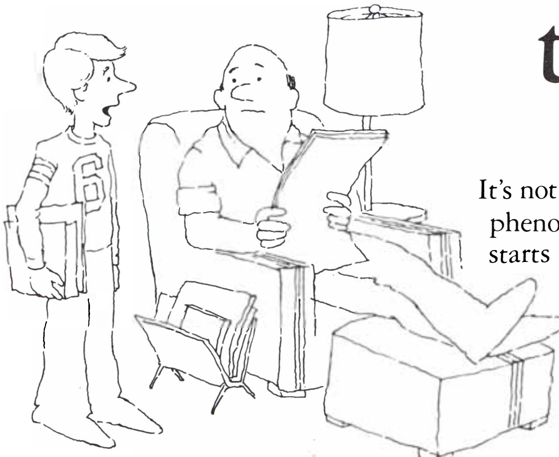
Each album will contain four long-playing records, presenting a program of artists and recordings carefully selected by the Foundation, and unique to this collection. And each album will be accompanied by specially prepared program notes, illustrated with photographs from the Foundation's permanent collection.

Available by subscription only

If you have a long-standing interest in America's country music . . . or are just discovering its rich and exciting sound . . . this Official Archive Collection is an indispensable treasure. A complete, comprehensive and authoritative collection of the greatest recordings in country music — on records of exceptional fidelity.

The collection may be acquired only by direct subscription to The Franklin Mint Record Society, Franklin Center, Pennsylvania 19091. It will not be sold through record stores. To enter your subscription, simply complete and return the accompanying application. Please note it should be mailed by February 28, 1982.

"Dad, can I use the IBM computer tonight?"



It's not an unusual phenomenon. It starts when your son asks to borrow a tie. Or when your daughter wants to use your metal racquet. Sometimes you let them. Often you don't. But when they start asking to use your IBM Personal Computer, it's better to say yes.

Because learning about computers is a subject your kids can study and enjoy at home.

It's also a fact that the IBM Personal Computer can be as useful in your home as it is in your office. To help plan the family budget, for instance. Or to compute anything from interest paid to calories consumed. You can even tap directly into the Dow Jones data bank with your telephone and an inexpensive adapter.

But as surely as an IBM Personal Computer can help you, it can also help your children. Because just by playing games or drawing colorful graphics, your son or daughter will discover what makes a computer tick—and what it can do. They can take the same word processing program you use to create business reports to write and edit book reports (and learn how to type in the process). Your kids might even get so "computer smart," they'll start writing their own programs in BASIC or Pascal.

Ultimately, an IBM Personal Computer can be one of the best investments you make in your family's future. And one of the least expensive. Starting at less than $1,600† there's a system that, with the addition of one simple device, hooks up to your home TV and uses your audio cassette recorder.

To introduce your family to the IBM Personal Computer, visit any ComputerLand® store or Sears Business Systems Center. Or see it all at one of our IBM Product Centers. (The IBM Data Processing Division will serve business customers who want to purchase in quantity.)

And remember. When your kids ask to use your IBM Personal Computer, let them. But just make sure you can get it back. After all, your son's still wearing that tie.

IBM®

The IBM Personal Computer and me.

†This price applies to IBM Product Centers. Prices may vary at other stores.

For the IBM Personal Computer dealer nearest you, call (800) 447-4700. In Illinois, (800) 322-4400. In Alaska or Hawaii, (800) 447-0890.

Established 1845 SCIENTIFIC AMERICAN February 1982 Volume 246 Number 2

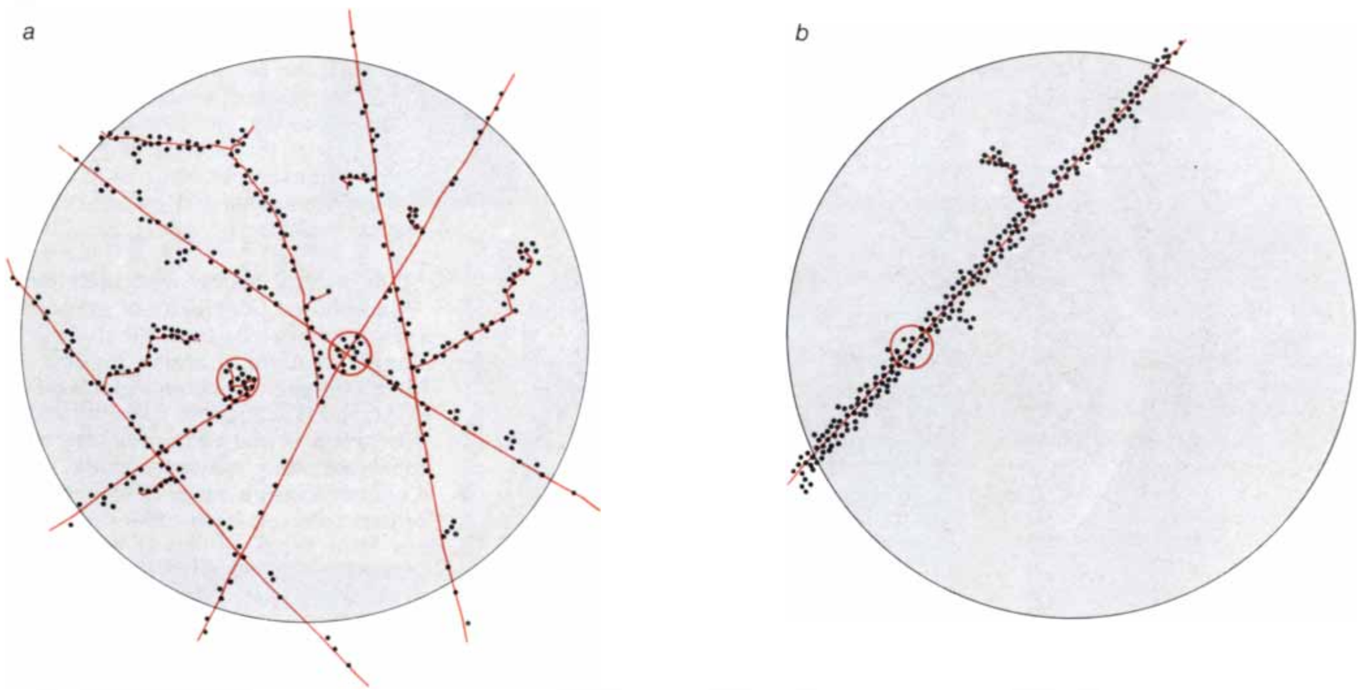
The Biological Effects of Low-Level Ionizing Radiation

What is the hazard to human beings of ubiquitous low-level radiation from natural and artificial sources? The evidence so far indicates that compared with other hazards it is slight

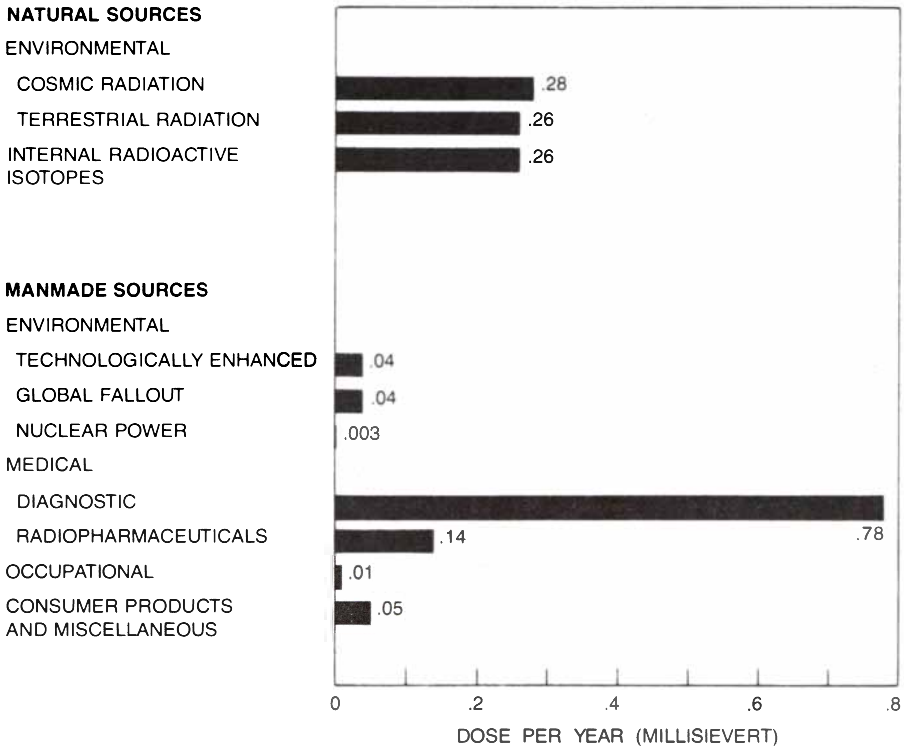
by Arthur C. Upton

Everyone is bathed throughout life with low-level ionizing radiation from natural sources. For the average person the lifetime exposure is virtually doubled by radiation from manmade sources, chiefly medical. Since radiation can cause cancers and genetic defects, the question arises of exactly how hazardous the low-level radiation from background and manmade sources is. The question has been intensively investigated for many years, and the evidence assembled so far indicates that the hazard, if it is real, is too small to be detectable. Hence compared with other hazards to which people are regularly exposed, the overall hazard from low-level ionizing radiation would not appear to be any cause for concern.

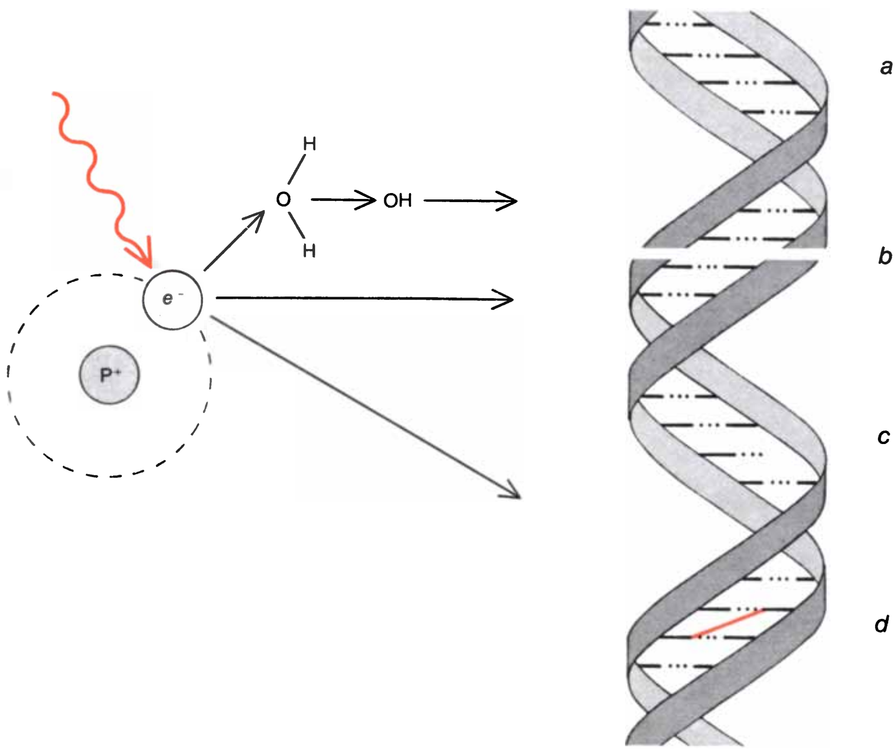
A related question has to do with thresholds. Is the effect of low-level ionizing radiation cumulative, or is ionizing radiation harmful only if it exceeds some threshold level? On this question the evidence is more ambiguous. It is clear that for certain types of radiation and certain effects the impact of radia-



ENERGY OF RADIATION impinging on a living cell is dissipated in different ways depending on the nature of the radiation. X rays and gamma rays generally have a low rate of linear energy transfer: they penetrate deeply but generate ions sparsely along their path. Particulate radiation, such as electrons and alpha particles, has a high rate of linear energy transfer. Here the two types are depicted schematically for equal doses of gamma rays (*a*) and particulate radiation (*b*). The dots represent ionizations caused by the dissipation of energy along radiation tracks (*colored lines*). High concentrations of energy-dissipation events occur in both cases within small volumes (*colored circles*) but per unit dose are less frequent in *a* than in *b*. The figure is based on work by G. W. Barendsen in the Netherlands.



RADIATION EXPOSURE of a typical person in the U.S. from natural and manmade sources is represented in millisieverts. One sievert is the amount of any radiation equal in biological effect to 100 rads (a rad is 100 ergs per gram of tissue) of gamma rays. The total average dose rate per year is .8 millisievert from natural sources and 1.06 millisieverts from manmade sources. A typical dental X ray delivers about one millisievert to the center of the cheek.



DAMAGE TO DNA is the most critical effect of low-level radiation. The effect can be direct or indirect. At the left a hydrogen atom, consisting of one proton and one electron, is ionized. (The electron is dislodged when the atom absorbs a photon of radiation.) In the indirect effect (*top arrow*) a secondary electron interacts with a molecule of water (H_2O) to give rise to a free radical (OH), which does the damage to the DNA. In the direct effect (*lower arrows*) the electron itself interacts with the DNA. The schematic representation of DNA at the right shows a normal segment of the molecule (*a*) and three of the many types of damage that can be caused by either direct or indirect effects of ionization: a double-strand break in the DNA double helix (*b*), the deletion of a base (*c*) and the chemical cross-linking of the two DNA strands (*d*).

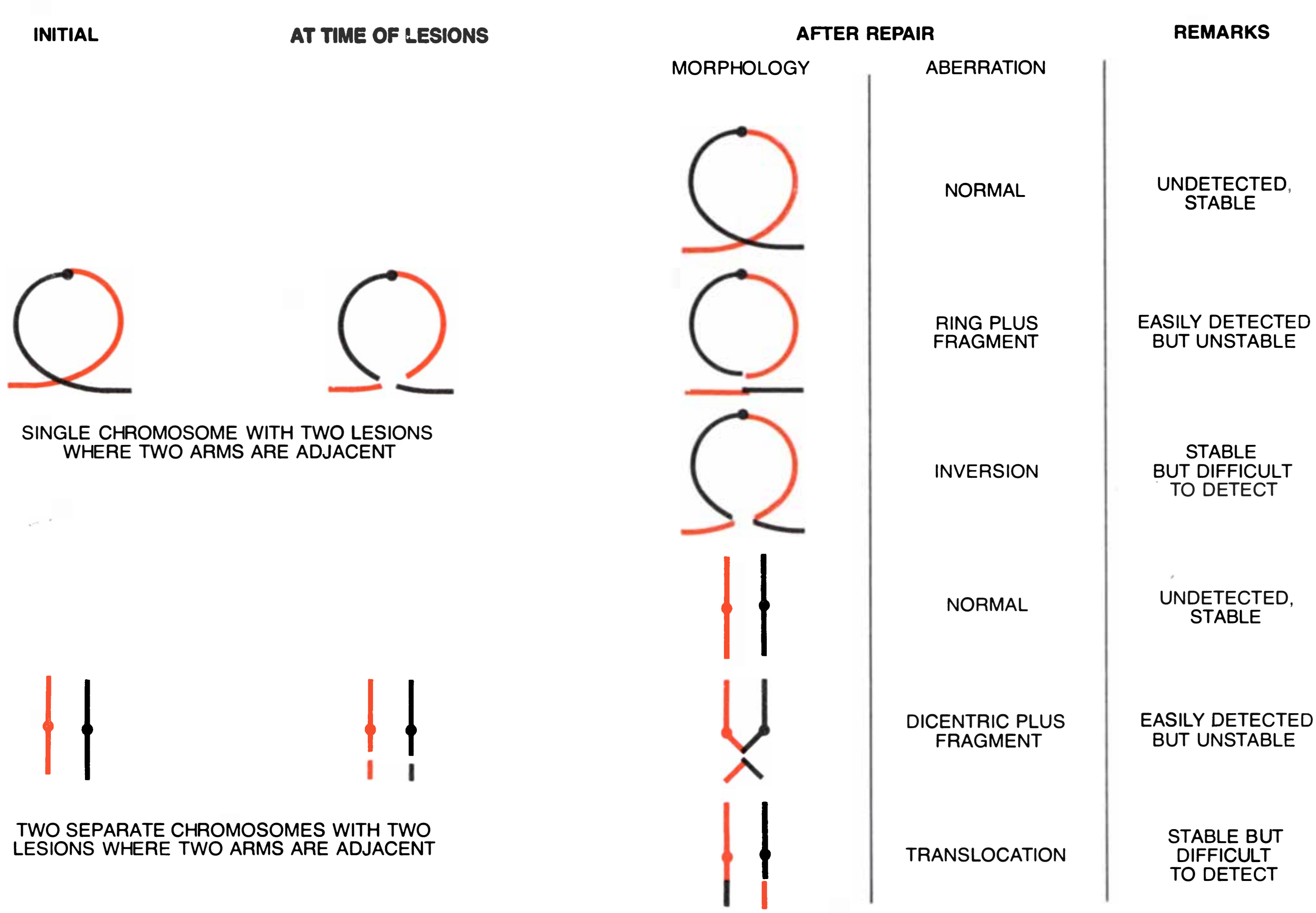
tion is indeed cumulative. For others, however, it seems more plausible to postulate a threshold.

Ionizing radiation, in contrast to other forms of radiant energy, can impart enough localized energy to an absorbing material to ionize (remove electrons from or add them to) atoms and molecules in its path. Some ionizing radiation is electromagnetic and some consists of subatomic particles of various masses and charges. The electromagnetic ionizing radiation includes X rays and gamma rays, which are characterized by relatively short wavelength and high frequency. The particulate radiations include electrons, protons, neutrons and alpha particles.

Quantities and doses of ionizing radiation are measured in several units. The oldest unit is the roentgen, which is a measure of the quantity of ionization induced in air. The principal units used to express doses absorbed by living matter are the rad (one rad equals 100 ergs per gram of tissue) and the gray (one gray equals one joule per kilogram of tissue, or 100 rads). Since for a given dose particulate radiations ordinarily cause more injury than X rays or gamma rays, the units called the rem and the sievert have been introduced so that doses of different kinds of radiation can be compared for their biological effects. One rem, defined loosely, is the amount of any radiation that produces a biological effect equivalent to that resulting from one rad of gamma rays. Similarly, one sievert is the amount of any radiation that is equivalent in biological effect to one gray of gamma rays; one sievert therefore equals 100 rems. The units employed for expressing collective doses are the person-rem and the person-sievert, each of which is obtained by a multiplication: the average dose per person times the number of people exposed. For example, one rem to each of 100 people equals 100 person-rems or one person-sievert.

The natural background radiation is composed principally of (1) cosmic rays; (2) emissions from the disintegration of uranium, thorium, radium and other radioactive elements in the earth's crust, and (3) emissions from potassium 40, carbon 14 and other radioactive isotopes occurring naturally in the body. Collectively the average dose to internal organs received from these three sources by a person living at sea level is about .8 millisievert (80 millirems) per year. A dose twice that size may be received by people living at higher elevations, where cosmic rays are more intense, or in regions where the radium content of the earth is fairly high.

The largest contribution from manmade sources comes from medical diagnosis, which is estimated to approach the dose received from the natural background. (The contributions from differ-



DAMAGE TO CHROMOSOMES by radiation is often repaired by the body but is sometimes misrepaired, resulting in chromosomal abnormalities. The situation is depicted for a single chromosome (*top*) with a radiation-induced lesion in each arm and for two separate chromosomes (*bottom*), each with a lesion in one arm. If the two breaks are not too close in space and time, the repair mechanism restores the chromosome to its normal condition. If such aberrations are close in space and time, abnormalities can arise from various kinds of misrepair. This depiction is based on work by Malcolm L. Randolph and J. G. Brewen of the Oak Ridge National Laboratory.

ent kinds of medical and dental examination vary widely.) Other manmade sources include radioactive minerals that are present naturally in certain types of crushed rock, building materials and phosphate fertilizers; radiation-emitting components of television sets, smoke detectors and other consumer products; fallout from atomic weapons, and leakage from nuclear reactors.

X rays were applied in the diagnosis and treatment of disease within months of their discovery by W. K. Röntgen in 1895. Harmful side effects of the radiation were encountered almost immediately. The first such effects, which often appeared within hours or days after exposure, were transitory injuries such as reddening and blistering of the skin. By 1902 it was recognized that a delayed effect of radiation could be the development of cancer. The first such case of cancer to become known arose at the site of long-standing skin damage on the hand of a radiologist. In ensuing decades, before adequate safety measures were adopted, scores of similar cancers appeared among pioneer workers with radiation. Because such cancers were characteristically preceded by a long history of progressive radiation-induced skin damage it was generally assumed that radiation gave rise to cancers only if the dose was large enough to cause gross injury. This was a form of the threshold concept.

The possibility that cancers might result from small doses of radiation was not considered seriously until the 1950's, when E. B. Lewis of the California Institute of Technology inferred (from the heightened frequency of leukemia in survivors of the atomic bombings of Japan, in radiologists and in certain groups of patients treated with X rays for noncancerous conditions) that the incidence of cancers might increase as a linear, nonthreshold function of the dose. Such a dose-effect relation seemed consistent with the hypothesis that leukemia could result from a radiation-induced mutation in a single blood-forming cell. On this concept Lewis postulated that from 10 to 20 percent of all cases of childhood leukemia might be attributable to natural background radiation.

Lewis' interpretation, implying that no dose of radiation is without some carcinogenic risk, generated concern and controversy that have continued to this day. Although the existence of a threshold for carcinogenic effects is no longer postulated in setting limits of exposure to ionizing radiation, the magnitude of the risk that may be attributable to low-level radiation continues to be under inquiry.

Heritable effects of ionizing radiation (effects transmitted to descendants) have long been postulated to have no threshold and therefore to represent a risk associated with low-level radiation. On the other hand, certain nonheritable effects have been documented clinically only at high levels; they therefore appear to have a threshold. Included in this category are the opacification of the lens of the eye, the impairment of fertility and the depression of the number of white cells in the blood. Similarly, radiation-induced disturbances in the development of the embryo apparently result from doses large enough to kill many cells in the affected tissues and so are regarded as unlikely to be caused by exposure at the low levels characteristic of natural background radiation.

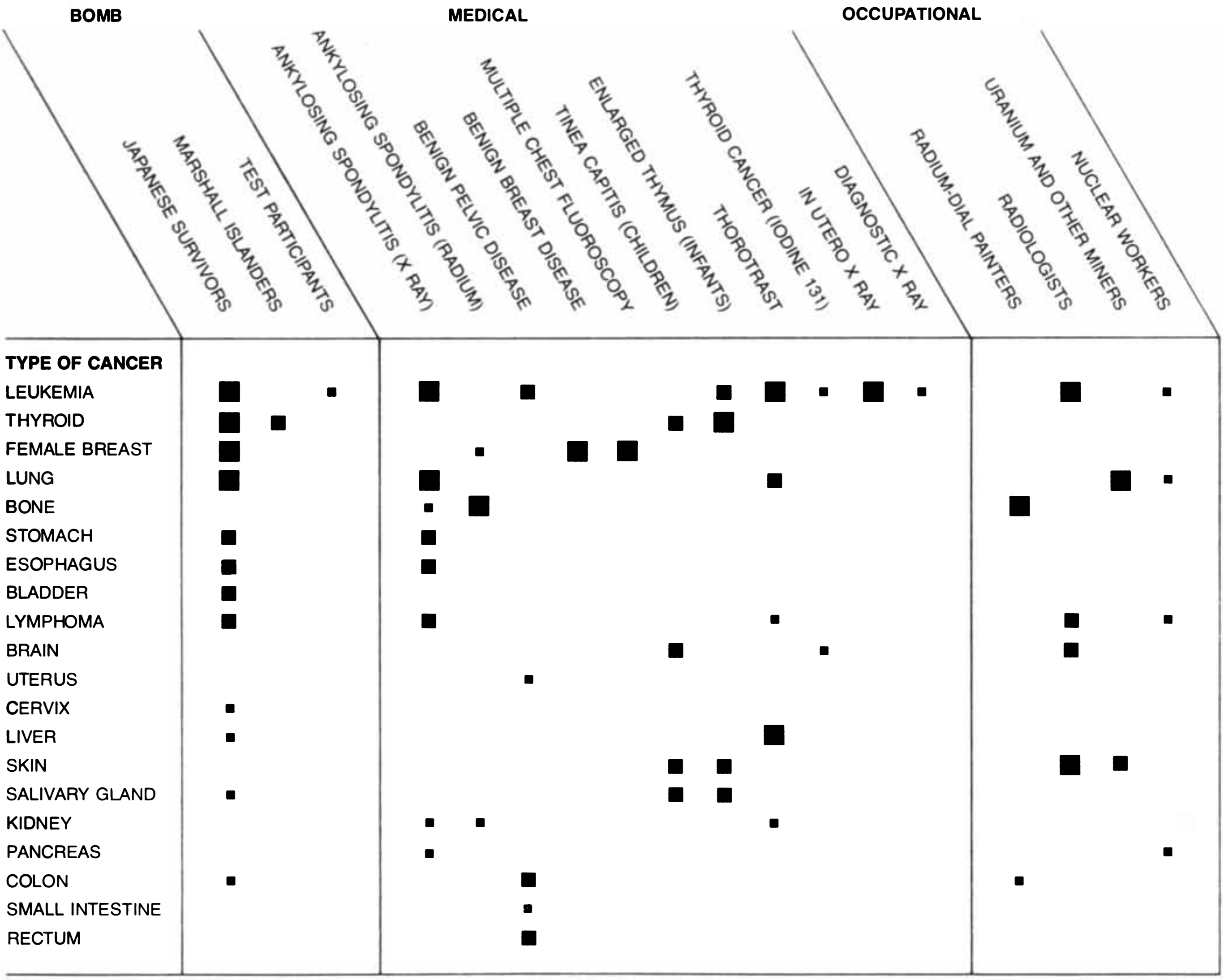
Carcinogenic and heritable effects therefore constitute the principal health hazards associated with small doses of radiation. The nature of these effects and their relation to an individual's exposure to radiation must be analyzed in assessing the risks of low-level irradiation. Such an analysis calls for somewhat more detailed consideration of the interactions of ionizing radiation with living cells.

As ionizing radiation penetrates living tissue it gives up its energy through a series of random collisions and interactions with the atoms and molecules in its path. The collisions and interactions give rise to ions and reactive chemical radicals, which break chemical bonds and cause other changes in molecules in their vicinity. The distribution of ionization events along the path of the impinging ray or particle depends on the energy, mass and electric charge of the radiation, and to a lesser extent on the density of the absorbing tissue. In general X rays and gamma rays (which are electrically neutral) are characterized by a low rate of linear energy transfer, that is, they generate ions sparsely along their tracks and penetrate deeply into tissue. Charged particles are characterized by a higher linear energy transfer and shallower penetration. The capacity to cause damage is correlated with the density of the deposition of energy. Radiations of high linear energy transfer (such as protons and alpha particles) are generally more damaging than radiations of low linear energy transfer (such as X rays and gamma rays).

An injury to a cell results from the molecular changes caused by radiation-induced ions and free radicals. Among the many types of molecule that are affected by ionizing radiation the most critical is DNA because of the limited redundancy of the genetic information encoded in it. The total amount of energy deposited by an acutely lethal dose of X rays affecting the whole body (about three to five sieverts) causes hundreds of breaks in DNA molecules in every cell of the body.

The damage to DNA along a track of radiation with low linear energy transfer is likely to consist of fairly simple lesions such as single-strand breaks in the two-strand double helix of DNA. Severer lesions are likely with radiation of higher linear energy transfer. Since the simple types of lesion can to a considerable extent be repaired by enzymes in the cells, the damage to DNA that ultimately results in a given instance may depend as much on the effectiveness of the repair processes as it does on the nature of the initial lesion. The effects of an unrepaired or misrepaired lesion in a single molecule of DNA may be amplified many times as the DNA is transcribed and translated; ultimately it can be transmitted to countless numbers of daughter cells. If only for this reason, damage to chromosomes and genes appears to figure more prominently than any other type of damage in the injuries resulting from low-level irradiation.

Chromosomal abnormalities constitute one of the most thoroughly studied effects of radiation. The abnormalities include changes in the number and structure of chromosomes. They result from the breakage and rearrangement



INCIDENCE OF CANCER associated with radiation is charted for people exposed to radiation in three different circumstances: the atomic bombings of Japan, radiation therapy and occupational exposure. A strong association of a certain type of cancer with a certain type of exposure is indicated by the large squares, a meaningful association by the squares of medium size and a suggestive but unconfirmed association by the small squares. The three groups are the ones in which the exposure to radiation is fairly well documented.

of chromosome fibers and from interference with the normal segregation of chromosomes when the cell divides.

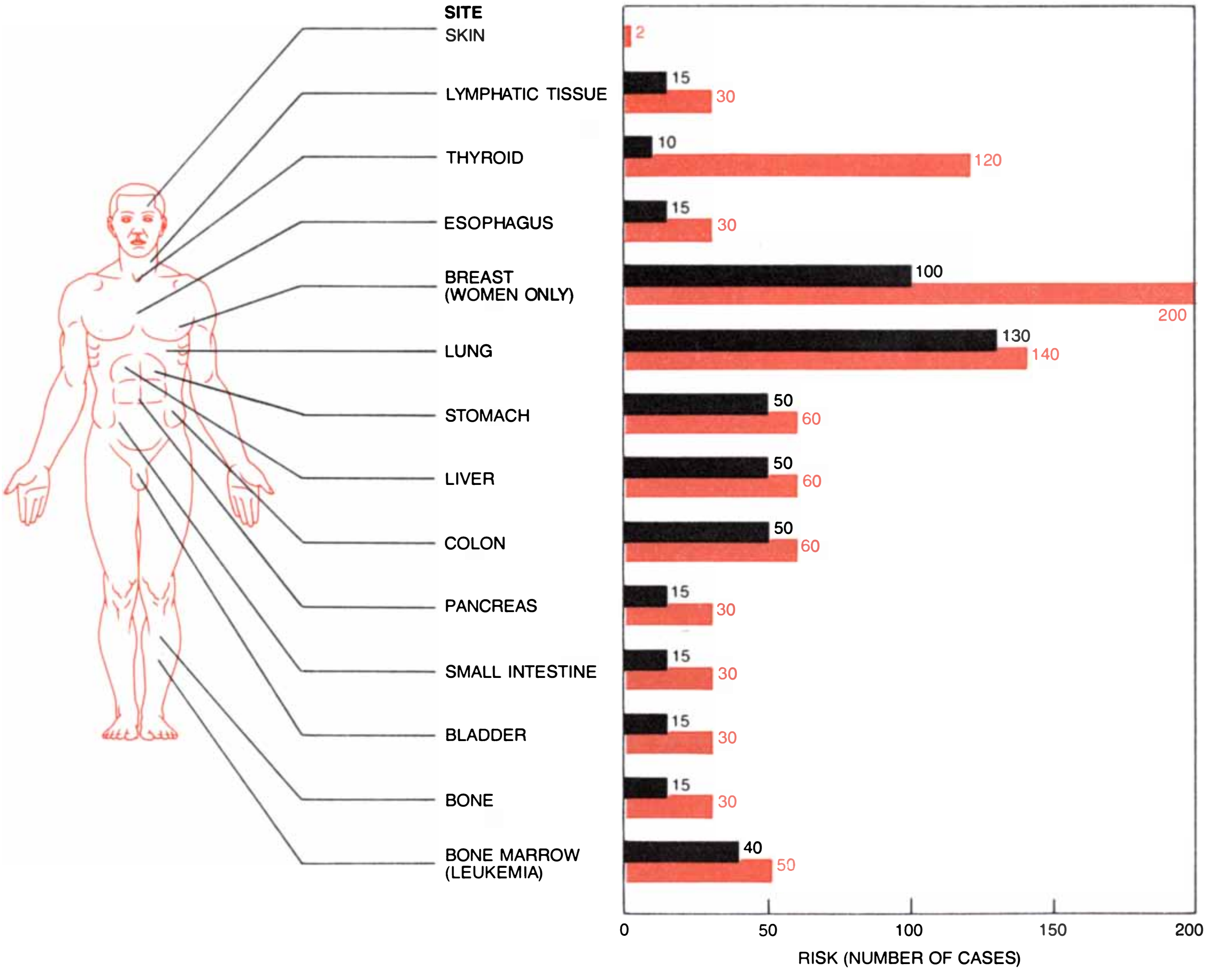
With metabolic repair processes a cell can rejoin the ends of a broken chromosome. When two breaks are close enough together in space and time, however, the broken ends from one break may be joined incorrectly with those from another, giving rise to translocations, inversions, rings and other types of structural rearrangement. The frequency of such chromosomal aberrations increases as a linear, nonthreshold function of the radiation dose in the low-to-intermediate range. When the number of chromosomal injuries is plotted as a graph, the slope of the line is steeper for radiation with high linear energy transfer than it is for radiation with low energy transfer. With increasing dose and dose rate the slope for low-transfer radiation increases, but it remains essentially unchanged for high-transfer radiation.

From these dose-effect relations it can be inferred that a chromosomal aberration has a small probability of being caused by low-transfer radiation unless two or more radiation tracks traverse a given region of the cell nucleus in close enough succession to enable the chromosomal lesions arising along one track to interact with those along another. In contrast, the lesions resulting from a single traversal of high-transfer radiation have a high probability of interacting with one another, independently of the lesions resulting from another traversal.

The frequency of chromosomal aberrations in lymphocytes (one type of white blood cells) taken from the circulating blood and maintained in a culture medium is detectably increased in people who work with radiation and in other people exposed to radiation. The combined frequency of such aberrations, scored soon after irradiation in vitro, approximates .1 per cell per sievert, from which the dose required to double their frequency can be calculated to approximate .05 sievert, or about 60 times the dose ordinarily received per year from natural background radiation. Such aberrations are also increased in frequency by certain viruses, chemicals and drugs. The majority of such aberrations are unstable in that they interfere with division in the affected cells; such cells fail to reproduce and gradually disappear.

The mutagenic action of radiation on genes has been studied extensively in a wide variety of organisms since its discovery in 1927 by H. J. Muller. Although no heritable effects of radiation have yet been demonstrated in human populations, the wealth of relevant information from other species provides a basis for predicting such effects in man. The most pertinent data come from experiments with mice.

In these experiments the dose-effect relation for the induction of mutations



LIFETIME RISK of various types of cancer from low-level radiation is shown on the basis of estimates from a number of investigators. The risk is per 10,000 person-sieverts, that is, the equivalent of one sievert of radiation to each of 10,000 people over a lifetime. The figures given are the maximum estimates for fatal cancers (*black*) and number of cancer cases (*color*). In each instance the estimates actually cover a range in which the minimum figure is much lower (for example, 15 cases and 15 deaths from bone-marrow cancer).

in spermatogonia and oocytes (maturing sperm and eggs) is quite similar to the one for chromosomal aberrations. With high-transfer radiation the frequency of mutations increases steeply in proportion to the dose but is relatively independent of the dose rate. (The dose is the total amount of radiation energy deposited in a tissue or cell regardless of time; the dose rate is the amount of radiation energy per unit of time.) With low-transfer radiation the frequency of mutations increases less steeply as a function of the dose but is highly dependent on the dose rate.

In considering these matters one's attention turns to the genetic effects of the atomic bombings and the natural background radiation. No detectable increase in genetic abnormalities has appeared among the children of people who survived the two atomic bombings. The failure to detect an increase is not incompatible with the induction rate observed in mice, considering the smallness of the sample (78,000 children) and the low average dose (.5 sievert) to the gonads of the parents.

The information available at present suggests that the dose required to double the human mutation rate lies between .2 sievert and 2.5 sieverts. On the basis of this value assessments have been made of the contribution of the natural background radiation to the frequency of various types of genetically determined disease in the population. Although such assessments are highly uncertain, since it is difficult to relate the amount of genetic disease in the population to the mutation rate alone, they imply that only a small amount (from .1 to 2 percent) of all genetically determined disease is attributable to the natural background radiation.

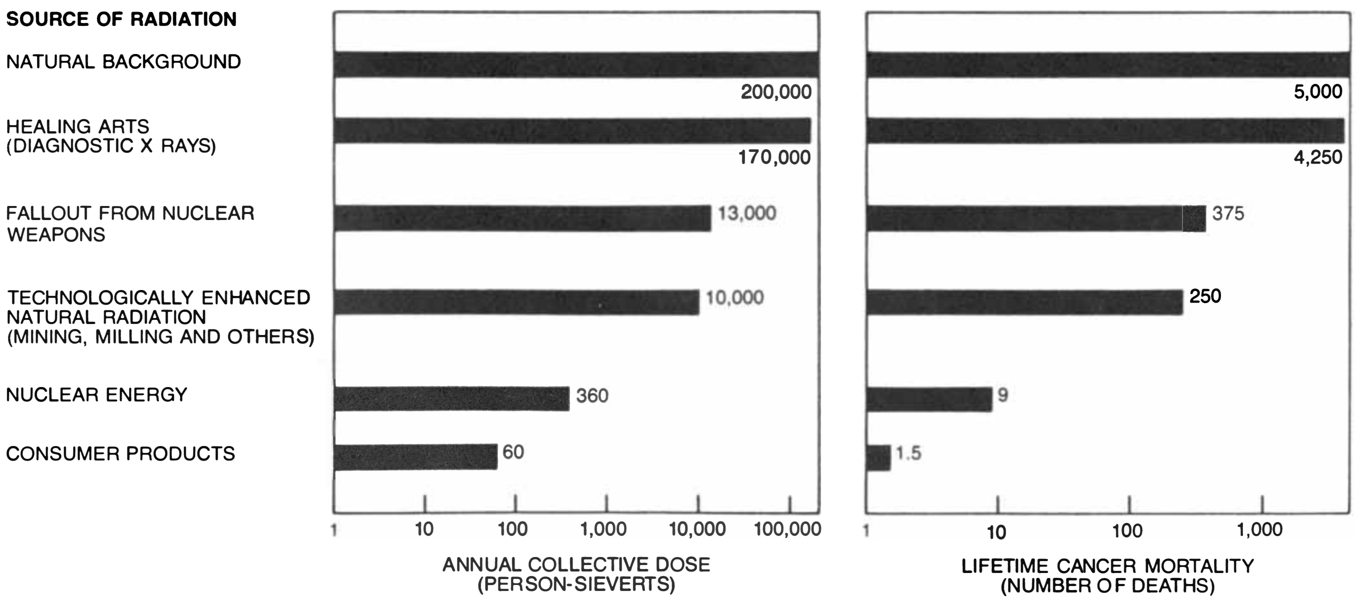
Most but not all types of cancer are increased in frequency by radiation. The increase depends on the conditions of exposure and on factors affecting the susceptibility of the exposed population. A serious difficulty is that radiation-induced cancers have no distinguishing features by which they can be recognized as such. Moreover, they do not appear until years or decades after the irradiation. A causal relation between a cancer and previous irradiation can therefore be inferred only on the basis of the appropriate epidemiological evidence. The three best sources of such evidence are studies made of people who survived the atomic bombings, patients exposed to radiation for medical purposes and workers exposed to radiation in their occupations.

Because the frequency of cancer of any one type is relatively low the analysis of the mathematical relation between dose and incidence calls for the study of sizable populations. Furthermore, because of the large number of subjects and the long time elapsing between irradiation and the appearance of cancer such studies are usually complicated by uncertainties about the size of the radiation dose to affected individuals and the possible influence of risks other than radiation. Nevertheless, different irradiated populations show a remarkable consistency in the extent to which their rates of cancer vary in relation to the estimated dose of radiation. At the same time the data come mainly from fairly high doses of radiation (.5 sievert to two sieverts), so that the carcinogenic risks of low-level radiation can be estimated only by an extrapolation based on tentative assumptions about the dose-incidence relation.

The information about this relation is more extensive for leukemia and for breast cancer in women than it is for any other human cancer. The data come from several populations and cover a considerable range of doses. For leukemia all types of the disease except the chronic lymphatic type increase in incidence with dose. In the first 25 years after irradiation the total cumulative excess of all types combined amounts to approximately one case per year per 10,000 persons at risk per sievert of radiation to the bone marrow.

The consistency from group to group in the figures for breast cancer is noteworthy in view of the differences among populations in the duration of exposure. Irradiation was almost instantaneous for people exposed in the atomic bombings, was generally received in from one to four exposures on successive days by women under radiation therapy and was accumulated over many months (in increments of from .01 to .05 sievert) by women who received repeated fluoroscopic examinations of the chest during treatment for pulmonary tuberculosis. (The treatment entailed the collapsing of a lung; the fluoroscopic examinations were made to ensure that the lung had not begun to expand again.) The fact that the excess of cancer cases per unit dose is not detectably smaller in the third group than it is in the others suggests that successive, small, widely spaced exposures are fully additive in their cumulative effect on the tissue of the breast. The absence of detectable repair between such small and widely spaced exposures is consistent with a linear nonthreshold relation between dose and incidence.

Further evidence of carcinogenic effects from low doses is available from



CONTRIBUTION OF RADIATION EXPOSURE to the burden of fatal cancers in the U.S. population is charted on the basis of a number of estimates. The annual dose from each radiation source is expressed in person-sieverts. The total number of fatal cancers is less than 10,000 annually, or about 2.5 percent of the annual mortality in the U.S. as a result of cancers from all known and unknown causes.

studies of thyroid cancer following radiation therapy and from an association between prenatal irradiation and childhood cancer. In several independent studies the incidence of thyroid tumors has been recorded as increasing after treatment of the head and neck with X rays in childhood. Two of the studies showed the excess of tumors to be evident at doses to the thyroid as low as from .06 to .2 sievert.

The association between prenatal irradiation and childhood cancer may not be a causal one, but other causes for it have yet to be demonstrated. Taken at face value, the data imply that diagnostic X rays during intrauterine development may increase a child's risk of developing cancer before puberty by as much as 50 percent. Since the average dose in the circumstances is only from .01 to .05 sievert, the number of excess cancers suggests that the embryo and the fetus are relatively susceptible to the carcinogenic effects of radiation.

Although many other types of cancer have been observed to be increased in frequency in irradiated populations, the dose-incidence data for such cancers are generally limited. Nevertheless, the data indicate that the carcinogenic effects of a given dose are usually larger with high-transfer radiation than they are with low. They also indicate that the excess of cancers per unit dose varies appreciably among cancers at different sites in the body.

SITUATION	CAUSE OF DEATH
TRAVELING 700 MILES BY AIR	ACCIDENT
CROSSING THE OCEAN BY AIR	CANCER FROM COSMIC RAYS
TRAVELING 60 MILES BY AUTOMOBILE	ACCIDENT
VISITING FOR TWO MONTHS IN DENVER	CANCER FROM COSMIC RAYS
LIVING FOR TWO MONTHS IN A STONE BUILDING	CANCER FROM RADIOACTIVITY
WORKING FOR 1.5 WEEKS IN A TYPICAL FACTORY	ACCIDENT
WORKING FOR THREE HOURS IN A COAL MINE	ACCIDENT
SMOKING FROM ONE TO THREE CIGARETTES	CANCER, HEART-LUNG DISEASE
ROCK-CLIMBING FOR 1.5 MINUTES	ACCIDENT
20 MINUTES BEING A MAN AGED 60	MORTALITY FROM ALL CAUSES

RISK FROM RADIATION is compared with the risk from other situations or activities. In each case the risk is a one-in-a-million chance of dying as a result of the exposure. The illustration is based on work by E. E. Pochin in Britain and Richard Wilson of Harvard University.

Analyses of the available data by groups studying the biological effects of ionizing radiation for the United Nations and the U.S. National Academy of Sciences have resulted in estimates that the total excess of cancers of all types is from four to 18 (from two to 10 of them fatal) per year per 10,000 people at risk per sievert, beginning some two to 10 years after whole-body irradiation and persisting thereafter for the remaining lifetime of the population at risk. The figures correspond to a cumulative lifetime excess of from 140 to 1,000 cancers (from 70 to 500 fatal ones) per 10,000 at risk per sievert.

These estimates have been criticized as understating the actual risk, possibly by a factor of 10, because they appear to neglect certain studies of populations exposed to low doses of radiation in which more cancers have been observed than could be predicted from the estimates. These studies have included workers at the nuclear plant in Hanford, Wash., nuclear workers at the naval shipyard in Portsmouth, N.H., military personnel exposed to atomic-weapons tests in Nevada, children downwind from a nuclear-test site in Utah and patients who had been given X rays in medical diagnosis.

Suffice it to say that none of the studies convincingly correlates the cancers in question with reliable estimates of radiation dose, nor does any of them adequately take into account the possible influence of exposure to cigarette smoke, drugs, industrial chemicals or other potentially confounding variables. In the opinion of most experts the studies therefore do not argue persuasively for increasing the present estimates of risk from radiation. Indeed, many experts see the present estimates as exaggerating the risk by failing to allow adequately for the repair of injury at low doses and low dose rates that has been observed in tumors induced in experimental animals.

With radiation of low linear energy transfer there are radiobiological grounds for expecting the risk per unit dose to decrease with decreasing dose and dose rate. In such instances, therefore, extrapolation by means of a linear nonthreshold model would tend to overestimate the risk of exposure to low levels of this radiation, and somewhat lower estimates would be in order. With radiation of high linear energy transfer, on the other hand, the data imply that the opposite may be true, since the carcinogenic effect of this radiation appears to increase with decreasing dose and dose rate.

Without a better understanding of the mechanisms of different forms of radiation-induced cancer any hypothetical dose-incidence model will probably be a gross oversimplification. With respect to doses that are in the range from intermediate to large, research on experimental animals indicates such doses may contribute to the risk of different cancers through diverse effects on endocrine balance, immunity, cell turnover and other physiological functions, with dose-incidence relations that can vary accordingly. Evidence that the dose-incidence relation may vary similarly among human cancers is provided by the data for skeletal tumors and for leukemia and breast cancer. Plotting the incidence on a graph yields similar curves for leukemia and breast cancer but a quite different curve for skeletal tumors.

Since the number of cancers attributable to low-level radiation in adults is small compared with the number attributable to other causes, any epidemiological study seeking to verify risk estimates for doses of from one to 50 millisieverts would have to include hundreds of thousands of subjects. Hence future efforts to refine human-risk estimates should not just rely on empirical observations but should include research into the mechanisms of carcinogenic effects in laboratory animals, cultured cells and other experimental systems.

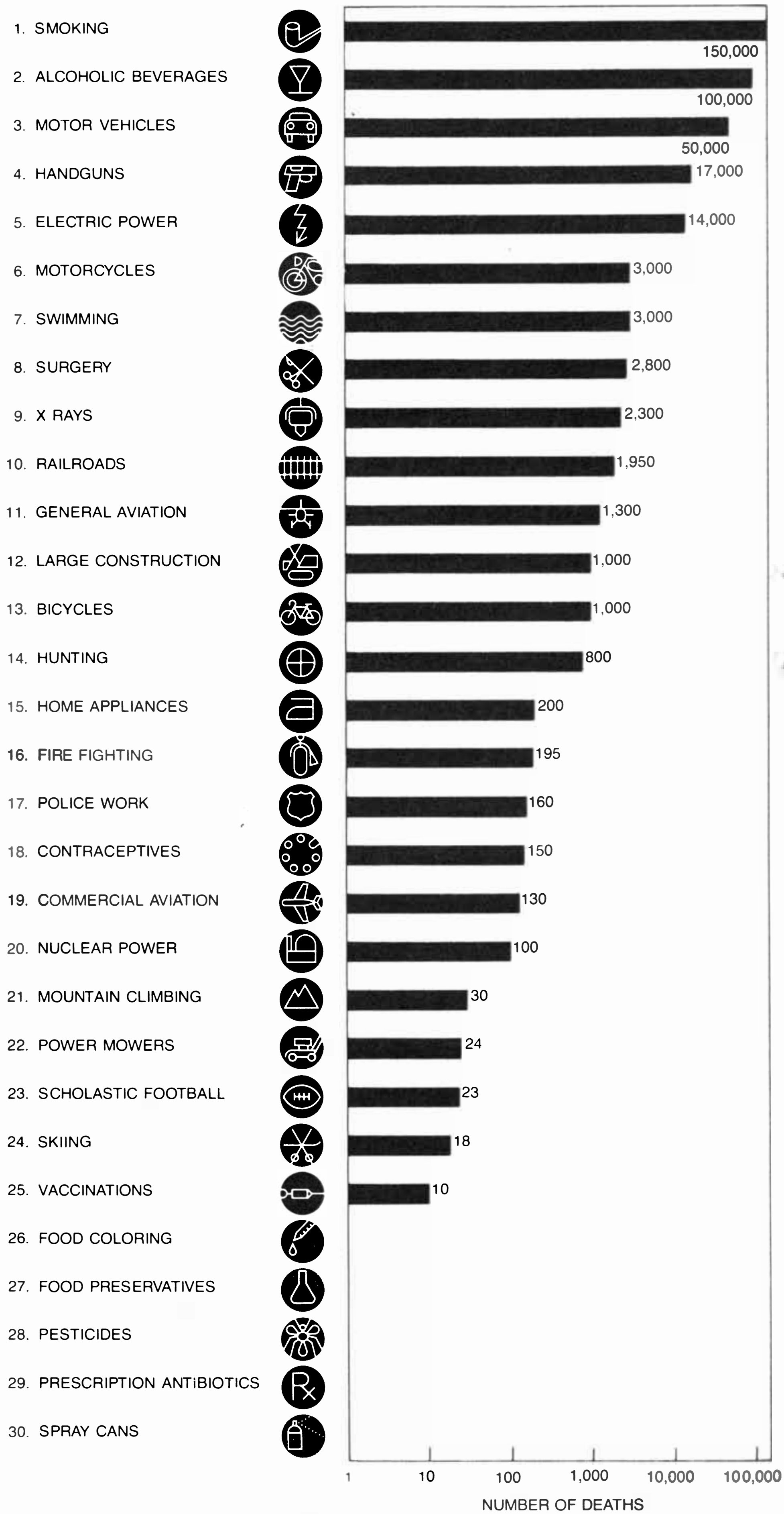
On the basis of the current risk estimates and on the assumption that risks vary as a linear nonthreshold function of the dose down to the level of the natural background radiation, it can be inferred that only a small fraction (from 1 to 3 percent) of all cancers in the general population are attributable to the natural background radiation. The inference is in consonance with empirical comparisons of cancer rates in geographic areas differing in background radiation levels. The estimates also imply that an appreciable fraction (perhaps 20 percent or more) of lung cancers in nonsmokers may be attributable to the inhalation of naturally occurring radioactive isotopes in the air.

The number of cancers attributable to occupational irradiation represents in most groups of radiation workers an increase of less than 1 percent over the natural incidence. Hence the average occupational risk to radiation workers, expressed in terms of life expectancy lost through carcinogenic and heritable effects combined, is comparable to the risk in occupations usually classified as being safe.

According to the view that no amount of radiation is entirely without effect, some risk must be presumed to be associated with any activity involving low-level irradiation, no matter how small the dose may be. Some such risks, compared with the other hazards of daily life, seem negligible [*see illustration on preceding page*]. Still, no risk can be regarded as acceptable if it is readily avoidable or if it is not accompanied by a commensurate benefit. In general the acceptability of a risk tends also to decrease to the extent that it is involuntary, particularly if it is imposed on those who do not share equitably in the related benefits. The weighing of risks and benefits calls for personal value judgments, which can vary widely among individuals.

Although the estimation of the health hazards arising from low-level radiation is a scientific problem, the acceptability of such hazards clearly depends on the perceived as well as the real magnitude of the risks and the associated benefits. Therefore the process of deciding about the acceptability of radiation hazards is one in which the society as a whole must be represented. Sound policy decisions will depend on public understanding of the issues as well as on the data provided by science.

Moreover, it is noteworthy that the problems encountered in evaluating the risks of low-level radiation are paralleled by similar problems in evaluating the risks of low-level exposure to toxic chemicals. The latter problems are even more complicated: they involve not only poorly defined variations among individuals and among species in the uptake, distribution, metabolism, detoxification and excretion of substances but also other sources of uncertainty. Nevertheless, the scientific bases for assessing the risks of radiation and chemicals are sufficiently similar for the concepts and principles derived from efforts to understand and protect against the biological effects of radiation to be helpful in developing strategies for protection against the health impact of toxic chemicals and other potentially harmful environmental agents.



PERCEPTION OF RISK varies widely among different groups of people. The chart at the right reflects a survey in which members of three groups were asked by Decision Research to rank the 30 sources of risk listed in the first column. There the sources are ranked according to their actual annual contribution to the number of deaths in the U.S. as they have been determined by actuarial estimates. That contribution appears in the second column; the remaining columns record the respective rankings by the three groups.

LEAGUE OF WOMEN VOTERS

NUCLEAR POWER
MOTOR VEHICLES
HANDGUNS
SMOKING
MOTORCYCLES
ALCOHOLIC BEVERAGES
GENERAL AVIATION
POLICE WORK
PESTICIDES
SURGERY
FIRE FIGHTING
LARGE CONSTRUCTION
HUNTING
SPRAY CANS
MOUNTAIN CLIMBING
BICYCLES
COMMERCIAL AVIATION
ELECTRIC POWER
SWIMMING
CONTRACEPTIVES
SKIING
X RAYS
SCHOLASTIC FOOTBALL
RAILROADS
FOOD PRESERVATIVES
FOOD COLORING
POWER MOWERS
PRESCRIPTION ANTIBIOTICS
HOME APPLIANCES
VACCINATIONS

COLLEGE STUDENTS

NUCLEAR POWER
HANDGUNS
SMOKING
PESTICIDES
MOTOR VEHICLES
MOTORCYCLES
ALCOHOLIC BEVERAGES
POLICE WORK
CONTRACEPTIVES
FIRE FIGHTING
SURGERY
FOOD PRESERVATIVES
SPRAY CANS
LARGE CONSTRUCTION
GENERAL AVIATION
COMMERCIAL AVIATION
X RAYS
HUNTING
ELECTRIC POWER
FOOD COLORING
PRESCRIPTION ANTIBIOTICS
MOUNTAIN CLIMBING
RAILROADS
BICYCLES
SKIING
SCHOLASTIC FOOTBALL
HOME APPLIANCES
POWER MOWERS
VACCINATIONS
SWIMMING

BUSINESS AND PROFESSIONAL CLUB MEMBERS

HANDGUNS
MOTORCYCLES
MOTOR VEHICLES
SMOKING
ALCOHOLIC BEVERAGES
FIRE FIGHTING
POLICE WORK
NUCLEAR POWER
SURGERY
HUNTING
GENERAL AVIATION
MOUNTAIN CLIMBING
LARGE CONSTRUCTION
BICYCLES
PESTICIDES
SKIING
SWIMMING
COMMERCIAL AVIATION
ELECTRIC POWER
RAILROADS
SCHOLASTIC FOOTBALL
CONTRACEPTIVES
SPRAY CANS
X RAYS
POWER MOWERS
PRESCRIPTION ANTIBIOTICS
HOME APPLIANCES
FOOD PRESERVATIVES
VACCINATIONS
FOOD COLORING

The Fossil Footprints of Laetoli

At this site in Tanzania thousands of animal tracks, including those of predecessors of man, are found in volcanic ash that fell some 3.5 million years ago

by Richard L. Hay and Mary D. Leakey

Near Lake Eyasi in Tanzania is a series of layers of volcanic ash notable for having yielded the remains of early hominids that are among the oldest known: they date back between 3.5 and 3.8 million years. The layers of ash hold an even more unusual example of preservation: fossil footprints. Several tens of thousands of animal tracks have now been discovered in these ash deposits. The survival of these normally ephemeral traces, ranging from early hominid footprints to the trail of a passing insect, gives a vivid glimpse of life on the African savanna well before the Pleistocene epoch. How did they escape obliteration?

The extensive formation known to geologists as the Laetolil Beds is exposed over some 1,500 square kilometers on the Eyasi Plateau, an uplifted fault block northwest of Lake Eyasi. The beds overlie ancient basement rocks of Precambrian age and are themselves bordered and overlain to the east by several large volcanoes. The exposures richest in fossils are found in the smaller area called Laetoli, about 70 square kilometers in extent, that lies at an elevation of 1,700 to 1,800 meters on and near the drainage divide that runs between Lake Eyasi and Olduvai Gorge to the north.

Fossils are found mainly in the upper 45 to 60 meters of the beds, which at Laetoli are at least 130 meters thick. About three-fourths of the upper part of the formation consists of eolian tuffs: beds of volcanic ash that was redeposited by the wind after it had fallen. Most of the other ash beds, which alternate with the eolian tuffs, are "air fall" tuffs, that is, deposits of ash that remained essentially undisturbed after it had settled out of the eruptive cloud. Consisting largely of particles the size of fine sand grains (.125 to .25 millimeter) and medium-size sand grains (.25 to .5 millimeter), the eolian ash buried animal bones and teeth, bird eggs, land snails and other objects exposed on the ground. All the ash, eolian and air-fall, came from one volcano: Sadiman, about 20 kilometers east of Laetoli.

In 1935 the area was explored by Louis Leakey, the geologist Peter Kent and one of us (Mary Leakey). The party found that the fossil animal remains at Laetoli were older than those discovered in the lowest level (Bed I) at Olduvai Gorge. In 1938–39 an expedition headed by Ludwig Kohl-Larsen of the University of Tübingen collected a large number of fossils in the area, including a fragment of a hominid upper jaw. In 1974 one of us (Mary Leakey) found additional hominid remains in a better state of preservation. As a result systematic investigations of the area have continued from 1975 to the present. They have yielded numerous hominid fossils, chiefly lower jaws and teeth, and a wide variety of other animal remains.

In 1976 Andrew Hill of Harvard University first came on animal tracks in a bed of tuff that since then has been called the Footprint Tuff. The next year Peter Jones of the University of Oxford and Philip Leakey found rather poorly preserved footprints, probably of hominid origin, in the same bed, and in 1978 Paul I. Abell of the University of Rhode Island discovered an unmistakably hominid footprint in the tuff at another place. Clearing of the surface layer there revealed more hominid footprints in two long parallel trails.

Several of the tuff beds contain the mineral biotite, which is rich in potassium. This makes it possible to date those beds by the potassium-argon method, which is based on the decay of the radioactive isotope potassium 40 into the noble gas argon. Working with samples of biotite from a bed under the lowest fossil-bearing deposits at the site and other samples from near the top of the formation, Garniss H. Curtis and Robert Drake of the University of California at Berkeley obtained potassium-argon dates that bracketed the Footprint Tuff. The samples from below it indicated that the tuff there was 3.8 million years old; the ages of samples from the Footprint Tuff itself and from above it clustered closely around 3.5 million years.

In general the animals preserved as fossils at Laetoli are similar in type to the animals found in the area today. Listed in order of their decreasing abundance, the commonest vertebrate remains are those of bovids (antelopes and related forms), lagomorphs (specifically hares), giraffes, rhinoceroses, horses, pigs and two kinds of proboscideans (elephants and dinotheres, a form now extinct). Significantly absent from the tuff deposits are the remains of crocodiles, hippopotamuses and other water-dwelling animals.

The air-fall tuffs that alternate with the eolian tuffs in the Laetolil Beds number about 50. Many of them are found throughout the fossil-rich area and can thus serve for stratigraphic correlations. Most of the air-fall tuffs range from one centimeter to 10 centimeters in thickness and are unlaminated; each tuff was evidently the product of a single eruption of Sadiman. Three of them, however, are laminated. They range from 12 to 30 centimeters in thickness and are evidently the result of a series of

PARALLEL TRACKS of hominid footprints extend for 25 meters across cemented volcanic ash at Site *G*, one of 16 localities at Laetoli where the tracks of various animals are found in the ancient series of ash layers known collectively as the Footprint Tuff. The tracks to the right of of the hominid trails are those of an extinct three-toed horse, *Hipparion*. The hominid trail at the left was made by the smallest of the three upright walkers; the other was made by a large individual whose prints were partially obliterated by a smaller individual who stepped in them. Whether the hominids walked together or at different times cannot be judged. The photograph appears through the courtesy of John Reader, © National Geographic Society.

PORTABLE SUNSHADE offered some protection to Mary Leakey (*left*), R. J. Clarke of the National Museum of Bloemfontein (*center*) and two technicians as they cleared overlying tuff from the hominid trails at Site *G* in 1979. Trails were discovered in 1978.

PARALLEL LINES, some six centimeters long, mark the passage of an insect across Laetoli ash. The insect may have been a dung beetle.

PAW PRINT, measuring 3.5 centimeters from front to back, is another of the thousands of tracks at Laetoli. It was made by a cat.

closely spaced eruptions. Whereas the ash particles forming the eolian tuffs are the size of fine sand to medium sand, the particles in the air-fall tuffs range in size from fine sand to coarse sand (.5 to one millimeter). Consisting of fragments of glassy lava, the particles are low in silica and contain the calcium-rich mineral melilite.

An additional constituent of most of the air-fall tuffs and perhaps all of them is carbonatite ash. This igneous material consists of calcium carbonate, with or without sodium carbonate. Carbonatite ash is highly unusual. Only one volcano in the world, Oldoinyo Lengai, 90 kilometers north of Laetoli, is known to have erupted carbonatite in recent times. The Oldoinyo Lengai carbonatite and the carbonatite of the Laetolil Beds are of the type that is rich in sodium carbonate. This fact provides a clue to the preservation of the tracks in the Footprint Tuff.

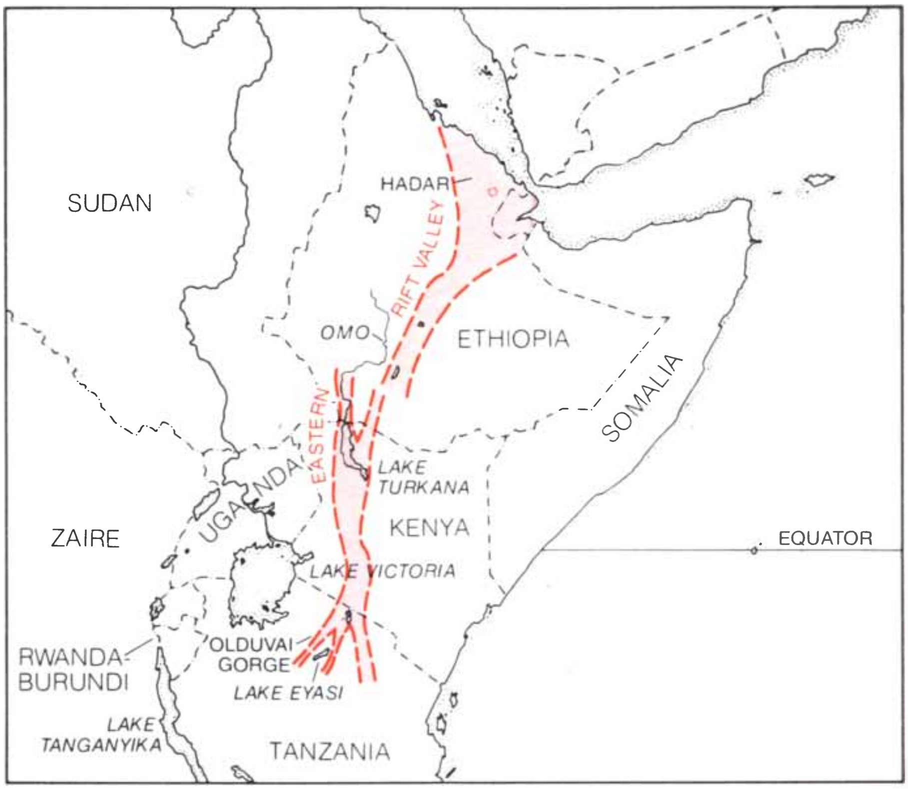
A carbonatite that contains both calcium carbonate and sodium carbonate reacts with water in a characteristic way. The calcium carbonate becomes a fine precipitate and the sodium carbonate goes into solution. When the solution evaporates, crystals of the mineral trona form; the crystals provide an instant cement for the ash layers.

The Footprint Tuff is one of the three laminated air-fall tuffs in the Laetolil Beds. Fossil footprints have now been found in all three of the laminated tuffs, and they are particularly abundant and widespread in this one. Generally between 12 and 15 centimeters thick, the Footprint Tuff lies near the top of the fossil-bearing zone and is overlain by a more massive deposit of ash, 50 centimeters thick in places. The thin ash layers that make up the Footprint Tuff accumulated in the early stages of an eruptive episode; the thicker overlying tuff represents the culmination of the same episode.

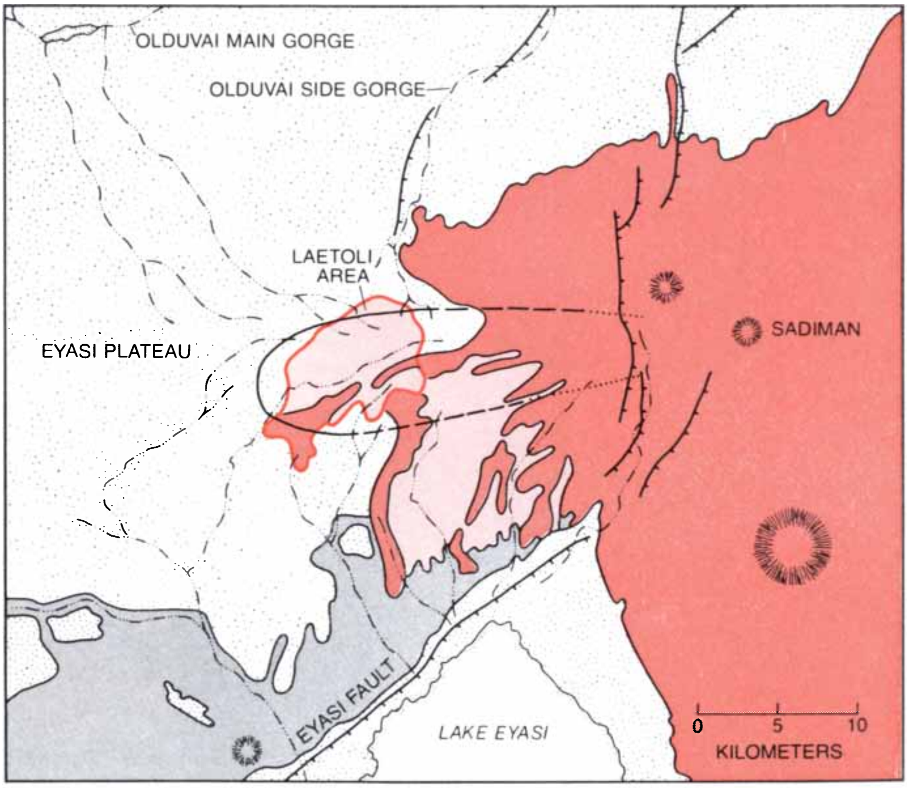
The Footprint Tuff is divisible into two major units. The lower of the two is seven to 10 centimeters thick; the upper is four to six centimeters thick. The contact between the units is sharp, and the top of the lower unit is eroded in places. The two units differ in their lithology, in their structure and in the footprints they contain.

The lower unit is subdivided into 14 thin layers, ranging from two to 15 millimeters in thickness. The layers are generally recognizable over the entire Laetoli area; each is very likely the product of a single volcanic eruption. Most of the layers cover surface irregularities such as footprints with little change in thickness. This shows that they remained essentially undisturbed where they were deposited.

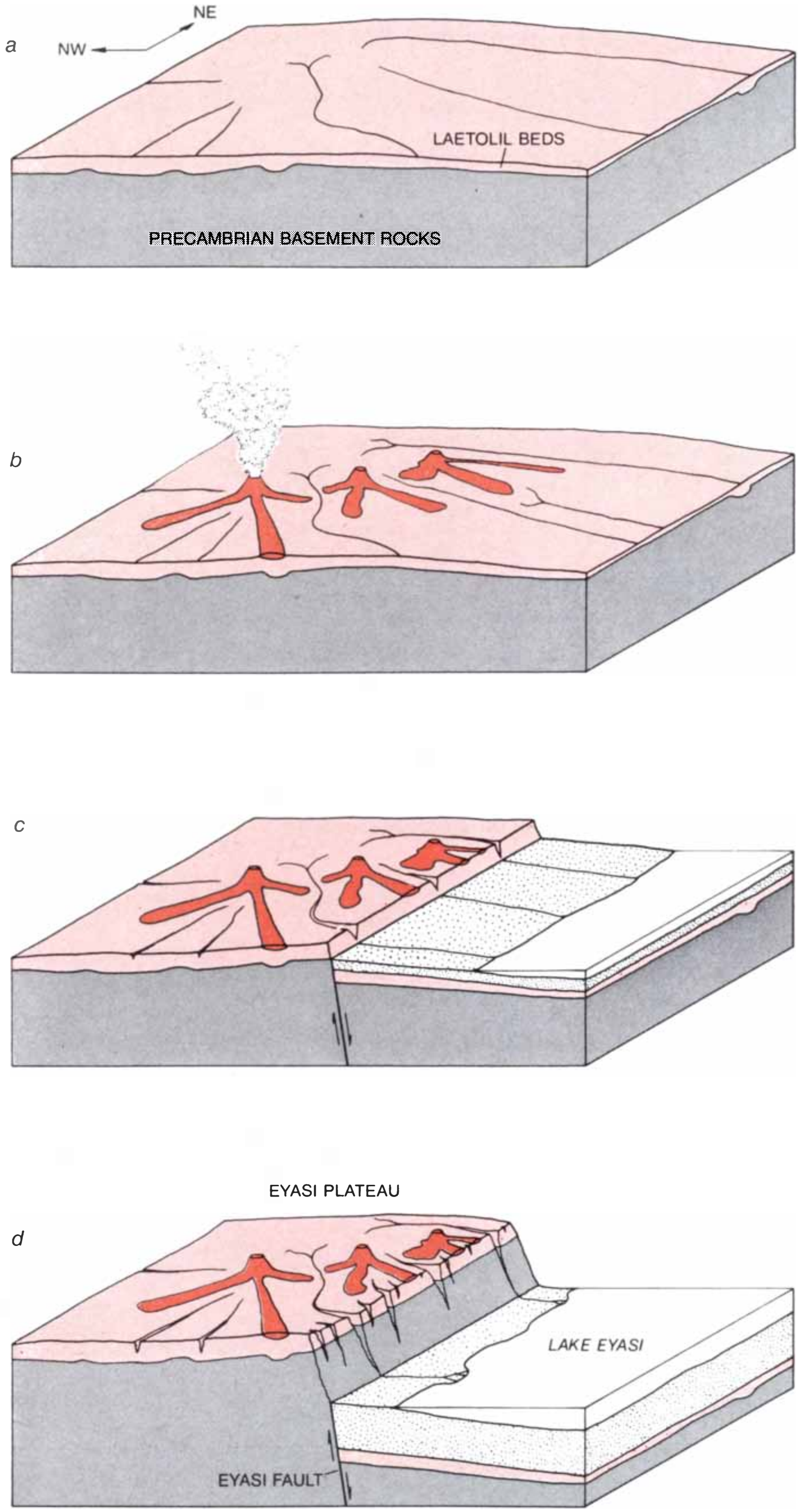
The surfaces of five layers are widely pockmarked by the impact of drops of rain; three others are only locally rain-



LAKE EYASI, southeast of Lake Victoria in Tanzania, is near the southern end of the Eastern Rift valley, which runs through Kenya and Ethiopia to the Red Sea. Early hominid remains have been found in large numbers in or near the valley. Some of the better-known hominid localities are at Hadar and Omo in Ethiopia, east of Lake Turkana in Kenya and at Olduvai Gorge and nearby Laetoli in Tanzania. Animal fossils and tracks abound in the Laetoli tuffs.



LAETOLIL BEDS, volcanic-ash deposits north of Lake Eyasi, are exposed over an area of 1,500 square kilometers (*light color*). They overlie Precambrian basement rocks (*gray*) to the west and are covered by volcanic rock (*dark color*) to the east. The beds richest in fossils are in the 70-square-kilometer Laetoli area, outlined in color. The Footprint Tuff, which lies near the top of the fossil-bearing strata, is generally more than 12 centimeters thick inside the black oval line. The volcanic ash deposited in the Laetolil Beds came from the volcano Sadiman.



FOUR BLOCK DIAGRAMS show the development of the Eyasi region. At the time when the ash of the Laetolil Beds was being deposited, between 3.8 and 3.5 million years ago (*a*), the region was relatively flat grassland savanna. Some 1.2 million years later (*b*) the region went through a period of major volcanic activity. After another 1.2 million years (*c*), when rift-valley faulting had begun, came the uplifting of the Eyasi plateau and the development of a lake below it. Further uplift and erosion (*d*) made the landscape of the Eyasi region as it is today.

printed. The rainprints are close together and well defined. They were evidently made by showers that were heavy enough to dampen the ash but not heavy enough to erode it. Footprints indent the surface of the topmost layer (No. 14) and the surfaces of seven of the 13 layers below it; they are particularly abundant and widespread on layers No. 9 and No. 14. Very few footprints are found within individual ash layers.

The upper unit of the Footprint Tuff presents a somewhat different picture. It consists of only four layers, each one representing either a single volcanic eruption or a closely spaced series of eruptions. Each of the layers has been extensively redeposited by water action: the runoff associated with heavy rainfall. The two lower layers cover surface irregularities, but together they are thicker over depressions such as hoofprints than they are over raised areas such as the rims of footprints. Most of the raised areas are truncated at the base of the third layer, which in many places is an erosional surface. Where the layers have been redeposited they are laminated, and the laminated tuff of each layer can fill the broad, shallow channels that have been scoured into the layers below it. Footprints in the upper unit are widespread on the surfaces of layers No. 1 and No. 2. Rainprints have been found at one place on the surface of layer No. 2.

As we noted above, most of the ash in the Footprint Tuff consists of particles of lava that contain the calcium-rich mineral melilite. The particles are ovoid or spheroidal in shape and mostly the size of fine to medium grains of sand (.1 to .3 millimeter). Particles of coarsely crystalline calcite, ranging in shape from rounded to flattish and generally between .02 to .01 millimeter in length, are widespread and locally abundant in the bottom layer of the lower unit (layer No. 1). Similar particles of calcite are found in a few places in layers No. 2 through No. 5.

The calcite contains relatively high concentrations of strontium and barium (a chemical "signature" of carbonatite). Textural evidence visible under the microscope, however, shows that the calcite has replaced what was originally a sodium-rich carbonatite mineral such as nyererite. The calcite particles undoubtedly represent only a small fraction of the carbonatite ash particles originally present; most of the ash would have disappeared in reacting with water (as it does, for example, in the process that leads to the formation of trona crystals).

Trona, being a water-soluble mineral, was long ago dissolved away; the Footprint Tuff layers are now cemented by calcite and, less commonly, by the mineral phillipsite (an aluminosilicate). Some layers consist largely of cement in which the ash particles appear to float. These layers are interpreted as being ash deposits where much of the space now filled with cement was originally occupied by carbonatite. The cement-rich layers are common in the lower unit of the tuff and are rare in the upper unit, which presumably received a smaller proportion of carbonatite ash.

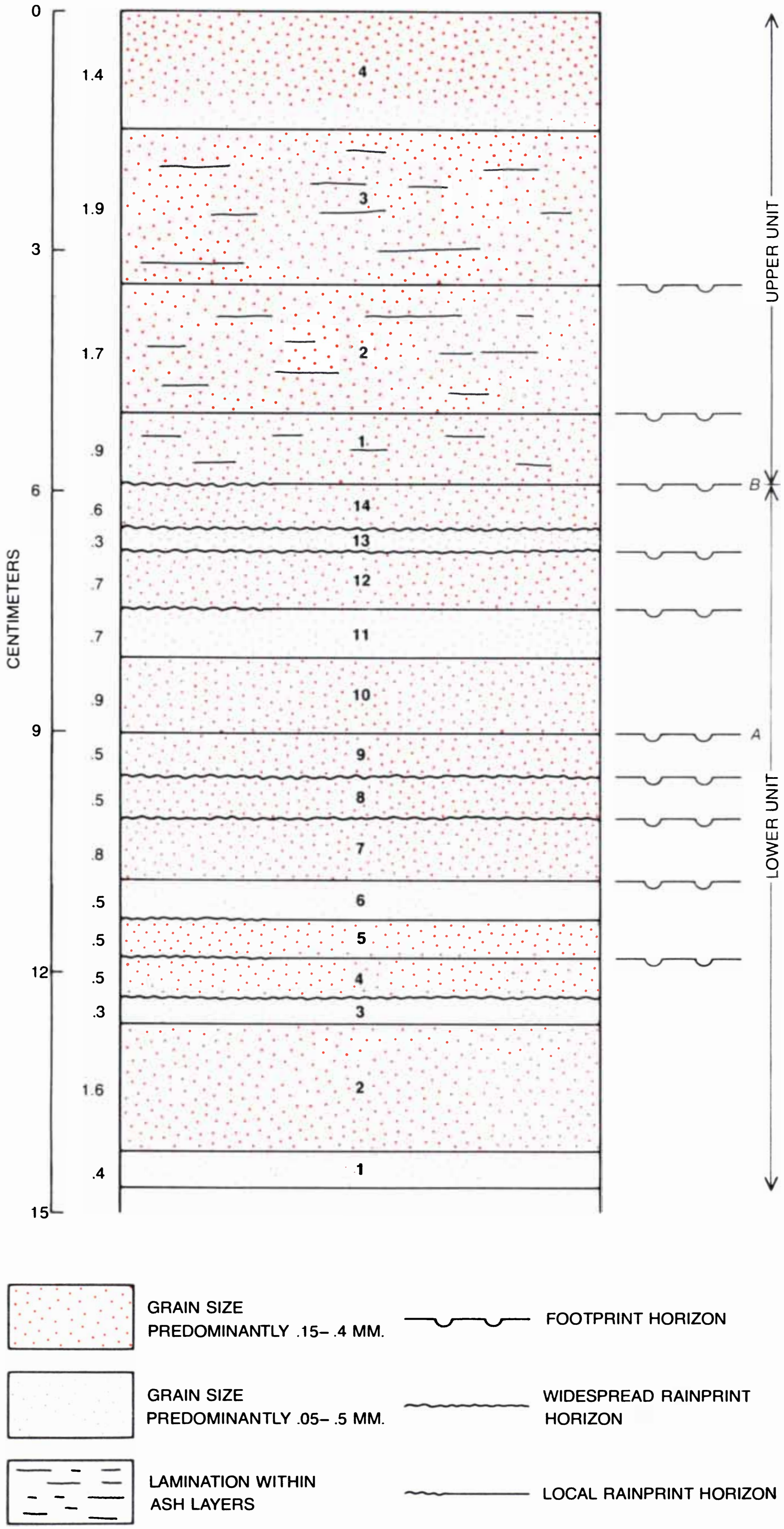
The land surface buried under the Footprint Tuff was nearly flat. Its largest irregularities were trenchlike depressions, some 10 centimeters in depth. They are interpreted as being game trails across the Pliocene savanna. The presence of grass is suggested by small fossilized rootlets under the Footprint Tuff. The ground must, however, have been essentially bare when the first layers of the Footprint Tuff were deposited; upright blades of grass would have disrupted the delicate layering of the ash. Most probably the savanna herbivores had grazed the grass down to a stubble.

The fossilized dung of hares, giraffes and the miniature antelope called the dik-dik is abundant at the base of the Footprint Tuff. Also found are the fossilized twigs and branches of trees and shrubs. A few fossilized sticks, thorns and leaves of the acacia, a typical savanna tree, have been found in overlying layers of the lower unit; one of these layers contains the dung of hares and possibly that of giraffes.

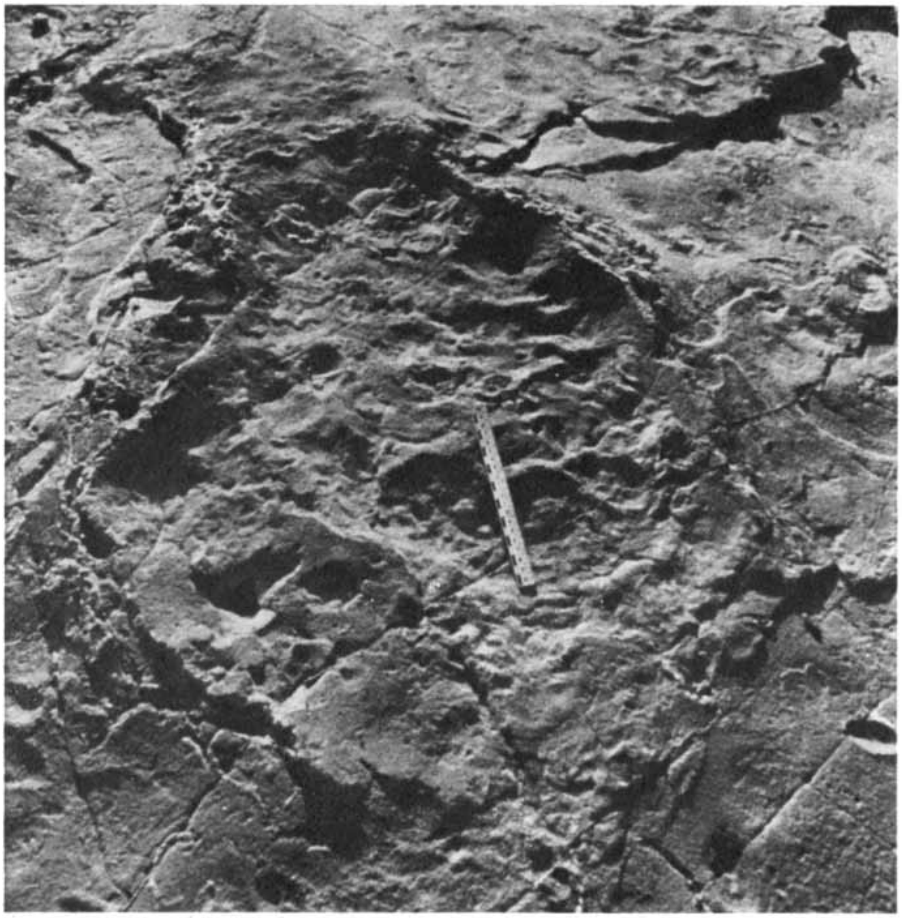
The Footprint Tuff was deposited over a short span, perhaps in a few weeks beginning near the end of the savanna dry season and extending into the early part of the rainy season. The excellent preservation of footprints and rainprints alike shows that the tuff layers must have been buried by fresh ash falls soon after the prints were made. Further evidence that the burial process was quick is provided by the continuity of layers, some only a few millimeters thick, over an area 70 square kilometers in extent.

There is other evidence of the short span of the event. The absence of grass blades at the base of the tuff suggests that the first volcanic eruptions came in the dry season. The rainprints in the lower unit were most probably made by brief showers as the dry season neared its end. In contrast, the redeposition of ash falls in the upper unit by runoff water points to a period of heavier rains, presumably the early part of the rainy season. Moreover, the widespread erosion at the base of the upper unit and between layers No. 2 and No. 3 of the upper unit is clearly attributable to heavy rainfall. Finally, the scarcity of rainprints on the surface of the upper-unit layers can be laid to the fact that rain-saturated sediments do not preserve such prints.

Since the Footprint Tuff is well ce-



FOOTPRINT TUFF is seen in section at a point where it is almost 15 centimeters thick. Of the 14 subdivisions of the lower unit eight are imprinted with tracks; the tracks are most abundant on horizons *A* and *B*. Two of the upper-unit subdivisions also bear tracks. The hominid tracks appear on horizon *B*. Numerals indicate the average thickness of the layers at Site *A*.



ELEPHANT FOOTPRINT is the large, shallow circular depression in this exposure of Upper Unit No. 1, a wet-season horizon. Additional footprints can be seen within the large one; they were probably made by small animals such as hares. The scale is 20 centimeters long.

mented and more resistant to erosion than the underlying and overlying deposits, it has survived exposure to erosion over wide areas. Where the tuff is weathered it splits readily along the bedding planes between layers, revealing any footprints on the surface of the layer exposed. The footprints have now been examined in 16 localities at Laetoli; the exposures have been labeled sites *A* through *P*.

The largest of the exposures is Site *A*, which preserves clear tracks over an area of about 800 square meters. We counted the number of individual prints in 17 square meters at Site *A* and in 20 square meters at Site *D* (which covers about 113 square meters) to arrive at an estimate of footprint density. The small oval prints of hares (and possibly of dikdiks) averaged 21 per square meter at Site *A* and 30 per square meter at Site *D*. Other prints, most of them larger, respectively averaged 1.8 and 3.6 per square meter. Taking an average of 23 footprints of all sizes per square meter, some 18,400 prints are preserved at Site *A* alone.

The footprints of animals belonging to a total of 17 separate families have been recognized. The animals make up a majority of those whose fossil remains have been found in the Laetolil Beds. The vertical distribution of the prints in the strata accords with the hypothesis that they were made during a transition from the dry season to the rainy one. Hence the prints in the lower unit are chiefly those of hares, guinea fowl, rhinoceroses and other animals that stay in the grassland savanna during the dry season. In contrast, the prints in the upper unit include those of wet-season visitors: proboscideans, horses, baboons and hominids. None of their prints are found in the lower unit. The prints of large antelopes are also much commoner in the upper unit than they are in the lower. They very probably record the wet-season migrations that are characteristic of the East African savanna. The tracks follow the preexisting game trails and indicate a normal walking gait. From this one may conclude that the ash falls from Sadiman were not heavy enough to disrupt the usual patterns of seasonal movement.

The five footprints found in 1977 that were probably made by hominids are at Site *A*. The two parallel tracks of hominid prints, first uncovered in 1978, are at Site *G*. The tracks are 25 centimeters apart and have been exposed over a distance of 25 meters. The best-defined of the footprints are from one centimeter to three centimeters deep and have clear margins. They show the rounded heel, uplifted arch and forward-pointing big toe typical of the human foot. One of the tracks was made by a single small individual. The other is a composite: the original prints were made by a comparatively large individual, and a second set of prints was superposed on the first set by a smaller individual who stepped in the original prints.

The hominid tracks are clear proof that 3.5 million years ago these East African precursors of early man walked fully upright with a bipedal human gait. This was at a time when both in stature and in brain size the hominids of Africa were still small by later human standards. Assuming that, as is true of modern human populations, the length of the Laetoli hominids' feet was about 15 percent of their height, then the smallest of the three who left their tracks at Site *G* was 1.2 meters tall (about three feet 10 inches). The next-largest would have been 1.4 meters tall (four feet seven inches). The length of the footprints made by the third and largest individual cannot be measured because they are overprinted and partly obliterated by the tracks of the next-largest.

An upright posture this early in the course of human evolution is of great importance. It freed the hands both for carrying and for toolmaking and tool use. In spite of diligent searching, no stone tools have been found in the Laetolil Beds. Hence it seems likely that the hominids who left their tracks in the Footprint Tuff had not arrived at the stage of making stone tools. The fact remains that their upright posture gave them full-time use of the first of all primate tools: unencumbered hands.

The nonhominid footprints preserved in the Footprint Tuff range from the very large to the minuscule. The large prints include those of elephants, dinotheres, giraffes and rhinoceroses. The smaller ones represent a small cat, hares, guinea fowl, francolins (birds somewhat smaller than guinea fowl) and one insect (possibly a dung beetle).

Both the making and the preservation of fossil footprints call for a rather special set of circumstances. First, the ground must be sufficiently soft and cohesive to retain clear prints. Second, the prints must be buried before destructive processes such as erosion can modify or obliterate them. Finally, the material covering them up must later separate cleanly from the surface that holds the prints. Most prints preserved in the fossil record were made in muddy sediments that became hard on drying and were then buried by sand or mud. The volcanic ash of the Footprint Tuff was quite different; it had the texture of fine- to medium-grained sand. For print preservation this is a most unusual surface.

The nature of a footprint made in unconsolidated sand depends to a great extent on how wet the sand is. Loose, dry sand will retain the tracks of small animals nicely. It will not clearly preserve the deep prints of large animals because the steep sides of the prints slump inward. Wet sand will preserve the prints of large animals but will not accept the tracks of small ones. To preserve both kinds of track the sand must be only slightly damp.

The clarity of the smaller animals' tracks in the Footprint Tuff shows that the ash was soft when the tracks were made. At the same time the well-defined vertical margins of some of the larger animals' prints show that the ash was fairly cohesive. (Some of the larger prints do have gentle inward slopes, suggestive of a minor slumping of loose ash.) Most of the larger prints also have a raised rim, squeezed upward by the pressure of the foot. The ash layers must have been firm only a few centimeters down, because the largest footprints, those of the rhinoceroses and the proboscidians, are generally no more than five centimeters deep.

It is difficult to imagine that when the various footprints were made, all 10 of the ash layers that preserve tracks could have been appropriately damp. Would that, however, have been strictly necessary? The answer is no. Just as mixing wood ash or portland cement with dry sand will make the sand more cohesive, so could the carbonatite have made the volcanic ash more cohesive by providing a fine-grained matrix for the ash particles. If the carbonatite had been fine-grained initially, a suitable primary matrix would have been supplied at the very time of deposition, and the subsequent reaction of the carbonatite with water would have provided an even finer-grained secondary matrix of precipitated calcium carbonate. Meanwhile, unless the sodium carbonate dissolved out of the carbonatite was entirely leached away by rainfall, it would crystallize into trona as the water evaporated, weakly cementing the surface layer of ash and stabilizing the footprints until they were buried by the next ash fall.

Over the ensuing thousands of years the ash layers of the Footprint Tuff were modified by a variety of processes. First the reaction of carbonatite with water produced an abundance of calcite that in the early stages would have weakly cemented the ash layers and would also have replaced the dung and the plant material in the ash. At this stage insects and other animals burrowed in the ash. As the tuffs were weathered, ash and melilite were transformed into montmorillonite clay, and the aluminosilicate mineral phillipsite was deposited locally as an additional cementing agent. Thereafter calcite filled nearly all the remaining pore space, leaving the ash layers a well-cemented hard rock. Calcite was also deposited between some ash layers in the form of thin, discontinuous sheets of limestone. Such limestone coated many of the hominid footprints at Site *G*, making it easier to excavate the prints without altering them.

To summarize, a most unusual set of conditions gave rise to the Footprint Tuff. Approximately 3.5 million years ago the Laetoli area supported an abundant and diverse animal population. The showers of volcanic ash that fell in the area were not heavy enough to drive the animals away. The excellent definition of a multitude of footprints, tracks and trails made in the fresh ash was a result of the admixture of carbonatite, an uncommon igneous material, with the more typical particles of volcanic ash; the imprinted ash layers were then buried at frequent intervals by fresh falls. Finally a heavy ash fall deeply buried the Footprint Tuff, protecting it from erosion, and the imprinted layers were cemented and became hard rock. When recent weathering split the layers of tuff along their bedding planes, it opened a unique window on the world of early hominids and the animals with which they lived.

THREE KINDS OF ANIMAL made the imprints seen in this part of the exposed tuff. Tracks of a bird appear at the lower left; those of a small animal, possibly a hare or a dik-dik, appear at the upper left and right. To the left of center is a print of a baboon's hind foot.

How an Animal Virus Gets into and out of Its Host Cell

Experiments with the Semliki Forest virus show in considerable detail how the cell is caused to manufacture new virus particles, including an outer membrane that is a piece of the cell's own

by Kai Simons, Henrik Garoff and Ari Helenius

For an animal or a plant the cell is the fundamental unit of structure and function. For a virus the cell is merely a means of making new virus particles. A virus particle consists of one or more strands of nucleic acid (DNA or RNA) enclosed in a protein shell called a capsid. In many viruses the capsid itself is enclosed in a membrane with protein molecules embedded in it, much like the outer membrane of an animal cell. The nucleic acid within the particle constitutes the genetic information needed for the virus to replicate itself. The virus particle, however, lacks the apparatus to transform this information into its own finished structure. That apparatus is supplied by the host cell. The virus particle enters the cell and expropriates the biosynthetic machinery to make its parts. When enough parts have been made, they are assembled into new virus particles that leave the cell to infect other cells.

The general nature of virus infection has been known for some years, but the precise molecular mechanisms by which the virus enters and leaves the cell are still being worked out. Although not all of the infective process is understood in detail, it has become strikingly clear that very few of the reactions of the viral life cycle are specific to the virus itself. For the most part the virus exploits the normal metabolic reactions of its host. In our study of the Semliki Forest virus, an RNA virus that infects animals, we have found that the virus enters the cell on the same pathway as many molecules required in the cell's metabolism.

After proteins of the outer viral membrane have become attached to the cell membrane, the cell membrane pinches off around the virus particle to form a vesicle. Inside the cell this vesicle merges with other vesicles; ultimately the virus particle is deposited in the large vacuole called a lysosome. (The acidic interior of the lysosome contains enzymes that can serve the function of degrading large molecules such as proteins and nucleic acids.) Only then does a reaction unique to the virus occur: the outer viral membrane fuses with that of the lysosome, releasing the capsid and its content of RNA into the cytoplasm of the cell and allowing them to escape the degradative enzymes of the lysosome. Then the process of replication of the virus particle can begin.

The duplication of the viral RNA and the manufacture of the capsid proteins and membrane proteins are accomplished almost entirely by the normal metabolic processes of the cell. The new RNA and capsid proteins are joined in the cytoplasm to form a new virus core: a nucleocapsid. The proteins destined for the outer membrane of the virus are synthesized, modified, transported and inserted in the cell membrane just as the cell's own membrane proteins are. The newly made nucleocapsid then moves toward the cell membrane; the capsid proteins bind to the viral proteins in the cell membrane, avoiding the host's own membrane proteins. As this binding proceeds the cell membrane curves around the nucleocapsid. When each molecule of capsid protein is attached to viral proteins in the cell membrane, the virus particle emerges from the cell. Semliki Forest virus is therefore a poor guest indeed: entering the host cell by means of the cell's normal ingestive processes, it expropriates the cell's machinery to make its own parts and leaves wrapped in a piece of the cell's membrane.

The Semliki Forest virus, abbreviated SFV, was first isolated in 1944 from the tissues of the mosquito and was named for the rain forest in southern Uganda where the mosquitoes were found. Viruses are generally grouped in three categories: those that infect animal cells, those that infect plant cells and those that infect bacteria. Because of the differences in the composition of the host cells the means of entry and exit of the three types of viruses are somewhat different. The Semliki Forest virus is an animal virus that infects a wide range of vertebrates and invertebrates. It is closely related to the viruses that cause yellow fever and dengue. Semliki Forest virus can cause encephalitis in mice but is not generally pathogenic for man.

The Semliki Forest virus is one of a class of viruses known as togaviruses. Togaviruses are small and spherical. They are also relatively simple in structure. The name of the group stems from the fact that its members all have a membrane—a "toga"—around the capsid. This is in contrast to those viruses in which the outer surface of the virus particle is formed by the capsid. At the center of the SFV particle is a single-strand molecule of RNA, a chain of 12,700 nucleotides. (Each of the nucleotides that make up an RNA molecule consists of a sugar molecule, a phosphate group and one of four nitrogenous bases.)

The relative simplicity of the Semliki Forest virus is manifested in the fact that each virus particle incorporates only four kinds of protein molecule. The capsid consists of 180 molecules of what is called the capsid protein, or *C* protein. Each *C*-protein molecule is a chain of 267 amino acids. The underlying structure of the capsid is that of an icosahedron. An icosahedron is a polygon with 20 faces each of which is an equilateral triangle. In the capsid each face of the icosahedron is divided into three smaller triangles; the surface of the capsid thus has 60 faces. A molecule of *C* protein is positioned at each corner of all 60 faces, yielding a total of 180 *C*-protein molecules.

The three other proteins of the Semliki Forest virus are joined to form spike-shaped protrusions that extend through the outer membrane of the virus. Biochemical studies have shown that each spike is composed of one molecule of each of the three proteins. The proteins have been designated *E*1 (438 amino acids long), $E2$ (422 amino acids long) and $E3$ (66 amino acids long). Carbohydrate side chains extend from each of the three proteins in the spike; proteins with such carbohydrate side chains are called glycoproteins.

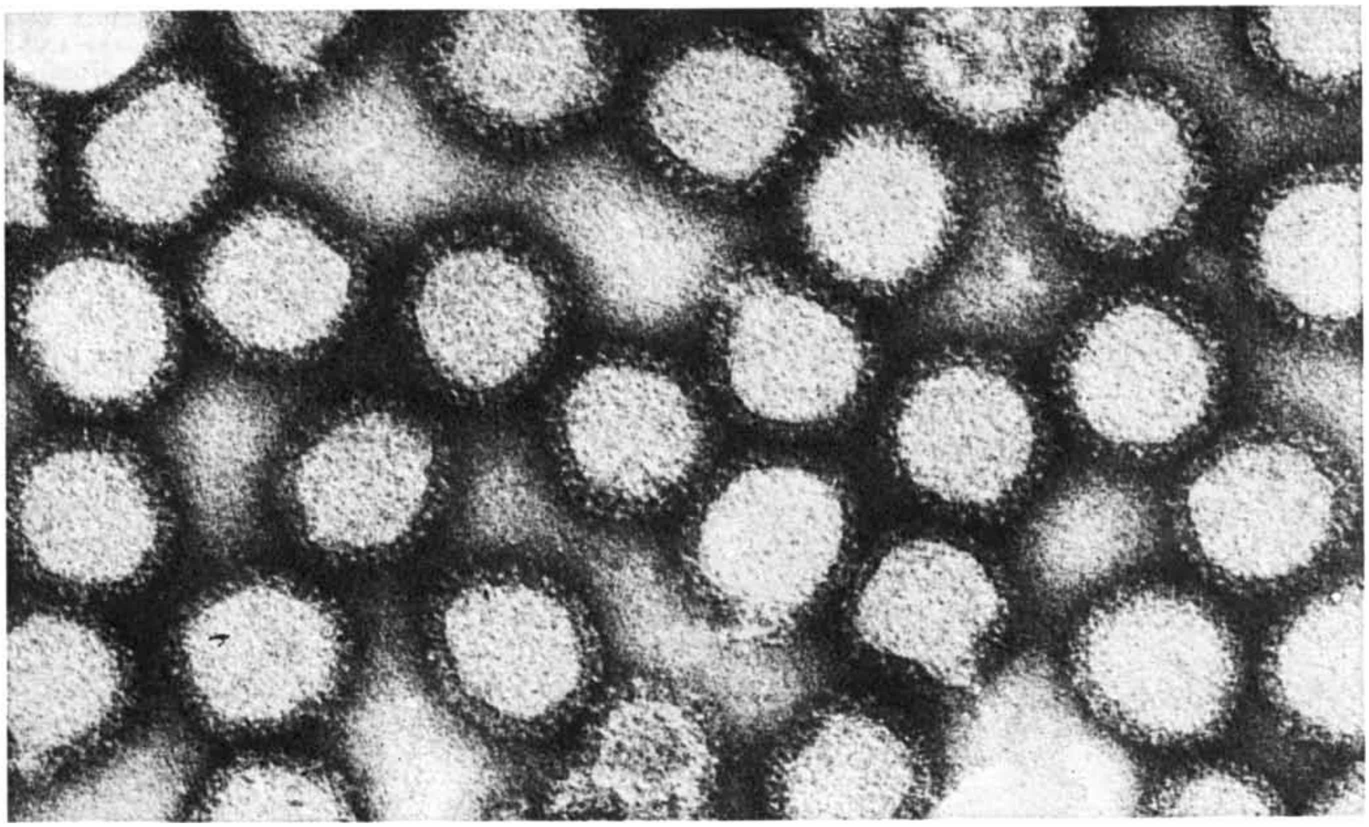
The glycoproteins of the SFV membrane spikes are twined about one another in a complicated way. As a result all of the $E3$ protein and about 90 percent of the $E1$ and $E2$ proteins are outside the outer surface of the viral membrane, forming the part of the spike visible in electron micrographs. Short segments of the $E1$ and $E2$ molecules pass through the membrane. The segment of the $E1$ molecule inside the membrane is two amino acids long and the segment of the $E2$ molecule 31 amino acids long. The short inner segments anchor the spike to the capsid.

The order of the amino acids in the spike proteins has been found by determining the sequence of nucleotides in the viral RNA. (Each amino acid is coded for by a sequence of three nucleotides.) It has proved difficult to determine the complete amino acid sequence of proteins that span membranes by analyzing the proteins themselves. The difficulty arises from the fact that it is rarely possible to obtain the protein unmixed with other substances in the quantity required for determining the sequence of amino acids. Methods of rapidly finding the order of nucleotides in an RNA or DNA molecule, however, have in recent years made the indirect approach to discovering the order of amino acids relatively straightforward. In collaboration with Anne-Marie Frischauf and Hans Lehrach of the European Molecular Biology Laboratory (EMBL) we have resorted to this method to analyze the proteins of the SFV membrane spikes.

Our work has shown that both the $E1$ and the $E2$ proteins have an unusual segment. The segment is made up predominantly of hydrophobic amino acids (amino acids that are repelled by water). An uninterrupted stretch of such amino acids is rarely found in the water-soluble proteins of the cell's cytoplasm. Such segments do, however, occur in membrane proteins other than those of the Semliki Forest virus. The order of the amino acids in a membrane glycoprotein that includes a hydrophobic segment was first determined by Jere P. Segrest, Vincent T. Marchesi and their colleagues at the National Institute of Arthritis, Metabolism, and Digestive Diseases, who used the direct method. Other membrane proteins (including immunoglobulin proteins that are bound to membranes, transplantation-antigen proteins and the spike proteins of other viruses) were subsequently analyzed by the indirect method.

All these proteins share important structural properties with the SFV spike proteins: when they are in the cell membrane, most of their mass remains outside the cell. The connection with the membrane is maintained by a segment consisting of between 20 and 30 hydrophobic amino acids that extends through the membrane; a small stub, consisting of a few amino acids, protrudes beyond the membrane into the cytoplasm. The close resemblance between the SFV spike proteins and the host cell's own membrane proteins will become even clearer when we describe the synthesis of the spike proteins.

In its protein molecules the Semliki Forest virus is quite exclusive; no host-cell proteins are found in the virus particle. In contrast the lipid molecules of the SFV membrane come from the host cell. The SFV membrane has the same form as the membrane of an animal cell: two layers of lipid molecules with a variety of inserted glycoproteins. (Each of the



SEMLIKI FOREST VIRUS PARTICLES are the round light areas in this electron micrograph made by Carl-Henrik von Bonsdorff of the University of Helsinki; the particles are enlarged 200,000 diameters. The Semliki Forest virus (SFV) is named for the rain forest in Uganda where it was first isolated. It infects many vertebrate and invertebrate animals. Related viruses cause encephalitis, but the Semliki Forest virus is not pathogenic for man. Like other viruses, the Semliki Forest virus lacks the apparatus needed to reproduce itself. Only by entering an animal cell and expropriating the cell's machinery of biosynthesis can the virus make new particles. Each particle consists of a single-strand molecule of RNA in a protein shell called a capsid. The capsid is covered by a membrane similar to that enclosing an animal cell. Embedded in the viral membrane are protein molecules, which appear as a light layer of fuzz around the particles.

lipid molecules has a hydrophilic "head" and two hydrophobic "tails" made up of extended fatty acid chains. In the membrane the hydrophilic heads face outward and the tails face inward toward one another.) The lipid molecules in the membrane are free to move laterally but not vertically. Their unconstrained movement in two dimensions endows the membrane with the properties of a liquid. The property is essential for many of the functions of the cell membrane, including the process that leads to the exit of the virus from the host cell.

Lipid molecules exist in a number of forms that differ according to the chemical groups included in the molecule. The differences among lipids have made it possible to identify the origin of the lipid molecules in the SFV membrane. Ossi Renkonen and Leeri Kääriäinen and their colleagues at the University of Helsinki have shown that the lipid composition of the membrane of the SFV particle depends directly on the host cell in which the virus particle has been assembled.

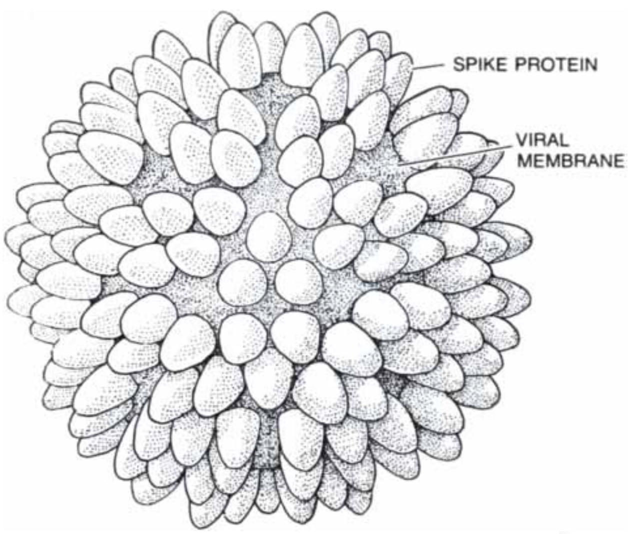
How does the Semliki Forest virus enter an animal cell? In our attempt to answer this question we have for the most part worked with hamster kidney cells cultured in the laboratory. Experimental systems based on cultured cells cannot entirely explain the process of virus infection in the intact organism. They can, however, help to clarify how a virus gets into and out of a cell. Although our knowledge of the later stages of virus infection in cultured cells was quite detailed, the means by which the genetic material of the virus gains entry to the cytoplasm of the host cell had remained obscure. An active mechanism must be in operation, because large molecules such as proteins and larger structures such as viruses do not simply diffuse through the cell membrane.

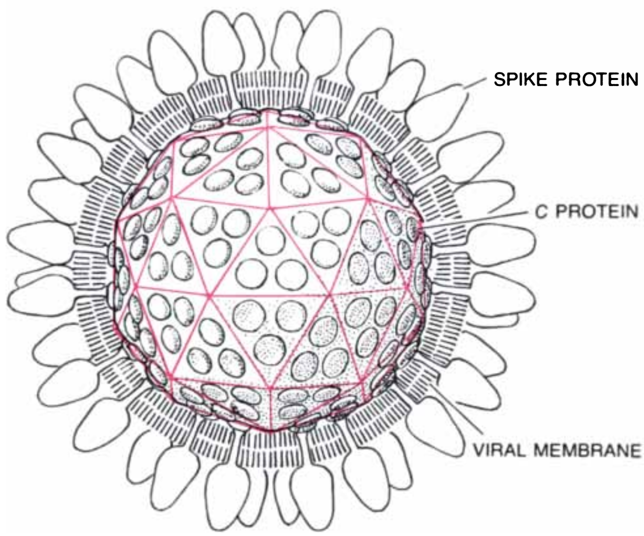
Two mechanisms have been proposed to explain how those viruses that have outer membranes get into animal cells. Entry might be gained by the fusion of the viral membrane with that of the cell. If the viral membrane is assimilated into the cell membrane, the nucleocapsid could be released into the cytoplasm. Such fusion has been observed among some viruses, but it has remained unclear whether the process could lead to infection. An alternative mechanism could be that the cell membrane forms a vesicle around the virus particle, whereupon the vesicle could detach itself from the membrane and carry the particle into the cytoplasm. Viruses are known to enter cells by this process, which is called endocytosis; it had not been shown, however, that endocytosis could lead to infection. Actually the route of entry of the Semliki Forest virus into the animal cell, which we discovered in collaboration with Erik Fries, Mark Marsh and Judith White of the EMBL and Jürgen Kartenbeck of the German Cancer Research Center, has features of both mechanisms.

Our work has also shown that the spike proteins and the viral membrane do not merely provide a protective coat for the nucleocapsid; they play a crucial role in the delivery of the particle into the cytoplasm. The first stage in entry is the attachment of the spike glycoproteins to the cell surface. Our results indicate that the SFV spike glycoproteins bind to glycoproteins in the cell membrane, including transplantation antigens (immune-system proteins responsible for recognizing foreign tissues).

On exposure to an animal cell the SFV particles bind preferentially to the surface of microvilli: elongated projections extending from the cell surface. After binding to a microvillus the virus particle moves down its shaft toward the center of the cell. Within a few minutes this motion takes the particle into a "coated pit," a specialized region of the cell membrane often seen near the base of a microvillus. On the cytoplasmic (inner) side of the coated-pit region of the membrane is the thick coat for which the region is named. The coat consists mainly of clathrin, a protein of high molecular weight that was first isolated by Barbara M. F. Pearse of the Medical Research Council Laboratory of Molecular Biology in Cambridge, England.

Coated pits are formed continually on the surface of the cell. After forming they fold inward and pinch off into the cytoplasm, forming a "coated vesicle." The clathrin that had been on the inner surface of the membrane now surrounds the vesicle, constituting a polygonal shell similar in form to the SFV capsid. As the coated vesicle moves into the cytoplasm it apparently loses its clathrin and fuses with an endosome, a large vacuole with a smooth outer surface. The endosome itself then fuses with a lysosome, the degradative vacuole mentioned above. The main function of this organelle of the cell is to degrade unneeded substances native to the cell and those that have entered the cell by endocytosis but must be broken down before being utilized. This action is ac-





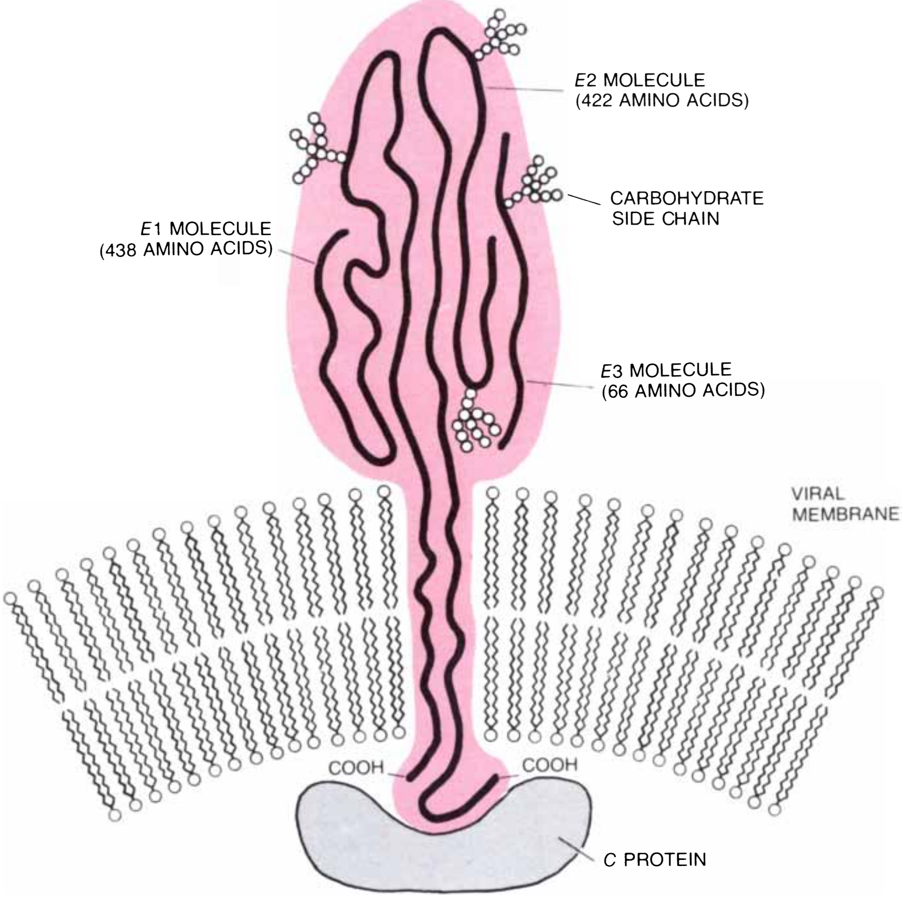
STRUCTURE OF THE SFV PARTICLE is relatively simple. The particle is 650 angstrom units in diameter. The RNA inside the capsid (not visible) is a chain of 12,700 nucleotides; the order of the nucleotides provides the genetic information needed for making the viral proteins. The capsid consists of 180 molecules of a protein designated the *C* protein. (Together the RNA and the capsid form a nucleocapsid.) The capsid has the form of a regular polygon with 60 faces; each face is an equilateral triangle. Three molecules of *C* protein are positioned in each triangle. The outer viral membrane is composed of two layers of lipid molecules, with their hydrophobic "heads" facing outward and their hydrophilic "tails" facing inward. The lipid molecules are free to move laterally, giving the membrane the properties of a liquid. Inserted into the membrane are 180 "spikes," each made up of three linked protein molecules. Each spike spans the membrane and is bound to a single *C*-protein molecule of the capsid. The protruding spikes give the SFV particle its studded appearance (*left*).

complished by a variety of digestive enzymes; it is aided by the fact that the *p*H in the lysosome is lower than that in the other regions of the cell, indicating a greater acidity.

In entering the cell the SFV particles are exploiting part of an important cellular cycle. Many molecules that are needed for cell regulation or nutrition enter the cell by endocytosis. The molecules that gain entry in this way include lipoproteins, polypeptide hormones, vitamins (as part of larger molecules) and growth factors. Many of these molecules end up in the lysosome. The most closely studied of those that do is serum low-density lipoprotein, the main carrier of cholesterol in the blood. Richard G. W. Anderson, Michael S. Brown and Joseph L. Goldstein of the University of Texas Health Science Center at Dallas have shown that the lipoprotein enters the cell by means of the coated pit and is degraded in the lysosome; the cholesterol is subsequently utilized in the assembly of new cellular membranes.

The capacity of animal cells to take up virus particles by endocytosis is large: a hamster kidney cell can take up some 3,000 particles per minute. An average of 1.3 virus particles is seen in each vesicle that contains a particle; therefore at least 2,300 coated vesicles must be formed at the surface of the cell every minute. The capacity of this system to transport substances into the cell is so great that the cell needs some way to move the membrane components that make up the coated vesicle back to the cell surface; otherwise the membrane would be depleted. It has been shown that endocytosis and the return of the membrane components to the surface form a cycle that is in continuous operation; in entering the cell the virus simply exploits the first leg of the cycle, which terminates in the lysosome.

How does the SFV particle escape the destructive action of the enzymes in the lysosome? The fact that animal viruses entering the cell by means of endocytosis are transferred to the lysosome has been known for some time. It had been thought, however, that the process led to the destruction of the virus particle rather than to infection. The first evidence to the contrary came from experiments employing lysosomal inhibitors. The inhibitors used were substances that are weakly basic and soluble in lipids when they are not electrically charged. In their uncharged form the inhibitors can pass through the outer membrane of the lysosome. Inside the lysosome the molecules acquire an electric charge as a result of the acidity of the interior of the vacuole. Since the inhibitor molecules are now charged, however, they cannot pass back out through the membrane of the lysosome. Therefore they accumulate inside the lysosome, decreasing its acidity and inhibiting its function.



SFV MEMBRANE SPIKE is composed of three protein molecules twisted about one another. The molecule designated *E*3, a chain of 66 amino acids, lies entirely outside the viral membrane. Most of the *E*1 molecule (438 amino acids long) and the *E*2 molecule (422 amino acids long) also lie outside the membrane. Short segments of the *E*1 and *E*2 proteins span the viral membrane. The short stub on the inner side of the viral membrane is made up of 31 amino acids of the *E*2 molecule and two amino acids of the *E*1 molecule. The stub anchors the spike to a *C*-protein molecule of the capsid. Each of the three spike proteins has carbohydrate side chains attached; such proteins are called glycoproteins. In structure and manufacture the glycoproteins of the SFV membrane are very similar to those embedded in the animal-cell membrane.

We have found that the application of such lysosomal inhibitors to animal cells interrupts the process by which the nucleocapsid is liberated into the cytoplasm of the host cell. The inhibitors do not affect the binding of the SFV particle to the cell membrane or the particle's movement into the lysosome. They do, however, keep the nucleocapsid from passing out of the lysosome into the cytoplasm. Further work in our laboratory has shown that the inhibitors interfere with the fusion of the viral and lysosomal membranes. In the absence of the inhibitors the fusion of the membranes allows the particle to pass through the lysosome and into the cytoplasm.

The fusion reaction apparently is triggered by the acidity of the interior of the lysosome. When the *p*H of a preparation containing SFV particles and pieces of free lipid membrane was lowered to a level similar to that in the lysosome, the SFV membrane fused instantaneously with the other membranes, allowing the nucleocapsid to pass through them. In an extension of this work synthetic vesicles were made out of lipids like those of the membrane of the lysosome. Ribonuclease, an enzyme that breaks down RNA molecules, was then introduced into the vesicles. When the vesicles were mixed with virus particles and the *p*H of the solution was lowered, the viral membrane fused with that of the vesicle. Electron micrographs revealed that the vesicles were covered with viral spike proteins; the nucleocapsids were inside the vesicles and the viral RNA was being broken down by the ribonuclease.

In the intact cell the virus particle moves out of the lysosome by this means. The acidity of the interior of the lysosome apparently induces a change in the viral membrane that enables it to fuse with the lysosomal membrane. The process takes place so quickly that the nucleocapsid is expelled into the cytoplasm of the cell without being destroyed by the degradative enzymes of the lysosome.

The exact mechanism by which the fusion is induced is not known. We suspect, however, that the *E*1 protein of the viral-membrane spike plays an important role. The *E*1 molecule has a hydrophobic segment made up of 16 amino acids in the part of the protein that lies outside the viral membrane. If the surrounding medium is acidic, the hydrophobic segment may become exposed and react with the membrane of the lysosome, bringing the membranes into close apposition and allowing the fusion reaction to occur.

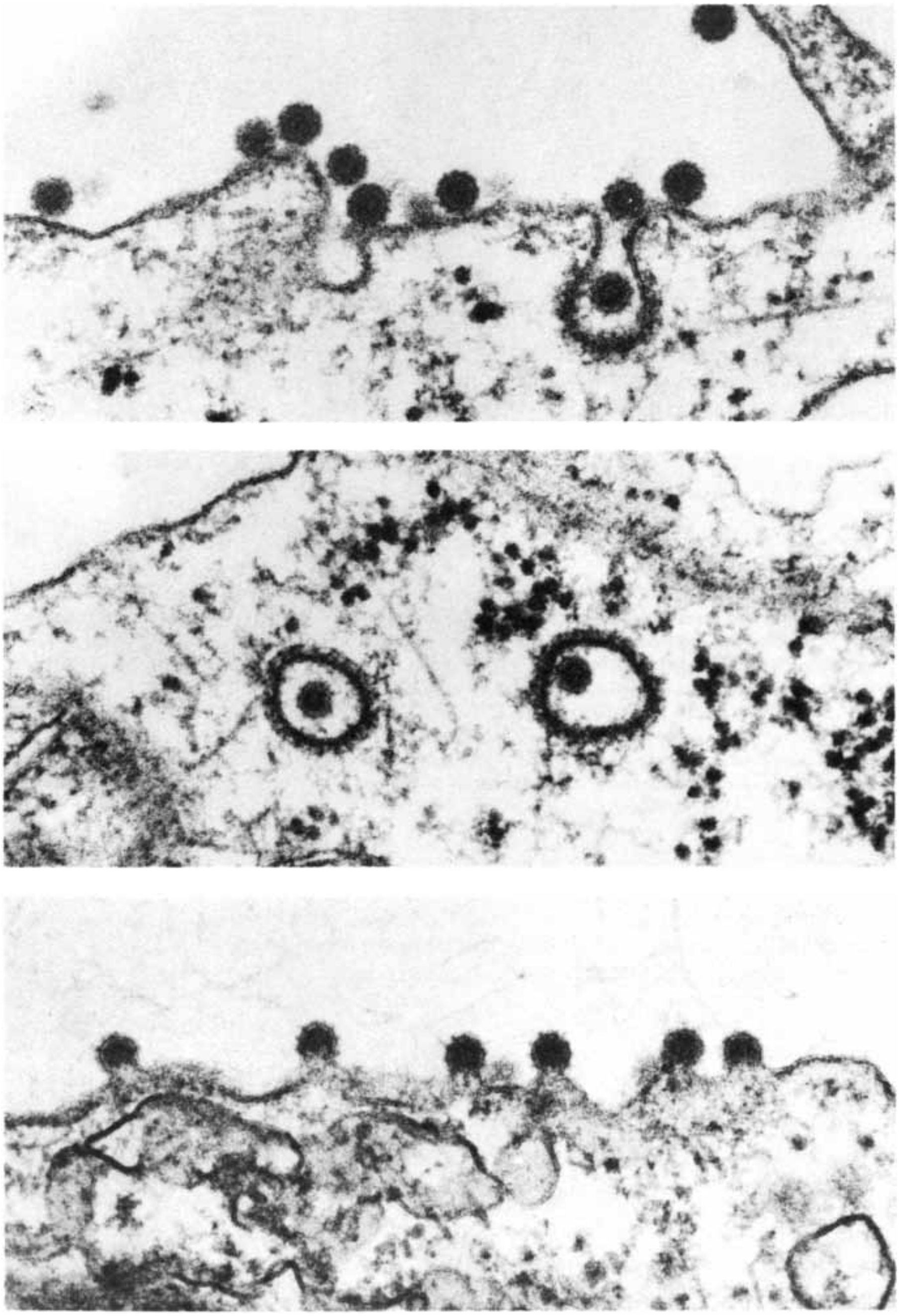
The work with the Semliki Forest virus marks the first time the route of a virus into an animal cell has been traced up to the point at which the nucleocapsid is released into the cytoplasm. The significance of the finding will be enhanced if this mechanism is found to be general among animal viruses. Recent work we have done with Karl Matlin and Hubert Reggio of the EMBL indicates that influenza viruses and the virus of the disease vesicular stomatitis (which is related to the virus that causes rabies) do follow this route. Both viruses, like the Semliki Forest virus, have an outer membrane with proteins inserted into it. Some bacterial toxins may also enter the cytoplasm of the cell by way of the lysosome.

The lysosome may indeed prove to be a point at which the animal cell is vulnerable to many pathogens. Investigation of the relation between viruses and lysosomes may therefore have important clinical implications. The drugs currently available for treating virus infections are of limited effectiveness and often have undesirable side effects. A drug that interfered with the fusion of the viral membrane and the lysosomal membrane might well be an effective antiviral agent.

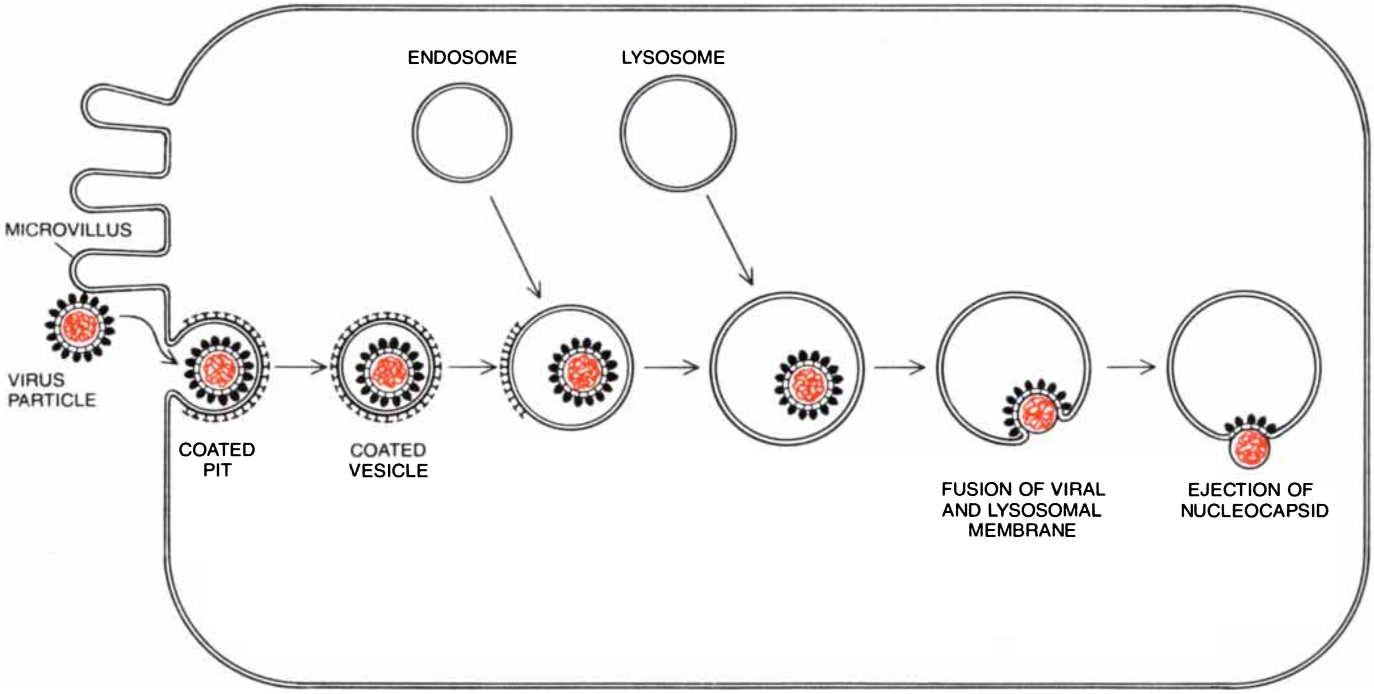
The release of the nucleocapsid from the lysosome completes the first phase of infection. The second phase includes the replication of the viral RNA, the manufacture and modification of the viral proteins and the assembly of the new nucleocapsid. These processes depend heavily on the machinery of protein synthesis contained in the host cell: the ribosomes (the small spherical organelles where proteins are synthesized) and a variety of enzymes and other molecules. In order for the initial replication of the viral RNA to begin the nucleic acid must be freed from the capsid. When that has been accomplished, several proteins are translated from the first segment of the viral RNA. One of them is an enzyme called RNA-dependent RNA polymerase. RNA polymerase catalyzes the duplication of RNA. It must be coded for by the genetic matter of the virus, because the enzyme does not exist in the host cell. Neither the release of the nucleic acid nor the making of the enzyme is as yet fully understood.

When the liberated nucleic acid and polymerase molecules are jointly present in the cytoplasm, replication begins. Two different kinds of RNA molecule are manufactured in large amounts. One is a copy of the complete viral nucleic acid molecule, 12,700 nucleotides long. It is designated the 42*S* RNA. (*S* is an abbreviation for Svedberg unit, a measure of the mass of the molecule.) The accumulating 42*S* molecules serve as the basis for making more RNA polymerase. Ultimately they are incorporated into newly made nucleocapsids.

The second kind of RNA is a copy of the last third of the whole viral-RNA molecule, beginning after the segment that codes for RNA polymerase. The shorter molecule, 4,170 nucleotides long, is the 26*S* RNA. It incorporates the genetic information needed for the synthesis of the four viral proteins. The coding sequences for the four proteins are arranged on the 26*S* strand in the order *C*, *E*3, *E*2 and *E*1. Within four



FIRST AND LAST STAGES OF INFECTION of an animal cell by the Semliki Forest virus are shown in electron micrographs made in the laboratory of the authors and that of Jürgen Kartenbeck. The first stage in entry is the attachment of the particle to the cell membrane (*top micrograph*). Proteins of the viral spike attach themselves to protein receptors in the cell membrane. The receptors are protein molecules whose configuration enables them to bind to a limited number of other proteins. The virus particle tends to bind to a microvillus, a fingerlike projection of the cell surface (*at right in top micrograph*). The particle then moves down into a "coated pit," named for the thick layer of protein on its inner surface. The coated pit pinches off around the particle to form a "coated vesicle"; the vesicle carries the particle into the cytoplasm of the cell (*middle micrograph*). When the virus particle has replicated itself, the new particles leave the host cell by budding outward from the cell membrane (*bottom micrograph*).



ENTRY OF SEMLIKI FOREST VIRUS into the animal cell requires the virus particle to pass through several vacuoles. The coated vesicle is covered with the protein called clathrin, which forms a shell around the particle similar in structure to the capsid around the viral RNA. Inside the cell the vesicle loses its clathrin coat and fuses with a large, smooth-surfaced vacuole called an endosome. Many molecules involved in normal cellular metabolism are taken into the cell by this process, which is termed endocytosis. The endosome merges with the vacuole called a lysosome. The lysosome contains degradative enzymes that break down substances that have been taken into the cell; many of the substances are thereby converted into usable form. The interior of the lysosome is more acidic than other regions of the cell are. It has long been known that virus particles can enter the cell this way. It was assumed, however, that the particles were broken down by the enzymes of the lysosome and could not reach the cytoplasm to initiate infection. It is now known that the acidity of the lysosome causes the outer viral membrane to fuse with the membrane of the lysosome. As a result of this process the nucleocapsid is freed into the cytoplasm. This is the first time that the path of a virus into an animal cell has been followed to the point where infection begins. The finding may prove to be highly significant if it can be demonstrated that other animal viruses initiate infection in the same way.

hours after the virus has been added to an animal cell the two forms of viral RNA comprise most of the messenger-RNA molecules in the cell. The cell has been converted into an assembly line for the manufacture of new virus particles.

The synthesis of the viral capsid and spike proteins begins when the 26*S* RNA is attached to a ribosome. The four structural proteins are translated consecutively from the 26*S* RNA. Translation begins at the "start" codon, a triplet of nucleotides preceding the sequence that codes for the *C* protein. The process continues without interruption until the "stop" codon at the end of the segment that codes for the *E*1 protein has been reached. The segment of the viral nucleic acid from the start codon to the stop codon is composed of 3,760 nucleotides.

In the living cell the four viral proteins are not simply released from the ribosome into the cytoplasm as a single chain. The new *C* protein, the first to be translated, is found in the cytoplasm. The *E*1 protein and an intermediate protein known as *p*62, consisting of the *E*2 and *E*3 proteins linked together, are found inserted into the membrane of the endoplasmic reticulum: a network of interconnected membranes that plays the central role in the synthesis and modification of host-cell proteins, including membrane proteins and secretory proteins (proteins the cell secretes). The manufacture of the viral membrane proteins is thus almost identical with that of the host-cell membrane proteins.

A good deal is known about the structure of glycoproteins in the membrane of animal cells and about how those molecules are built. Many of the cell glycoproteins have a distinctive group of between 20 and 30 amino acids at the end of the protein chain known as the amino terminal end; the amino terminal end is the first part of the chain to be made on the ribosome. This segment, the "signal sequence," has a special affinity for a site on the surface of the membrane of the endoplasmic reticulum. As soon as the 20 to 30 amino acids have been translated by a ribosome moving freely in the cytoplasm, the ribosome attaches itself to the endoplasmic reticulum. The bound ribosome proceeds with the translation of messenger RNA: the RNA that conveys genetic information from the cell's DNA to its ribosomes. The growing chain of amino acids does not, however, extend into the cytoplasm. It is now threaded through the membrane of the endoplasmic reticulum and extends into the interior of the reticulum.

With an assay developed by Günter Blobel and Bernhard Dobberstein of Rockefeller University we have been able to show that the same process is involved in the synthesis of the membrane proteins of the Semliki Forest virus. We worked with a cell-free system including ribosomes, enzymes, transfer RNA's, amino acids and the other substances needed for protein synthesis. Without endoplasmic reticulum the system was unable to manufacture the viral spike proteins correctly. When vesicles of endoplasmic reticulum were added, however, the spike proteins were synthesized just as they are in a cell that has been infected by the virus. Three proteins were made: *C*, *E*1 and *p*62. Moreover, the amino acid chains were in the right places. The *C* protein was outside the vesicles, in the region corresponding to cytoplasm in the intact cell. The *E*1 and *p*62 chains, however, were inserted into the membrane of the vesicle.

As a result of these experiments and others it is now known that immediately after the *C* protein is translated in the intact cell it is cleaved from the growing chain by a protease, an enzyme that cuts proteins. The cleavage of the *C* protein exposes a signal sequence at the end of the *p*62 chain. The signal sequence binds the ribosome to the endoplasmic reticulum and thus initiates the insertion of the protein chain through the membrane. Almost all of the *p*62 chain is then translated from the viral nucleic acid and passes into the interior of the endoplasmic reticulum. The movement across the membrane is apparently arrested by the group of hydrophobic amino acids close to the tail of the chain (near the end of the *E*2 molecule). This group cannot pass through the hydrophobic lipid membrane, and it anchors the protein chain to the endoplasmic reticulum. When the translation of *p*62 is complete, the chain is fixed to the membrane, with 31 amino acids extending into the cytoplasm and the rest of the chain in the interior of the endoplasmic reticulum.

The *E*1 protein is synthesized next. It apparently has its own signal sequence, coded for by a segment of the viral RNA between the stretches that correspond to the *E*1 and *E*2 proteins. The signal sequence, 60 amino acids long, is designated 6*K*. It is removed from the growing *E*1 chain by two enzymatically catalyzed cuts after the chain has been bound to the membrane of the endoplasmic reticulum. The passage of the *E*1 protein across the membrane, like that of the *p*62 protein, is arrested by a sequence of hydrophobic amino acids near the end of the chain. The chain remains fixed to the endoplasmic reticulum, with a small segment extending into the cytoplasm. The translation of the *E*1 protein, however, is less fully understood than the other processes in the synthesis of the viral proteins.

The carbohydrate side chains of the viral glycoproteins are attached in the course of translation; the side chains are later chemically modified. The same process operates in the synthesis of some of the cell's own glycoproteins. In fact, the translation of the viral proteins does not rely on unique mechanisms. The only unusual feature of the process is the location of the signal sequences that initiate the attachment of the ribosome to the endoplasmic reticulum. With most of the animal-cell glycoproteins whose synthesis has been closely examined the signal sequence is coded for by nucleotides that immediately follow the start codon; the signal sequences of the viral *E*1 and *p*62 proteins are coded for by nucleotides in the interior of the messenger-RNA chain. Too few messenger-RNA molecules coding for cell proteins have been fully characterized, however, for one to exclude the possibility that some of them may also include interior regions coding for signal sequences.

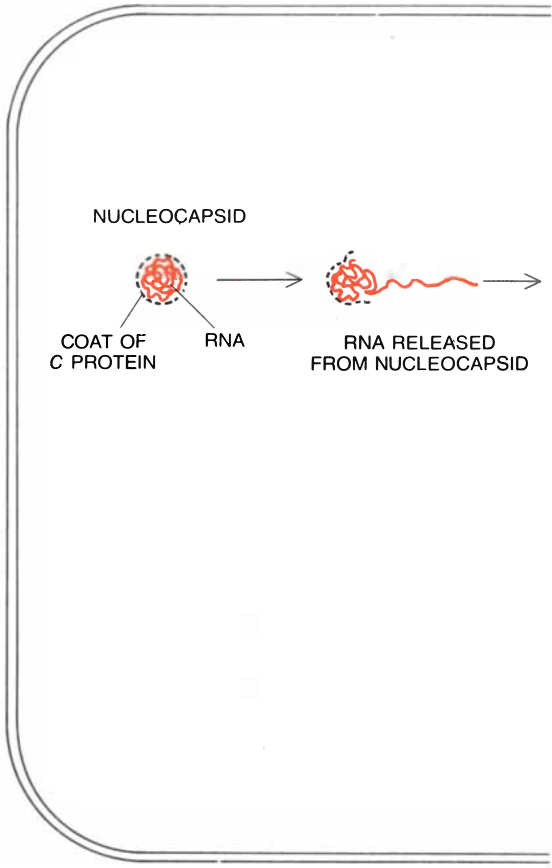
The reliance of the virus on the normal processes of the cell continues as the spike proteins are modified, transported and inserted into the cell membrane. Soon after the *E*1 and *p*62 proteins are made they are attached to one another in the membrane of the endoplasmic reticulum. The complex they make up is transported to the surface of the cell on the same path followed by secretory proteins and some membrane proteins: from the endoplasmic reticulum the spike proteins move to the Golgi apparatus, a complex organelle with several interior spaces enclosed by membranes. In the Golgi apparatus the carbohydrate side chains are modified. Michael F. G. Schmidt and Milton J. Schlesinger of the Washington University School of Medicine have shown that in the Golgi apparatus molecules with long hydrophobic side chains are attached to each protein. Such molecules are known to be attached to cell-membrane glycoproteins and to the membrane proteins of certain other viruses, but the significance of the modification is not yet known.

In its passage through the Golgi apparatus the complex consisting of the *E*1 and *p*62 proteins retains a single orientation, which is in a sense the reverse of the one it will have in the viral membrane. The segment that will form the spike protruding outward from the viral membrane is put into the internal compartment of the Golgi apparatus; the segment that will traverse the viral membrane is inserted into the membranes of the apparatus; the short segments of the protein molecules that will be attached to the capsid extend out of the apparatus into the cytoplasm.

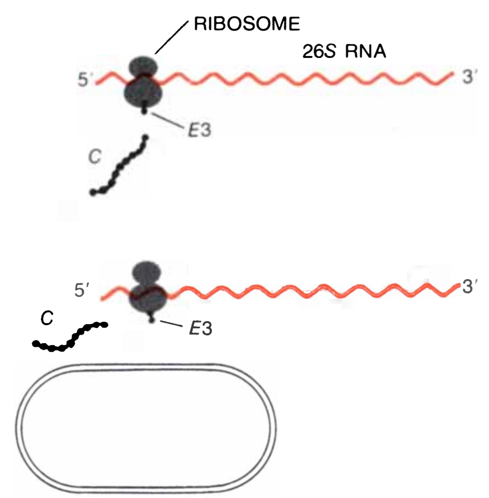
After the joined glycoproteins have passed through the Golgi apparatus they are carried to the surface of the host cell and inserted into the cell membrane, but little is known of how such movement into and out of organelles takes place. The cleavage of the *p*62 protein into the *E*2 and *E*3 chains occurs at about the time the complex reaches the cell surface. The cleavage is accomplished by a protease. Many secretory proteins, including hormones and blood proteins, are converted into their final form in a similar way.

By the time the complex of spike glycoproteins has been inserted into the cell membrane the nucleocapsid has been assembled in the cytoplasm out of *C* protein and nucleic acid. The final phase in the transit of the Semliki Forest virus through the animal cell can then begin: the formation of the complete virus particle and its release from the cell. Nicholas H. Acheson and Igor Tamm of Rockefeller University first showed that the SFV particle is assembled and released by a budding outward of the cell membrane. A section of the membrane is wrapped around the nucleocapsid and the bud is released from the cell. Most other viruses that have membranes, including influenza viruses, measles virus and rabies virus, are also liberated from the host cell by budding.

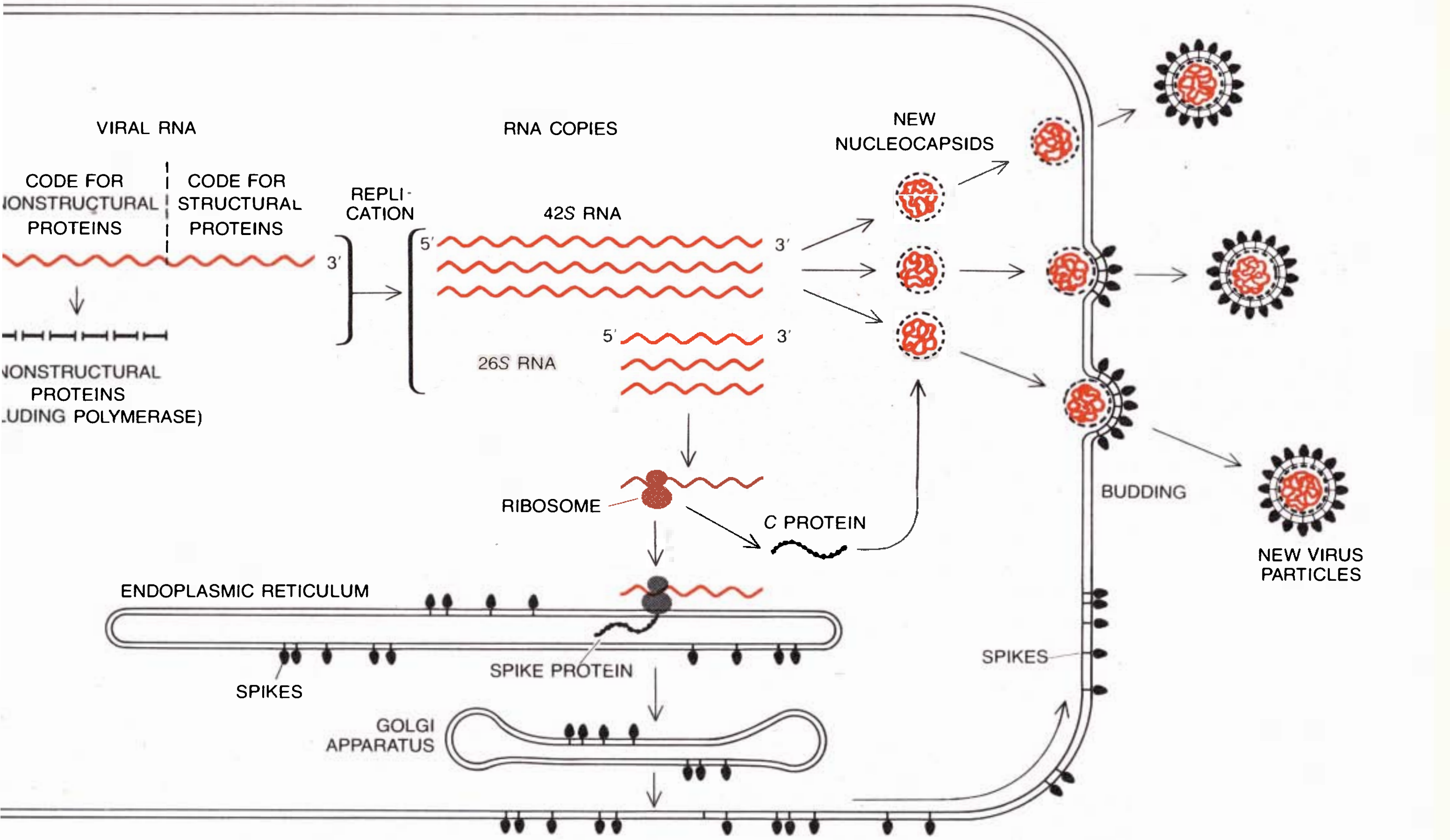
The key to understanding the final assembly of the Semliki Forest virus was obtained when we found that the viral proteins span the cell membrane and extend into the cytoplasm. The spike proteins in the cell membrane have the same orientation they will assume in the



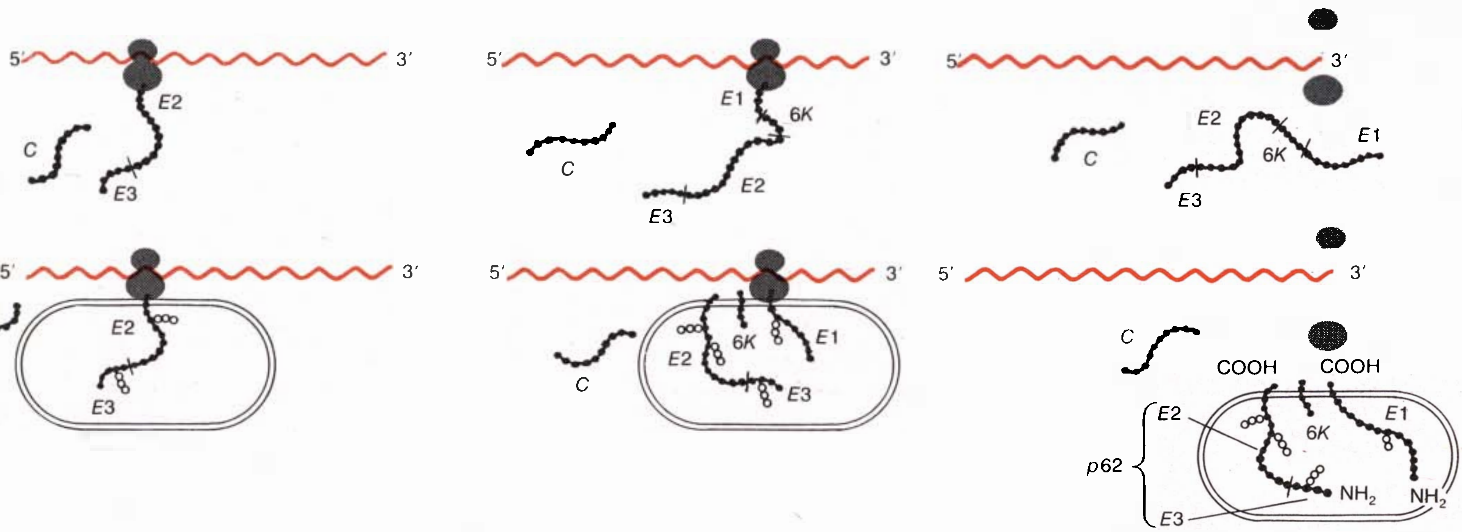
MANUFACTURE OF COMPONENTS of the SFV particle relies on the cell's apparatus for making proteins. After the nucleocapsid is freed from the lysosome replication begins. The viral RNA loses its protein shell; immediately afterward an RNA polymerase specific to the virus is made. (RNA polymerase is the enzyme that catalyzes the replication of RNA.) The polymerase and the RNA combine to make two kinds of new viral-



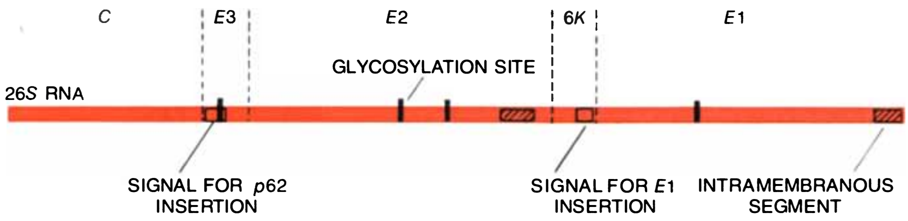
ENDOPLASMIC RETICULUM plays a significant role in the synthesis of the SFV spike proteins. If 26*S* RNA molecules are placed in a solution containing ribosomes, enzymes and the other substances required for protein synthesis, two amino acid chains are made: the *C*-protein molecule and a long chain that includes the *E*3, *E*2 and *E*1 molecules. This is not the way the spike proteins are made in the host cell. When vesicles made of membranes



RNA molecules. The 42*S* RNA is a copy of the complete viral-RNA molecule; the 26*S* RNA is a shorter molecule that incorporates the genetic code for the four viral proteins. (*S* stands for Svedberg unit, a measure of the molecule's mass.) Within three to four hours these are the most abundant RNA molecules in the cell. Now the 26*S* RNA becomes attached to a ribosome and begins to act as messenger RNA for the viral proteins. The four proteins are made continuously in the order *C*, *E*3, *E*2 and *E*1. As soon as the *C* protein is made it is clipped off by an enzyme. In the cytoplasm it joins a 42*S* RNA molecule to form a new nucleocapsid. The *E*3 protein is made next. When the first few amino acids of the *E*3 chain are assembled, the ribosome attaches itself to the endoplasmic reticulum, a network of interconnected membranes within the cell. The growing *E*3 chain is threaded through the membrane of the endoplasmic reticulum; assembly continues with the chain extending into the interior of the reticulum. When the three spike-protein molecules have been made, they are combined to form the membrane spike. They are then transported to the organelle called the Golgi apparatus, where they are modified. The spike is inserted into the cell membrane, where it meets a new nucleocapsid. Finally the finished virus particle buds out of the cell.



of endoplasmic reticulum are added to the solution, however, the viral proteins are manufactured as they are in the cell. The *C* protein is made and cleaved from the chain. The first few amino acids of the *E*3 chain constitute a "signal sequence" that binds the ribosome to the vesicle. Translation proceeds with the *E*3 chain threaded through the membrane. The *E*3 and *E*2 molecules are made consecutively and remain connected in a protein called *p*62. Translation proceeds until a point near the end of the *p*62 chain is reached. A stretch of hydrophobic amino acids there prevents the chain from passing through the membrane; the chain remains fixed to the membrane as translation continues. When translation of *p*62 is complete, the *p*62 chain is clipped off and the *E*1 molecule begins to be made. The *E*1 chain has its own signal sequence, 6*K*, which is later clipped off. Assembly of the *E*1 molecule proceeds until its passage through the membrane is stopped by a hydrophobic segment. As a result the *C* protein ends up outside the vesicle, in the region corresponding to the cytoplasm. Most of the *p*62 and *E*1 chains are inside the vesicle in the region corresponding to the interior of the endoplasmic reticulum. When the proteins are made in a cell, the *p*62 chain is later cut to form the *E*2 and *E*3 chains; the three proteins are then assembled into the spike.



26*S* RNA MOLECULE carries the genetic information needed to make the four viral proteins. The RNA molecule provides the information that directs the ribosome which amino acids to add to the growing chain. Glycosylation sites represent the points on the protein where carbohydrate side chains will be attached. The signal sequence recognizes the spot on the endoplasmic reticulum to which the ribosome attaches. Many messenger-RNA molecules in animal cells have regions that code for signal sequences. In the RNA's that have been examined in detail, however, these regions are found at the end of the RNA molecule, rather than in the middle as the *E*3 and 6*K* sequences are. Few animal-cell messenger RNA's have been studied closely enough to determine their entire genetic code; some may contain such internal regions coding for signal sequences as the viral RNA does. Apart from this possible anomaly synthesis of the viral proteins precisely follows the pattern of synthesis of many of the cell's own proteins.

viral membrane, with the spike protruding outward and the short stubs of the *E*1 and *E*2 proteins extending into the cell. Our work suggests that the driving force of the budding process is provided by the chemical interaction of the stubs with the *C* protein.

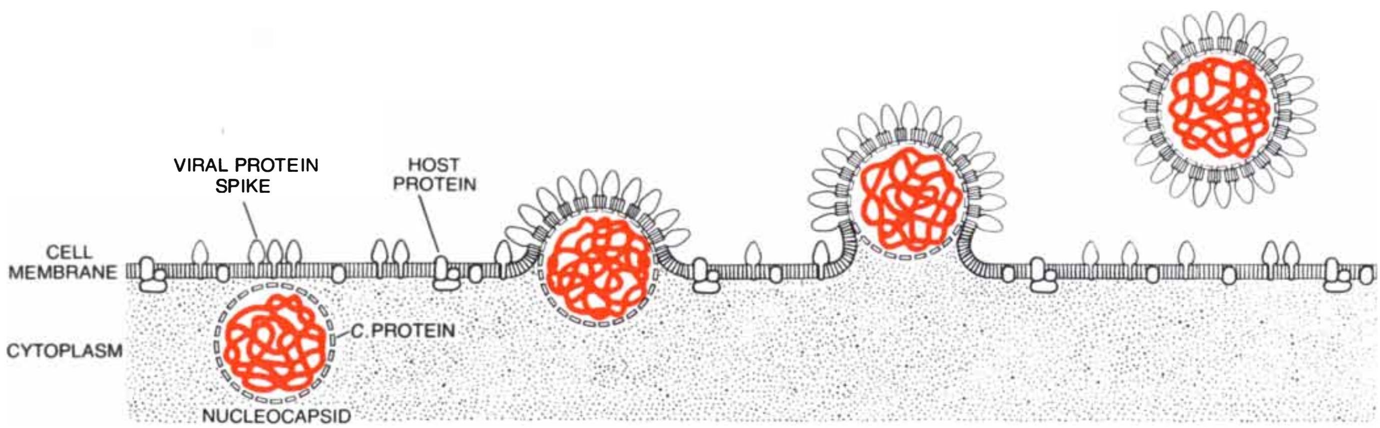
Budding begins when the nucleocapsid diffuses from the site in the cytoplasm where it is made to a point just under the cell membrane. We hypothesize that one *C*-protein molecule binds to the stubs of a single spike. The bond between the stubs and the *C* protein anchors the nucleocapsid to the membrane. The nucleocapsid then serves as a point of attachment for more spikes. Because of the two-dimensional fluid character of the cell membrane both the viral glycoproteins and the cell glycoproteins are free to move laterally. When a spike moves over the spot where the nucleocapsid is anchored, it binds to a vacant *C* protein. As more bonds form between spike proteins and *C* proteins the membrane of the cell is forced to curve around the nucleocapsid. When all 180 capsid molecules are attached to spikes, the membrane has been wrapped completely around the nucleocapsid. The membrane then breaks off, releasing the virus particle into the space between the cells. No hole in the membrane remains after the outward passage of the particle because lipid membranes can reseal spontaneously.

The mechanism we have proposed provides a ready explanation for the fact that the lipids composing the viral membrane are the same as those of the host-cell membrane. It is not as clear, however, why no host-cell membrane glycoproteins are found in the viral membrane. It may be that the host proteins are pushed away from the budding site because their chemical affinity for the *C* proteins is less than that of the viral glycoproteins.

A few hours after the entry of the Semliki Forest virus into an animal cell newly formed virus particles have begun to bud outward from the cell membrane. Once the particles are in the intercellular space they move outward toward other cells and spread the infection. When the virus cycle is considered from entry to exit, the paucity of mechanisms unique to the virus itself is striking. The binding of the particle to the cell surface and the fusion of its membrane with the membrane of the lysosome depend on the viral membrane proteins; an RNA polymerase specific to the virus catalyzes the replication of the viral nucleic acid; the exit from the cell requires the interaction of the viral spike proteins with the viral *C* protein.

For the rest the virus relies on the normal processes of the animal cell. It enters the cell by a route that is taken by many large molecules needed for metabolism. The path from the ribosome to the endoplasmic reticulum to the Golgi apparatus to the cell membrane is the one taken by numerous cell proteins. The manufacture of the viral-RNA molecules and the four viral structural proteins exploits the apparatus of protein synthesis that is normally operating in the animal cell.

The heavy reliance of viruses, the Semliki Forest virus among them, on the normal functions of the host cell suggests the reason that antiviral agents have been less effective than antibacterial agents: it is difficult to prepare a drug that interferes with the replication of the virus without simultaneously damaging the host cell. Work directed toward useful antiviral agents should probably be aimed at interrupting the few functions that are specific to the virus. On the other hand, the heavy reliance of the virus on normal cellular functions implies that work with viruses can provide clues not only to clinical practice but also to a deeper understanding of the life of the animal cell.



EXIT OF SFV PARTICLE from the host cell is achieved by the budding outward of the cell membrane. The process begins when a new nucleocapsid moves to a spot just under the membrane. A *C* protein binds to the short stub of the membrane spike, anchoring the nucleocapsid to the membrane. Bonds form between other *C*-protein molecules and other spikes. The spikes are free to move laterally in the cell membrane. They appear to have a greater chemical affinity for the *C* proteins than the cell's own membrane proteins do; the cell's proteins are therefore pushed away from the nucleocapsid. As binding continues the cell membrane curves around the nucleocapsid. When all 180 *C*-protein molecules are bound to spikes, the membrane has curved completely around the nucleocapsid and breaks off. The virus particle is thus freed into the space between cells to renew the process of infection. The exit mechanism explains why the lipid molecules of the viral membrane are of the same kinds as those of the host cell, and why the SFV proteins are specific to the virus.

As easy as apple pie.

And it's quicker and more forgiving about time and temperature. The new line of Kodak Ektaflex PCT products represents a significant refinement of the electron-transfer chemistry and the marvelously predictable diffusion processes that go on in the color-sensing layers of our instant color film.

This technology will allow thousands the pleasure of making their own color prints at home without the need for an elaborate darkroom and without complicated processing procedures.

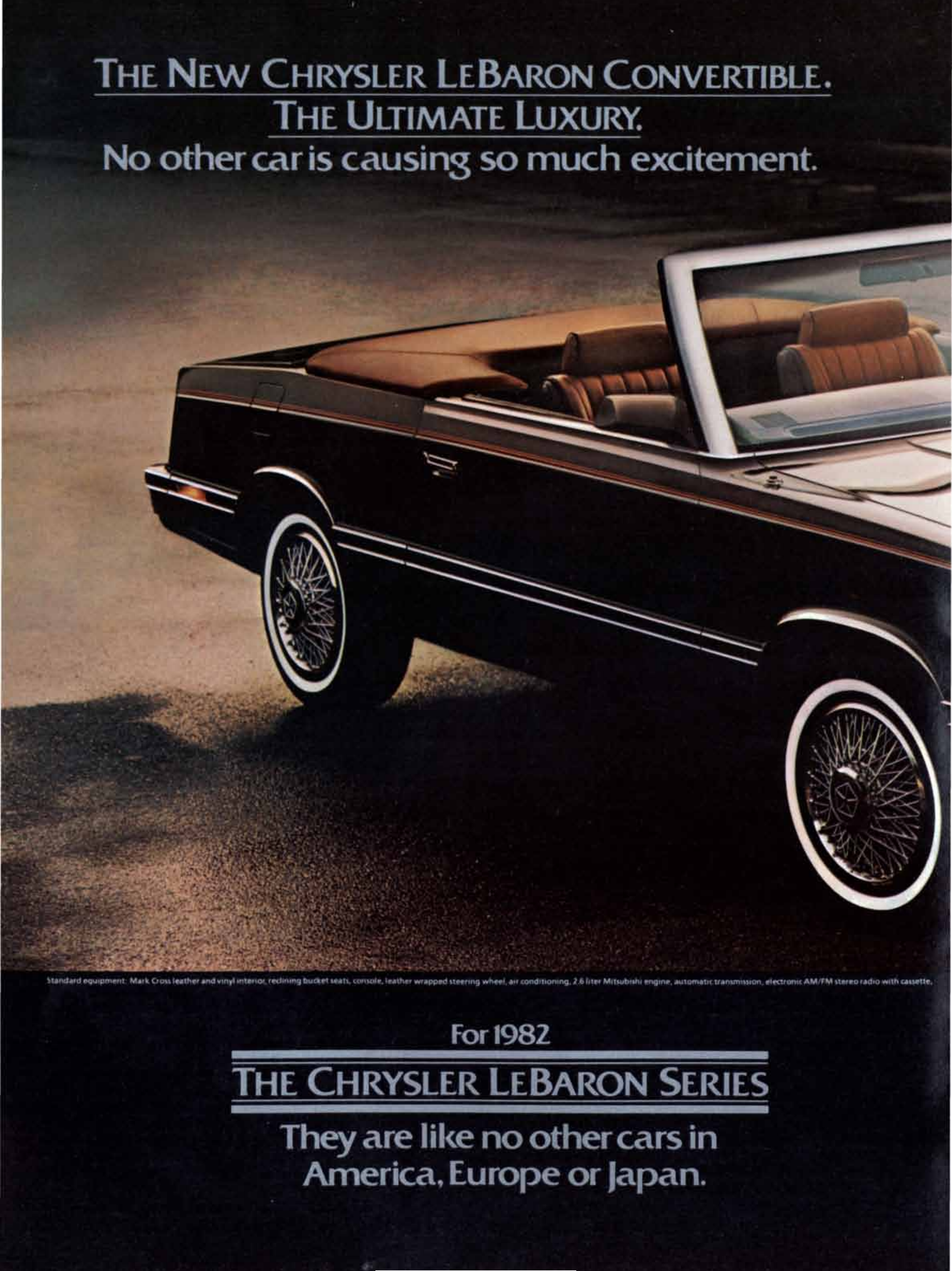
From the time an original negative or slide is exposed via traditional color enlargement projection, a finished color print is ready in 6 to 12 minutes. At any room temperature between 65 and 80° F.

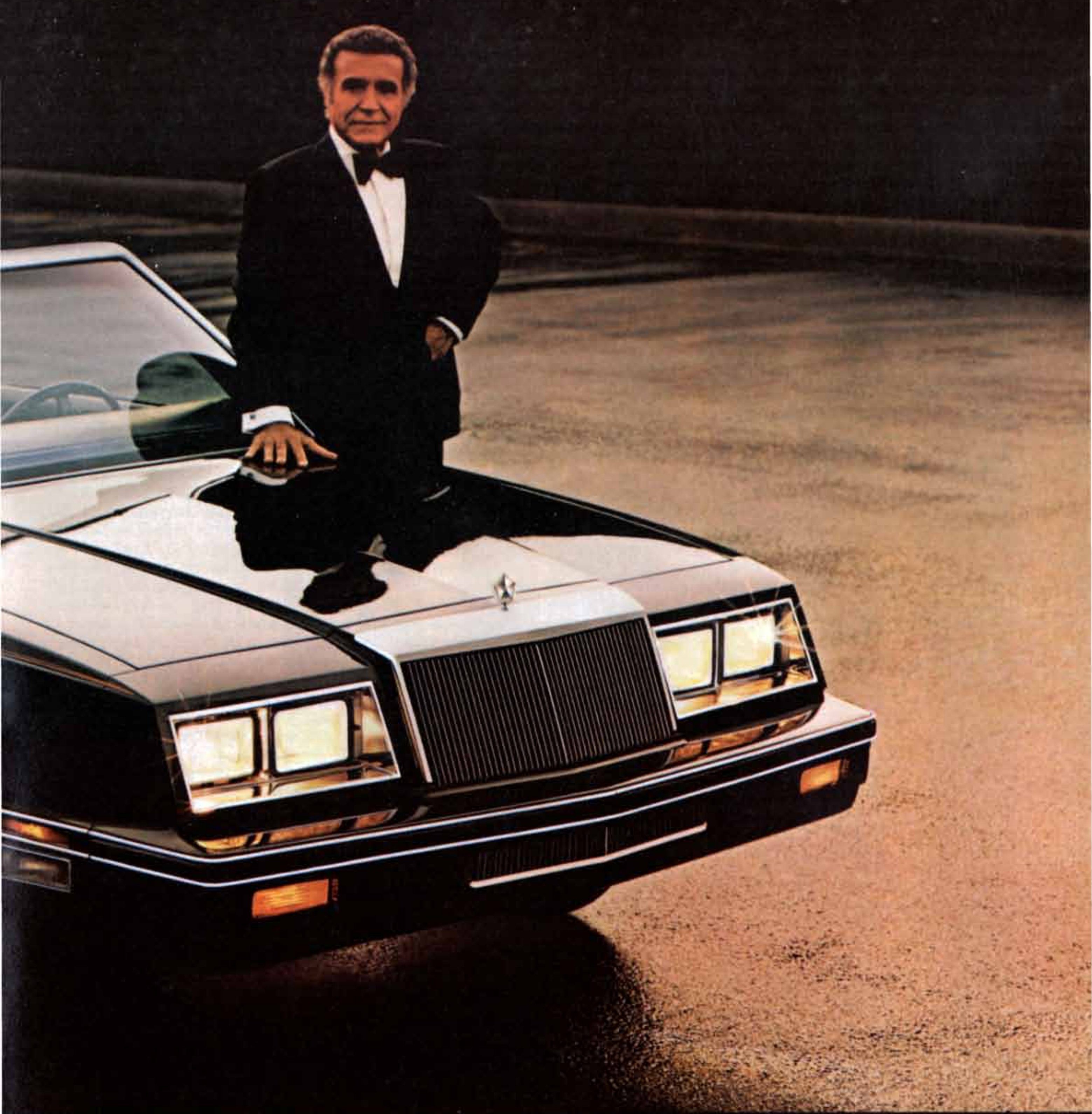
And print quality and stability are comparable with that of our line of Ektacolor papers. Note the brilliant reds reproduced in the apples and the soft, accurate hues captured in the complexions of our models.

The Ektaflex printmaker made the image reproduced here.

If you'd like to know more, write to us for more information at Eastman Kodak Company, Dept. A7-GB, Rochester, NY 14650.

Kodak


THE NEW CHRYSLER LEBARON CONVERTIBLE.
THE ULTIMATE LUXURY.
No other car is causing so much excitement.
Standard equipment: Mark Cross leather and vinyl interior, reclining bucket seats, console, leather wrapped steering wheel, air conditioning, 2.6 liter Mitsubishi engine, automatic transmission, electronic AM/FM stereo radio with cassette.
For 1982
THE CHRYSLER LEBARON SERIES
They are like no other cars in
America, Europe or Japan.



power windows, door locks, deck lid, rack and pinion power steering, power front disc brakes/rear drum brakes, electronic digital clock, WSW steel belted radial tires, dual remote mirrors, cornering lamps, and concealed wipers.

Lee Iacocca's dream to combine high mileage and luxury in a series of cars has come true with the 1982 Chrysler LeBaron Series. And the LeBaron convertible is the ultimate luxury. It exudes an elegance, a sense of style that starts at the tip of its crystal hood ornament and continues through to its plush interior.

The LeBaron Mark Cross Edition comes so fully equipped all you need to do is just select your color, choice of sound system, and either wheel covers or aluminum road wheels. It even has a power operated top that folds behind its leather and vinyl rear seat.

Chrysler's advanced front-wheel-drive technology, precise assembly by computerized robots, a refined suspension system and special body reinforcements help the LeBaron convertible achieve remarkable stability and control. The Chrysler LeBaron convertible. High mileage. Ultimate luxury.

Reservations are now being accepted for spring delivery. For information see your Chrysler-Plymouth dealer.

31 EST HWY PROJECTED | 23 EPA EST MPG* PROJECTED | $13,900** PRICE AS SHOWN

The New Chrysler Corporation

Use this Chrysler projected EPA est. MPG for comparison. Actual mileage may vary, depending on speed, trip length and weather. Hwy. MPG probably lower. Projected Calif. est. Hwy. 28, EPA est. MPG 21. **Sticker price as of December 11, 1981, publication closing date, excluding taxes, title and destination charges.

SCIENCE AND THE CITIZEN

The Business of Research

Is basic research in the sciences in danger of being subverted when universities or their faculty members seek to profit from its commercial applications? The question has been raised by at least two events in recent months. In response to increasing involvement by its faculty in business enterprises, particularly in biological technology, Harvard University has adopted guidelines governing conflicts of interest and commitment. Stanford University has secured a patent on a basic technique of genetic manipulation and is selling the right to apply the technique commercially.

The background to these events is the decrease in Federal support for research. At the same time the development of techniques for inserting an exogenous gene into a cell has created the possibility of a profitable new industry. The commercial potential of the new genetic methods has led to a large increase in the amount of money individuals and corporations invest in biological research. In 1981 private investment in such research amounted to about $300 million, according to Nelson M. Schneider of E. F. Hutton and Company. Schneider adds that within a few years the amount of private funding in biology may approximate the funding available from Federal sources.

The financial potential of genetic engineering is apparent not only to investors but also to workers in the field. A number of investigators have established business relations with biotechnology firms. Walter Gilbert of Harvard, who developed methods of deciphering the sequence of nucleotides in DNA, has taken a year's leave of absence from the university to serve as an executive of Biogen. Christian B. Anfinsen left the National Institutes of Health last year to work for DNA Science, a company formed by E. F. Hutton that has since been dissolved. Raymond C. Valentine of the University of California at Davis has become a vice-president of Calgene. Mark Ptashne of Harvard is the founder of Genetics Institute, Inc.; Winston J. Brill of the University of Wisconsin at Madison has an association with the Cetus Corporation; Timothy C. Hall, on leave from the University of Wisconsin, is associated with Agrigenetics.

A research worker serving in the management of a company whose business is closely related to his research suggests several concerns. Among them are the possibility that basic research will be diverted from questions contributing to the fundamental understanding of nature to those whose answers offer profit; that corporate secrecy will replace the relatively free exchange of information in the academic community, and that the quality of advanced teaching will deteriorate as professors teach less or guide their students toward financially rewarding subjects.

The guidelines adopted by the Harvard faculty in December are meant to resolve such conflicts without "particularly condemn[ing] entrepreneurial enterprise," according to Paul C. Martin, associate dean of arts and sciences and one of the authors of the rules. Under the guidelines a faculty member may spend no more than 20 percent of his time doing outside work (a rule widely accepted previously as an informal standard). Activities that "seem likely to present an unacceptable conflict" include assuming "executive responsibilities for an outside organization that might seriously divert... attention from University duties," exploiting unpublished results of university research for financial gain and passing unpublished results to an outside organization for its gain. On the other hand, the guidelines note that directing a student into work from which the professor hopes to realize financial gain is by itself an activity in which "no irreconcilable conflict of interest or commitment is likely to be involved."

According to Derek C. Bok, the president of Harvard, the rule forbidding faculty members from assuming executive positions was made because "certain types of responsibilities are incompatible with an academic position." He added that the participation of professors in profit-making enterprises related to their fields presents a greater threat to academic research than the increase in corporate funding of such research does. "We know how to control bilateral research agreements with the corporations; it seems to me that the problems are considerably more serious in relationships that give professors an incentive to make money rather than to do good science." Keeping research results secret and directing research into areas that are profitable rather than scientifically important are the greatest risks of such relations, he said.

The events at Stanford concern the commitments of the university rather than those of its faculty. The university has obtained a patent on the technique that allows a segment of DNA from a virus or a plasmid (a free bacterial gene) to be recombined with other genes and inserted into a cell. The expression of the recombinant genetic material could have many applications in the chemical and pharmaceutical industries and in agriculture. The first commercially important products are likely to be human hormones, including insulin. The technique was developed by Stanley N. Cohen of the Stanford University School of Medicine and Herbert W. Boyer of the University of California Medical Center in San Francisco.

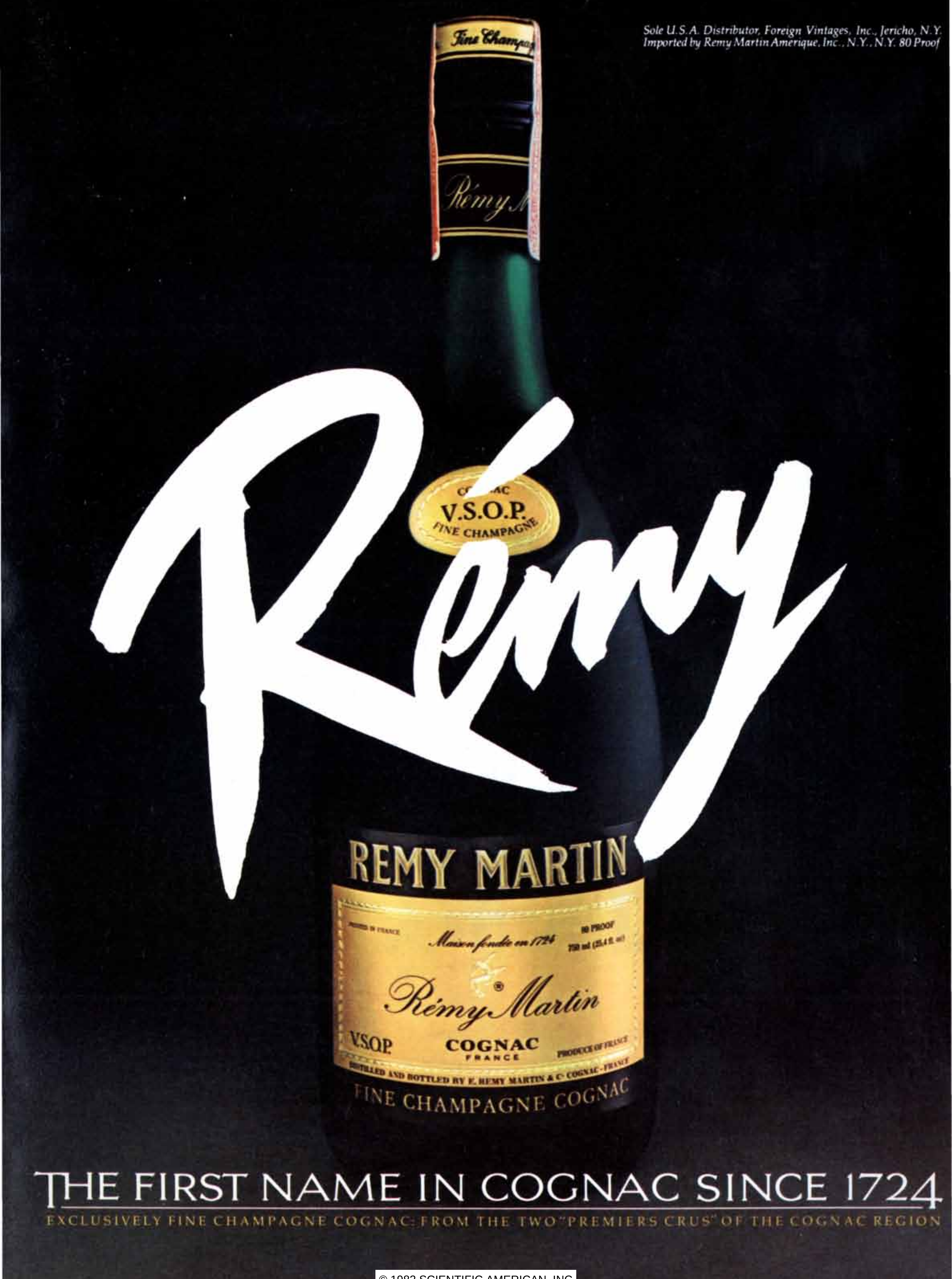
Boyer and Cohen have assigned their rights under the patent to their universities. Nonexclusive licenses to employ the method are being offered for $10,000, and in addition the licensees will pay an annual fee of $10,000. Stanford is to administer the program; after direct costs are paid the royalties are to be divided by the two universities. So far 73 licenses have been sold.

Donald Kennedy, the president of Stanford, points out that patenting arrangements can to some extent compensate for reductions in Federal funding. One of the most serious problems faced by universities is that their laboratories are becoming outdated, and workers may be tempted to move to better-equipped and better-staffed private research facilities, Kennedy said. The result would be the separation of research and graduate training; the combination of the two is now an important source of strength for basic research. Preventing undue faculty involvement in private research is "very hard to do through a strictly regulatory approach; the best thing to do is to develop strong incentives to keep things inside [the university]," Kennedy said.

In the recent controversies over private enterprise in biology it has often been overlooked that there is ample precedent in other disciplines for such ventures. The practice of faculty members serving as consultants to industry is well established and rarely criticized. The practice is particularly widespread in chemistry and engineering, where it has not adversely affected research in any obvious way.

The boundary between business and the university has been crossed with notable frequency in electronics, lasers and semiconductors. The firms along Route 128 near Boston and in the "Silicon Valley" of northern California owe much of their vitality to the participation of faculty from nearby universities. Furthermore, universities have in the past utilized patents on innovations by their faculty as a source of revenue. Indiana University has derived substantial income from its patent on the formula for the fluoride solution used in dentistry. The University of Florida has benefited from its patent on the formula for the drink sold as Gatorade.

From the point of view of those concerned with the continued well-being of the sciences, the danger in the current situation is not that some individuals or institutions will enrich themselves through basic research. On the contrary,


Sole U.S.A. Distributor, Foreign Vintages, Inc., Jericho, N.Y.
Imported by Remy Martin Amerique, Inc., N.Y., N.Y. 80 Proof
Rémy
V.S.O.P.
FINE CHAMPAGNE
REMY MARTIN
Maison fondée en 1724
80 PROOF
Rémy Martin
V.S.O.P.
COGNAC
FRANCE
PRODUCE OF FRANCE
FINE CHAMPAGNE COGNAC
THE FIRST NAME IN COGNAC SINCE 1724
EXCLUSIVELY FINE CHAMPAGNE COGNAC: FROM THE TWO "PREMIERS CRUS" OF THE COGNAC REGION

"Planet Earth has a furnace inside," says Glen E. Campbell, Gulf's Geothermal Supervisor. "The heat is generated by the natural radioactive decay of elements deep in the earth, and is conducted outward toward the surface.

"But the earth's crust isn't the same thickness everywhere. So in some places, this geothermal energy is buried miles deep, and in other places it comes right to the surface as volcanoes, geysers or hot springs.

"In many areas of the world, including the western U.S., geothermal resources hot enough to generate electricity come fairly close to the surface. In one such area, at Steamboat Springs, near Reno, Nevada, Gulf and another energy company are working on various techniques for extracting heat from the earth and putting it to work generating electricity.

"The potential is tremendous. Government figures show that the geothermal resources within the U.S. could free roughly a million barrels of oil per day for other uses."

Geothermal energy could be a big step toward energy independence. It's part of the business Gulf is in: energy for tomorrow.

Gulf people: energy for tomorrow.

© Gulf Oil Corporation 1981

Electricity from Earth's body-heat.

"Generating electricity from geothermal resources could save America a million barrels of oil every day."

There's a special coming in late January
"DISCOVER: The World of Science"
sponsored by the
Gulf Oil Corporation
check your local TV listing for time & station.

the threat is that a fundamental change in the relations of the university with its faculty or with businesses will impair its capacity to foster such research. In basic (as opposed to applied) research the investigator must be free to examine any important question suggested by work in progress. This freedom in turn depends largely on the independence of the university.

Two developments could reduce this independence. The first is the formation by universities of corporations aimed at exploiting research results for profit. Several universities have reportedly considered the possibility of forming such corporations, in which senior faculty members would serve as staff. They have rejected the idea for the time being. If a profit-making organization were made a part of the university, it would be difficult to exclude commercial considerations from decisions on hiring, promotions and salaries. The second danger is that in an attempt to procure corporate funds the research agenda of the university may be shifted toward applied research. Because basic research rarely yields a direct return only a small proportion of corporate funds are directed there. If universities begin to compete strenuously for private funds, applied research may come to predominate in academic departments.

The Big Blank

A group of four astronomers has found evidence of a great void in space: a region of the universe notably lacking in galaxies. If the comparative emptiness of the region is confirmed, it will be the largest void known; indeed, it will constitute the first evidence that the universe has inhomogeneities on a scale as large as a few percent of its overall size. The finding has several implications for cosmology.

The investigators are Robert Kirshner of the University of Michigan, Augustus Oemler, Jr., of Yale University, Paul L. Schechter of the Kitt Peak National Observatory and Stephen A. Shectman of the Mount Wilson and Las Campañas observatories. They had set out to study the large-scale distribution of galaxies in the universe. To do so they picked three squares in the sky, each square being 1.4 degrees on a side. The squares lay in the direction of the north pole of the Milky Way, a direction chosen to avoid the central plane of our galaxy and in particular the central lane of dust that blots out the light from objects farther away. In each square the investigators endeavored to count all the galaxies brighter than a certain limit. For most such galaxies they could determine the shift of the spectral lines in the radiation toward the red end of the spectrum. The shift reveals the velocity with which the galaxy is moving away from our own. According to a hypothesis first advanced by Edwin P. Hubble the velocity is proportional to the distance of the galaxy; the proportionality is given by the Hubble constant.

The squares had surprising statistics. If galaxies were scattered more or less evenly throughout the universe, the greatest number of galaxies in the squares would have red shifts that correspond to recessional velocities of between 12,000 and 18,000 kilometers per second. In particular the investigators could have expected to find 25 galaxies with recessional velocities in that range. Instead they found only one. Equally surprising, each square had a concentration of galaxies with recessional velocities of about 9,000 kilometers per second and a second concentration of galaxies with recessional velocities of about 21,000 kilometers per second.

The three squares are separated from one another by arcs of 35 degrees across the sky. Still, the investigators propose that the scarcity of galaxies at the same range of red shifts in all three squares may well be evidence of a single great void in the universe. In order to specify the position and size of the void it is necessary to assign a numerical value to the Hubble constant; a plausible value is 50 kilometers per second per megaparsec. (A megaparsec is 3.26 million light-years.) In that case a galaxy receding at a velocity of 15,000 kilometers per second would be 300 megaparsecs from the solar system. The void is presumably centered there. It has a depth on the order of 100 megaparsecs and a volume on the order of a million cubic megaparsecs.

What would the presence of the empty region signify? The assumption that galaxies are scattered evenly throughout the universe is an approximation. As long as five decades ago it was recognized that galaxies tend to lie in clusters; thus the universe has inhomogeneities on the order of a megaparsec in size. Then it was recognized that clusters lie in superclusters, which represent inhomogeneities on the order of 10 megaparsecs in size. The discovery of the void suggests there is at least one inhomogeneity on the order of 100 megaparsecs. In comparison the horizon of the universe (the distance at which a receding galaxy would have the speed of light, if that were possible) is 6,000 megaparsecs away, assuming the value of 50 for the Hubble constant. Hence the newly discovered void has only a few percent of the characteristic length of the universe.

The investigators note that the discovery of a void this size does little to subvert the hypothesis that the universe expanded from an explosion: the big bang. Even if the universe was almost homogeneous early in its expansion, some regions would have been slightly denser than others; their gravitational self-attraction would then have been greater than the average, so that they would have resisted the universal expansion to a greater degree. In this way a slight concentration of matter would have grown more pronounced as the universe expanded. Conversely, a region slightly deficient in matter would have become more deficient. The problem is that the variations in density should affect such "fossils" of the early universe as the cosmic background of electromagnetic radiation. The background radiation is quite homogeneous. If inhomogeneities in the distribution of galaxies on a scale of 100 megaparsecs turn out to be common in the universe, the background radiation should be observably "lumpy."

Cellular Treadmill

Microtubules are long, thin, hollow cylinders that play various structural and motor roles in every eukaryotic (or nucleated) cell. Along with other fibrous components they provide the cell with an ever changing three-dimensional framework. They are the main structural members of such motor organelles as cilia and flagella. They also function, either as cables or draglines or simply as tracks, to move things from place to place in the cell. A mechanism that has lately come to be understood may explain how the microtubules accomplish some of these tasks. In certain contexts, at least, it seems that a microtubule can act as a treadmill or an escalator and might thereby carry attached materials along in its motion.

Microtubules are polymers of the protein tubulin, which itself is a double molecule composed (like a squat snowman) of two similar globular proteins, alpha and beta tubulin. The subunits aggregate end to end to form a strand, and 13 strands wrap up to form the cylindrical microtubule. Most microtubules are extremely labile: they are continually being assembled from a cellular pool of tubulin subunits and then disassembled.

The process of assembly and disassembly has been studied primarily in laboratory preparations rather than in living cells. It became clear that subunits are added to and removed from the microtubule at both ends. A few years ago Gary G. Borisy of the University of Wisconsin at Madison and his colleagues found that when the microtubule is growing, the rate of growth is different at the two ends: there is an intrinsic polarity of growth. In 1978 Robert L. Margolis of the Fred Hutchinson Cancer Research Center in Seattle and Leslie Wilson of the University of California at Santa Barbara discovered a related phenomenon under steady-state conditions, that is, when the amount of tubulin incorporated into microtubules and the amount free in the medium are in equilibrium. Under these circumstances the overall rates of assembly and disassembly must be equal, but Margolis and Wilson found that subunits are mainly added at one end and mainly lost at the other end. The result is a unidirectional flux of subunits along the microtubule. In a recent article in *Nature* Margolis and Wilson review what has been learned about the microtubule treadmill and how it may work in the cell.

Treadmilling was discovered when Margolis and Wilson investigated the mechanism whereby the drug colchicine, which binds to tubulin, inhibits the growth of microtubules and so interferes with their functioning. They recognized that the blocking of polymerization could not require the binding of a colchicine molecule to every tubulin subunit because growth is inhibited when only a small proportion of the free tubulin is bound. Instead a few colchicine-tubulin complexes must assemble at the end of the microtubule and somehow block further assembly. That would be possible, however, only if subunits are not readily disassembled at the drug-complex site; otherwise the complexes would drop off and uncomplexed subunits could continue the polymerization. There must be, Margolis and Wilson reasoned, a net assembly of subunits at one end (which can therefore be blocked by colchicine) and a net disassembly at the other end. The subunits must get on at the assembly end, travel the length of the microtubule and be dropped off at the disassembly end.

The reality of the treadmill (at least as a laboratory phenomenon) was confirmed by experiments in which free subunits were labeled with radioactive atoms and then timed as they moved onto and eventually off microtubules of known average length. Assembly can predominate over disassembly only when the concentration of free tubulin subunits exceeds a critical level. In a laboratory medium the concentration is essentially the same everywhere, and so the existence of treadmilling implies that the critical concentration needed to favor assembly is lower at the assembly end than it is at the disassembly end. Energy to maintain the disparity is supplied by the breakdown of the nucleotide guanosine triphosphate (GTP), in the absence of which treadmilling does not take place.

Clearly a treadmill can do work. For a flux of subunits to move an organelle through the cell, however, the organelle must be attached to the treadmill and the microtubule must be anchored in the cell, and yet both ends must remain free to gain and lose subunits. Margolis and Wilson suggest that these conditions can be met if microtubules are attached both to their anchors and to whatever they are moving by means of lateral linkages projecting like oars from the subunits. As an example they discuss the operation of the mitotic spindle, the fibrillar apparatus that moves the two sets of chromosomes toward opposite poles of a dividing cell.

The mitotic spindle is composed of two kinds of microtubules. A set of interpolar microtubules extends from each pole toward the other pole, overlapping at the cell's equator; a set of kinetochore-to-pole microtubules extends from each pole to the kinetochore, or point of attachment, of each chromosome. In each half-spindle both kinds of microtubules have the same polarity; the model proposed by Margolis and Wilson assumes that the disassembly ends are at the poles. Two kinds of motion are observed in mitosis: the poles separate and the kinetochore microtubules shorten. Margolis and Wilson suggest that opposing sets of treadmilling interpolar microtubules slide or ratchet past each other, separating the poles. At the same time the chromosomes are reeled in toward the poles as the kinetochore microtubules (whose assembly at the chromosome end has been blocked) disassemble at the polar end.

Treadmilling could function in similar ways to move such things as viruses, secretory granules or mitochondria; it could also mediate changes in the skeletal framework of the cell. Margolis and Wilson think the actual work of translocation is probably done not by the breakdown of GTP in the treadmilling reaction itself but rather by some array of auxiliary machinery such as the proposed oarlike lateral elements. Moreover, treadmilling is not the only means by which microtubules might operate in translocation. They could remain stationary and propel organelles along their length with armlike attachments, or they could serve simply as tracks for some independent locomotive system.

Battle of the RAM's

The business of making the chips that are the building blocks of computers and other electronic devices has been dominated from the outset by the U.S. semiconductor industry. The American manufacturers are now facing their first serious competitive challenge. It comes from Japan, whose well-organized and rapidly growing semiconductor industry appears to have established a position from which it can dominate the market for the latest generation of electronic memory devices.

A chip is a small piece of silicon on which large numbers of transistors and other components are combined to form a complete electronic device. Chips have various functions, perhaps the most glamorous one being that of the microprocessor, which incorporates all the elements of the central processing unit of a computer. In terms of sales volume and profitability, however, the main interest is in the chips that serve as memories; in a computer a single microprocessor is generally supported by dozens or hundreds of memory chips. For the most part the chips are random-access memories, or RAM's, meaning that each memory cell of the chip can be addressed independently.

The capacity of the memory chips is measured in kilobits, or units of 2^{10} (1,024) binary digits. Since 1970 the capacity of the chips with the largest share of the market has increased fourfold every few years: from 1K to 4K to 16K. The transition to 64K chips is now under way. The present competition is over the 64K dynamic RAM, which actually has a capacity of 65,536 bits. It is called dynamic because it stores information in the form of evanescent electric charges, which must be refreshed periodically. The next generation, the 256K dynamic RAM, is expected to reach the market in about three years.

Several factors can be discerned in the Japanese success with the 64K chip. One is Japan's concerted effort, guided by the Ministry of International Trade and Industry, to develop an internationally competitive semiconductor industry. The effort was abetted by the government, which established a 10 percent tariff on imported semiconductors (the U.S. tariff is 5.8 percent) and restricted foreign investment in Japanese industry. Another factor has to do with technological strategy. In the U.S. several manufacturers believed the 64K chip represented such a big leap that it would require new concepts in design. The Japanese companies took the approach of extending the techniques that had been employed at the 16K level, fitting more circuit elements onto the chip by making the individual elements smaller or by making the chip larger. The result is that Japanese companies have what is widely agreed to be at least 50 percent of the market for 64K chips; some estimates range as high as 70 percent. Among American manufacturers the one with the largest share of the market (about 20 percent) is Motorola, Inc., which also decided to adapt existing methods of fabrication.

The Semiconductor Industry Association, an American trade group, has published an analysis of the competition titled *The International Microelectronic Challenge.* The report concludes that American manufacturers need help from the Government. The association calls for negotiations to reduce the tariffs of both the U.S. and Japan to 4.2 percent and to bring about "equal national treatment" for U.S. companies in Japan. According to the association, the Government should also adopt tax policies that encourage capital investment and research in the industry.

Although the U.S. appears to be strongly challenged in the market for the 64K chip and may for similar reasons be under pressure in the forthcoming competition to dominate the 256K market, the future does not look entirely bleak. Calculations of the market share include only the "merchant companies," which sell chips in the open market. In the U.S. 64K chips are also being made in large quantities by "captive companies," such as the International Business Machines Corporation and the Western Electric Company, which install the chips in their own products. Indeed, these companies may have been the first to develop the 64K RAM, and they have been meeting their own needs for the chips for some three years.

The market for the 64K chip is only beginning to build up, and the pattern may change as the demand grows. Moreover, at least one group in the U.S. is at work on a chip that would leapfrog the 256K memory by two generations. It is a four-million-bit (four-megabit) dynamic RAM.

And Littler Bugs inside 'Em

Entomologists have known for more than a century that the ability of the termite to metabolize cellulose can be attributed to enzymes synthesized by microorganisms in the termite's gut. Now a group of investigators has begun a census of the microorganisms present in one species of termite, *Pterotermes occidentis.* They have found a microscopic ecosystem of astonishing complexity: the termite is host to at least 100 species of protozoa and bacteria, of which more than 30 have so far been identified. The investigators are Leleng P. To and Lynn Margulis of Boston University, David Chase of the Sepulveda Veterans Administration Hospital in Sepulveda, Calif., and William L. Nutting of the University of Arizona; they present the results of their observations in *BioSystems.*

The inhabitants of the termite gut provide the ecologist with a clear example of the unit on which the science is based: a community of organisms that interact strongly with one another but whose interactions with conditions outside the community can be controlled. Moreover, the study of such a microscopic community has an advantage over the study of other ecological systems not only because its boundaries are well defined but also because the "field work" for the investigation can be done in the laboratory. The entire community can be kept in a covered jar and fed scraps of brown paper.

P. occidentis is a rare species found only in the Sonoran desert of northwestern Mexico and the U.S. Southwest. The complexity of the inner world of the termite contrasts sharply with its austere habitat. Colonies of the insects confine their activities to the dead interior wood of five or six species of desert plant, including the saguaro cactus and the palo verde tree. Simply finding the termites and collecting them has called for painstaking effort. In one study done by Nutting and his co-workers all the dead wood of 18 trees was sawed into short lengths and split; then the termites were extracted one by one.

Most of the microorganisms that inhabit the gut of the desert termite are found in the paunch of the insect, the third segment of its digestive tract. Four species of protozoa have been identified in the paunch, all of which have been found to engulf pieces of wood. The protozoa vary in length from three to about 300 micrometers; the smaller of them can be seen swimming among the undulating flagella of the larger ones. The population density of the protozoa is between a thousand and a million individuals per milliliter of the fluid in the termite gut.

In the same gut fluid the bacterial density is between 10 million and 100 billion cells per milliliter. Bacteria are found freely swimming in the chamber of the paunch, adhering to its wall, adhering to or inside the protozoa and adhering to or inside other bacteria. It has not yet been determined whether the wood-digesting enzymes in the termites are produced by the protozoa or by the bacteria that inhabit the protozoa.

It is likely that most of the symbiotic relations in the community are devoted to maintaining the internal environment of the gut. Many of the microorganisms, for example, cannot live in the presence of oxygen and have never been found outside the termite gut. Such anaerobic organisms probably evolved from free-living aerobic ones, but they have adapted to the oxygen-poor conditions inside the termite. According to To and her co-workers, however, there are probably patches of relatively high oxygen concentration in the termite gut; the anaerobes may depend on bacteria that serve as oxygen scavengers to protect them from such toxic concentrations.

Other examples of species interdependence in the community apparently maintain the concentrations of essential nutrients and the *p*H of the gut fluid. Spiral-shaped bacteria may feed on the waste products of protozoan metabolism, and nonmotile bacteria attach themselves to motile bacteria or to protozoa. Hence although the breakdown of cellulose is the most important outcome of the symbiosis for the termite, it may well be incidental to the functioning of the community as a whole.

Meddling with the Midden

The archaeologist takes a professional interest in the rubbish heap and often ascribes some meaning to the layered sequence of artifacts in the refuse of an earlier society. Children, who have always been more numerous than archaeologists, may also consider the trash pile an interesting site for exploration. What influence might the activities of children have on the later findings of archaeologists? The question has been investigated experimentally by two workers who between them qualify for membership in both groups.

The results of the investigation are reported in *American Antiquity* by Gawain Hammond, who is identified as the senior author, and Norman Hammond, of Rutgers University, the junior author. Field work for the study was "carried out in Berkeley, Calif., in January 1977 by the senior author, with the junior author serving as recorder.... An artificial trash pile... was constructed, using only nonbiodegradable trash selected from the available supply." The trash was modern, but, as in a midden composed mainly of potsherds, it was dominated by discarded containers. The commonest items were half-gallon wine jars and aluminum beer cans.

"During the following three days, the senior author, at the time 1.2 years old, engaged in 'child-play' activities at and around the trash pile for a total of three 30-minute periods; concentration on the task for more than 30 minutes at one time was difficult, although it was, even in the solitary mode, one with which the experimenter was familiar. All locomotion during the experiment was quadrupedal or tripedal (when one hand was used to move an artifact)....

"The initial pattern of behavior observed was a localized dismemberment of the trash pile, with the largest objects, the wine jars, being preferentially removed from it and rolled a short distance downhill.... During the second session (31–60 minutes) the remaining nucleus of trash was further disturbed.... Three beer cans were tossed into the air in a casual manner.... One of the wine jars previously rolled was picked up, the screw cap removed, and various pieces of bark and twig... inserted into the jar. The discovery of such unexpected vessel contents in many archaeological contexts would be regarded as the result of structured 'ritual' behavior; the present observation shows that similarly nonlogical circumstances can result from unstructured 'child-play.'...

"In the third session (61–90 minutes) the original nucleus was further depleted, with all remaining artifacts being moved, most for only short distances to the west. At the end of the experiment the trash formed a rough circle, with density slightly higher on the uphill eastern side close to the original locus of the trash pile."

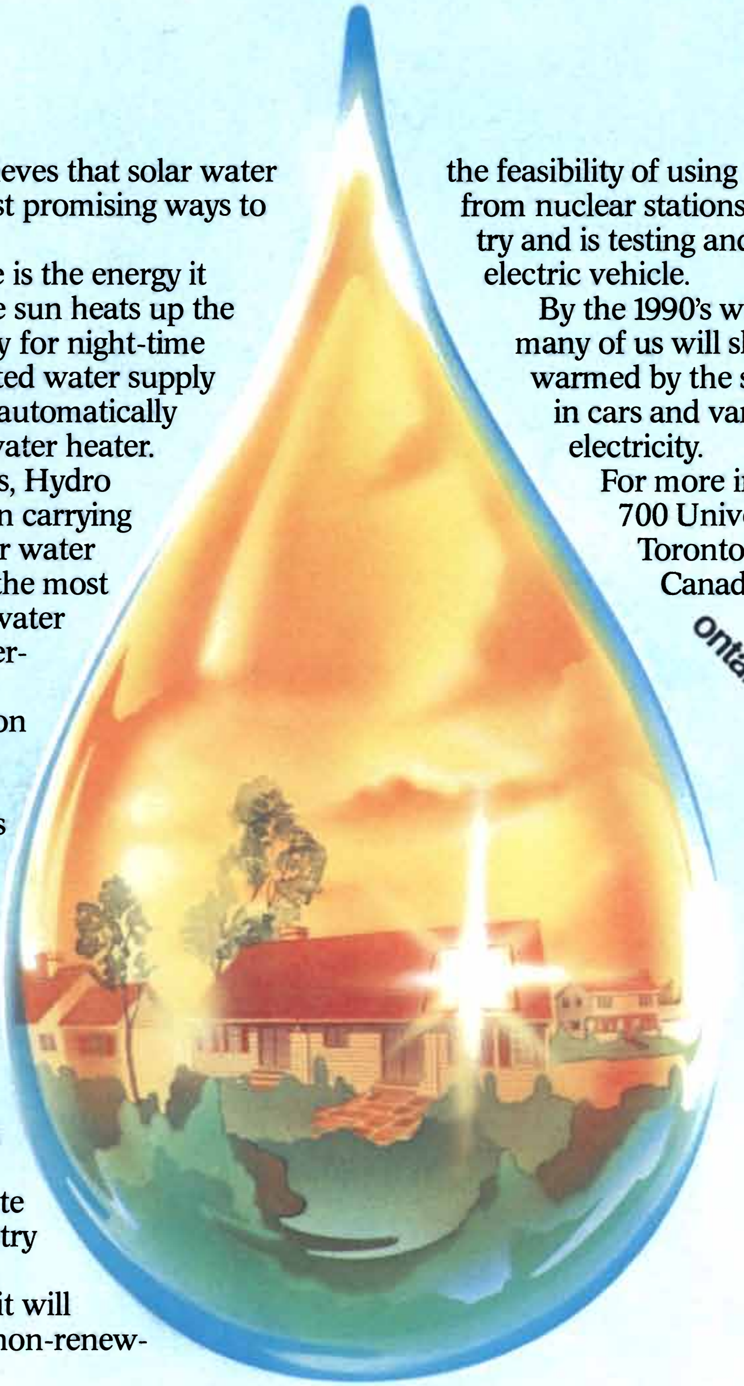
The experimenter and his recorder conclude that an episode of child play "may profoundly modify the initial archaeological pattern... and transmute it into an arbitrary pattern with a visible but unrelated structure; the causes of this change must be allowed for in investigations of artifact, and particularly trash, distributions if interpretive error is not to result."

CANADIAN TECHNOLOGY

The above image of Canadarm, the NASA Shuttle Remote Manipulator System developed and built in Canada, was created by an artist on Canada's computerized, two-way graphic and text distribution medium, Telidon. The following survey reports on these and other Canadian contributions to world technology.

SCIENTIFIC AMERICAN

Hot water from the sun? Just one of the sources we're tapping.



Ontario Hydro believes that solar water heating is one of the most promising ways to use the sun's energy.

The biggest advantage is the energy it saves. During the day the sun heats up the water and builds a supply for night-time use. When the solar-heated water supply isn't enough, the system automatically switches to the regular water heater.

For the past four years, Hydro research teams have been carrying out in-depth tests of solar water heaters. The program is the most complete study of solar water heater systems ever undertaken in Canada, underscoring Hydro's dedication to explore renewable energy possibilities.

By the end of 1983, it's expected the program will represent a multi million dollar investment shared equally by Ontario Hydro and the Ontario and Federal governments.

Investment in the development of practical solar water heating systems should help stimulate the Canadian solar industry in world markets. If that happens, a second benefit will be the saving of a lot of non-renewable energy world-wide.

Ontario Hydro supplies more than 111 billion kilowatt hours of electrical power annually to more than eight million people from a mix of generation sources including the highly successful CANDU nuclear system.

In the search for new and innovative sources of energy, Hydro plays a leading role and is exploring the possibilities of energy from many sources, such as hydrogen, wind, burning refuse and by growing algae culture. Hydro is also investigating the feasibility of using hot water or steam from nuclear stations to power industry and is testing and exploring the electric vehicle.

By the 1990's who knows how many of us will shower in water warmed by the sun or drive around in cars and vans fuelled by electricity.

For more information write:
700 University Avenue
Toronto, Ontario
Canada M5G 1X6

ontario hydro

We do more than make electricity.

Through the latter part of the last century, the technologies of steam and steel made Canada possible. In the latter part of this century, after more than a decade of analysis and debate, the technologies of energy and electronics have become Canada's conscious choices for growth.

To succeed the traditional *"three solitudes"* of government, industry and academia are entering an era of unprecedented collaboration. Concrete examples of the country's commitment to technological innovation abound – the Industrial Innovation Centres in Quebec and Ontario, the National Research Council's Integrated Research Centres in Quebec and Newfoundland, the Discovery Parks in British Columbia, the Innovation Place in Saskatoon and the Alberta Oil Sands Technology and Research Authority.

Provincial governments – especially Ontario and Quebec – are now acting on their belief that competitiveness in manufacturing depends on a sound and comprehensive technological base. The Federal Government has announced its determination to reverse the trend of the 1970's, and to expand national R & D activity by at least 50 percent during the next five years.

The reasons are not hard to find. Secondary manufacturing in Canada has over the years come to resemble a one-tenth scale model of the U.S. manufacturing sector. Through most of the last fifteen years, these *"miniature replica"* and branch plants have been in a state of economic stagnation. Canadians have lived well off their natural resources, while imports of manufactured goods mounted year by year. In 1980, the country's trade deficit in end products was a sobering $18 billion – six times the deficit of the early 1970's. Canada's deficit in services in 1980 was almost an additional $11 billion.

The country's trade surplus in raw materials and fabricated products has for many years fallen far short of covering these deficits. The difference has been met by net external borrowing – thereby compounding the national trade deficit in services in coming years. This deterioration in Canada's international payments position can be seen in international trading statistics. From 1969 to 1979, for example, the country's share of world exports declined fully 33 percent – the worst performance among industrialized countries during this period.

Increasingly overcrowded world markets in a widening range of commodities, and the emergence of newly industrialized countries with highly competitive process technology and labour costs seem to dictate one action – the dismantling of large parts of the mature industries which have historically been central to Canada's economy, replacing them with dynamic and competitive new industries.

But which industries? During the 1970's, investigations by the Science Council of Canada, the Ministry of State for Science and Technology, and the Senate Special Committee on Science Policy all concluded that a significant transformation in Canadian industry should occur through the selective development of industries utilizing high technology, and the encouragement of companies with innovative capability (whatever their degree of technological intensity). In other words, the development strategy should be selective, based on existing technological strengths and on particular sectors in which Canada already enjoys an international competitive advantage.

As John Roberts, Federal Minister of State for Science and Technology sees it, *"These are the areas which offer the best prospect for the ultimate edge for Canada – technological self-reliance. The sectoral approach best promises to mobilize the economic potential of this country and to meet its long-term needs. First, the sectoral strategy recognizes that our economy is largely resource-based and that we must use science and technology as the means of maintaining and enhancing our competitive edge in the development of these resources. We are currently piloting research efforts in areas which could have global impact in the future. Biotechnology, for example, particularly as it is applied to forestry and agriculture, holds enormous development potential and promises even more efficient use of our abundant natural resources. Second, the sectoral strategy would expand our resource orientation by applying the common-sense principle of processing our resources where that would be most efficient – near the resource base. Technology will play a key role in developing the petrochemical industry in Alberta and improving its international competitiveness. Third, a sectoral strategy would foster very new emerging technologies in which Canada clearly has a strong capability. Canadian know-how leads the world in wind energy, and by maintaining our present advantage, we will be able to capture future world markets in wind generators. We are also internationally recognized for our telecommunications technology. The Telidon videotex system has been incorporated into the standards of many prestigious organizations in North America and Europe and holds great promise for a wide range of applications. Fourth, a sectoral strategy would help to capitalize on the spinoffs from the megaprojects Canada will be undertaking in the next decade. Tar sands and pipeline developments, for instance, will require new breakthroughs in machinery and tooling equipment technology. Skillfully applied, this technology could yield fantastic advantages in other high potential industries."*

There is now an evident commitment by federal and provincial governments – and by industry itself – to take part in a "new industrial revolution" by moving increasingly into knowledge-intensive and technology-intensive industrial sectors. What follows is an account of these current developments in the inevitably overlapping fields of resource management, energy generation and advanced electronics.

ENERGY – TOWARD SELF-RELIANCE

Energy technology has always been important to Canadians for two obvious national needs – space heating requirements in a severe climate, and transportation fuels to move people and goods across vast distances. The steady depletion of the world's oil reserves has caused increasing emphasis on developing natural energy resources in which Canada is rich: natural uranium from large deposits in Ontario and

THE AUTHORS

Hugh C. McIntyre is a freelance writer on science and technology subjects. He received his B.S. in chemistry in 1949 from Mount Allison University in Sackville, New Brunswick and his M.A. in English from the University of Toronto in 1952. He presently lives in Toronto.

Frank J. Kelly is a policy consultant to industry and government. From 1970 to 1975 he was a science advisor to the Science Council of Canada.

THE COVER

Designer Rick Allen of St. Clair Videotex Design Inc., has symbolized Canada's contributions to world technology with the image he has created and the medium in which he worked. The image is of Canadarm, the remote manipulator system built for Canada's National Research Council as a contribution to the NASA Space Shuttle Program. The medium is Telidon, Canada's computerized, two-way graphic and text distribution medium.

Art direction, supervision of text and illustration and coordination of production were provided by Camp Associates Advertising Ltd., Toronto.

Saskatchewan for nuclear power, synthetic crude from the immense bitumen deposits of the Athabasca and related oil sands, new finds of natural crude at Beaufort Sea and Hibernia sites and still untapped hydro potential in the James Bay complex.

CANDU: Exporting Expertise

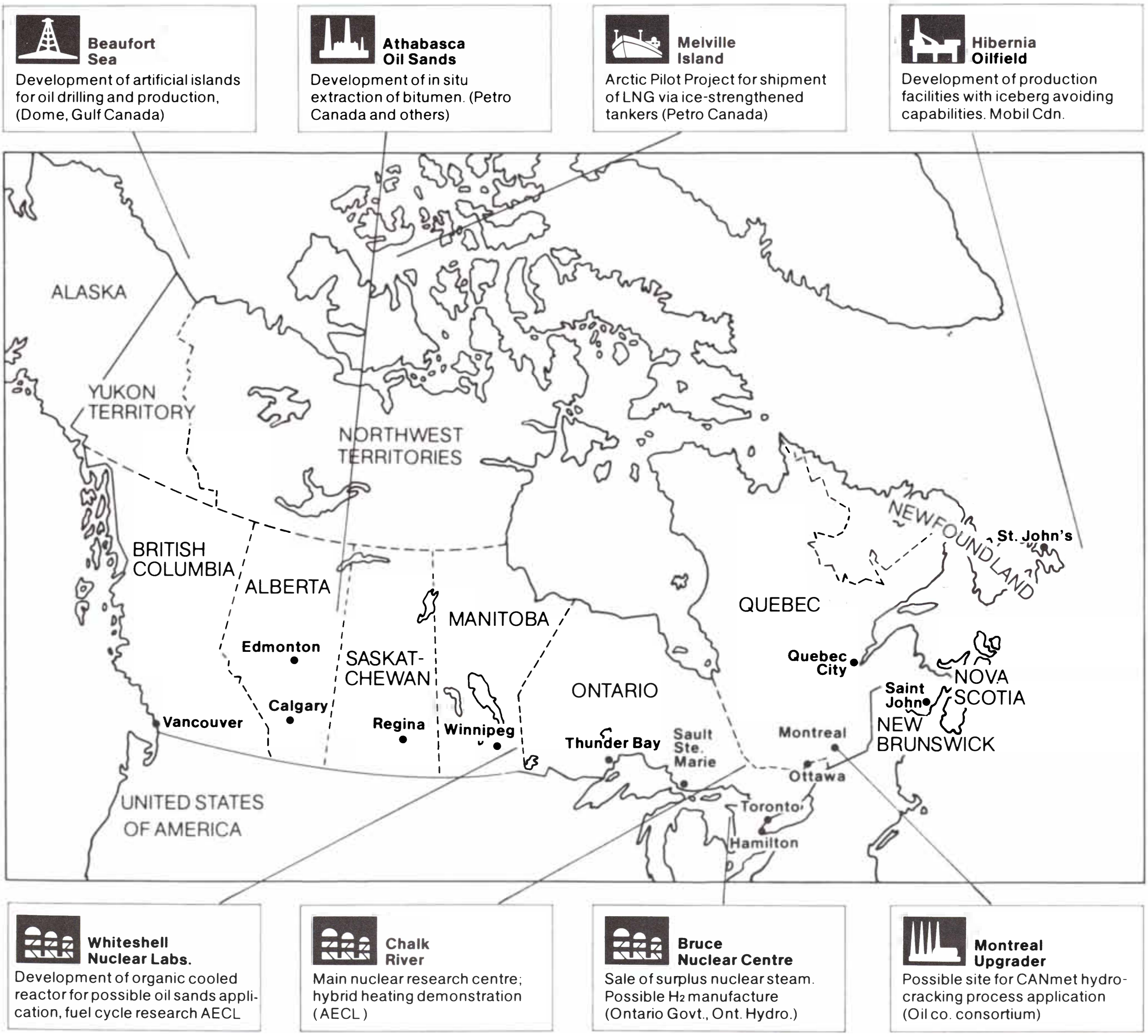
The development of Canada's unique type of nuclear reactor, the CANDU, fuelled by natural uranium and moderated by heavy water, and developed by government-owned Atomic Energy of Canada Ltd., is well known (see Scientific American, Oct. 1975).

The CANDU is now well established domestically, producing some 37% of the electricity generated by Ontario Hydro, the country's largest utility. The type is also gaining international acceptance. CANDUs in Argentina and South Korea are ready to go on power shortly, while the Peoples Republic of Romania has completed commerical agreements for the first two in a planned series of reactors.

The limits of the Canadian market for large nuclear stations have impelled AECL to continue to stress reactor exports, in order to maintain a viable nuclear industry in Canada. This, in turn, has forced AECL to become expert in the process of technology transfer; a topic which, it is now recognized, has subtle but powerful social and political ramifications.

Responsibility for technology transfer has been given to the Atomic Energy of Canada Research Company (AECRC), one of five units into which the corporation has been recently subdivided. Technology transfer is an important factor in domestic sales also, points out Robert Hart, Executive Vice President, Research for AECL and head of AECRC, *"Technology has to be transferred to each new utility as it enters the nuclear field, and to additional consultant, manufacturing, and university person-*

Canada:
Sites of Energy Technology Development




TELIDON
Infomart
The World's Leading Supplier of
Telidon Systems and Services
TORONTO CALGARY WINNIPEG OTTAWA STAMFORD MUNICH

nel in the other provinces." (Quebec and New Brunswick have CANDUs under construction – none of the other seven provinces has yet undertaken a nuclear program.) AECL's export philosophy similarly assumes that countries importing CANDU technology will wish to develop an industrial structure of their own similar to Canada's, where manufacture of nuclear components is a billion dollar a year industry, employing 30,000 people directly and 50,000 indirectly.

The activity of the Research Company, while based on support of the existing CANDU-PHW (Pressurized Heavy Water) system, and further development of that reactor type, also embraces the study of new potential applications of nuclear energy and isotopes.

Advanced Fuel Cycles

Specifically in the nuclear area, AECRC researchers are continuing to investigate advanced fuel cycles. *"Using the natural uranium cycle,"* Hart points out, *"it is possible to recover only about one percent of the energy potential available. Advanced fuel cycles, using thorium in conjunction with either the plutonium produced in current CANDU reactors, or enriched uranium, will, however, allow energy recovery to be increased many-fold.*

These advanced cycles make the cost of nuclear energy less sensitive to variations in the price of uranium and thorium, thus promising a secure supply of low cost energy for centuries. The Canadian approach is extremely cost-effective since it requires only the development of a new fuel cycle, while the alternative approach, the fast breeder, requires the development of a new fuel cycle and a new reactor."

'Slowpoke', 'Powershare' & Electrolytic Hydrogen

The most rapidly growing program of AECRC, which promises the most immediate impact on the day-to-day life of Canadians, is simply described as *"New Applications."* The goal of this work, says Hart, *"is to assess the economics and status of other energy options so that AECL remains up-to-date on the competitive positions of its current technology and is alert to new applications for it. The major thrust at this time is to identify areas where nuclear energy can displace oil and to develop the appropriate technology."*

The most direct approach to this has been to evaluate the feasibility of *"hybrid heating"*– using reactor-supplied, base-load electricity for space heating with fossil fuels being used for peaking purposes. Last winter the town of Deep River, the dormitory community for the Chalk River laboratories, was the site of a major hybrid heating evaluation. Forty Deep River houses with oil/forced-air heating systems were equipped with electric heaters installed in the hot air plenum above the oil furnace. The heaters ranged from 2.5 to 15 kW in power, the latter being equipped with a cutoff to avoid overloading the standard 100 ampere service when other electrical appliances were in use. A two stage thermostat cut in the oil burner only in very cold weather when electric heat was insufficient. The results of the winter-long test showed the 2.5 kW heater replaced 45.5% of normal heating oil requirements, the 5.0 kW heater 75.8%, and the 15 kW interruptable heater 85.9%.

Ontario Hydro is enthusiastically supporting the hybrid heating concept. Both Ontario's Board of Industrial Leadership & Development (BILD) and the federal government provide funds for total or partial residential/commercial conversion to electric heating.

As explained by Chief Economist David Drinkwalter of Ontario Hydro, hybrid heat-

Advertisement

ing, if extensively applied, would be of benefit to the oil industry, the utility, and AECL. Fuel oil suppliers could at least retain a toe-hold in their traditional market area. The use of hybrid electric heaters would provide a large market for interruptable bulk power, without the steep swings in demand that all-electric heating would impose. And such a steady base load would advance the program for the deployment of nuclear reactors in Ontario.

For the more frigid areas of the Canadian north, where remoteness entails a very high cost of bulk fuels, AECRC is proposing a different solution: direct nuclear district heating, using a small supersafe pool-type nuclear reactor called *"Slowpoke."* Originally developed for research institutions and neutron activation work, the low excess reactivity and negative temperature coefficient for the little reactor, whose core is only 22 cm. in diameter and 22.6 cm. in height, render it so inherently safe that no full-time operator is required. One proposal for heating of a remote military base in the far north is estimated to pay for itself in fuel savings within two years.

The Ontario Government, as well, is seeking to market nuclear energy in new forms. A new Crown owned corporation, Ontario Powershare Ltd. has been set up to market nuclear-raised process steam. Ontario Hydro has contracted to sell 250,000 lbs. per hour of nuclear steam at the boundary of its Bruce Nuclear Power Development at prices rising from $1.90 to $2.76 per million BTU in the 1982-1987 period. Even with line costs and infrastructure rolled in, Powershare expects to be able to sell steam at $3.50–$4.00 per million BTUs by 1983, about half the output cost of a gas or fossil-fuelled steam plant. The Domtar Chemicals Group is presently considering locating a $50 million chemical plant near the site, which would use about 100,000 lbs. per hour of the nuclear steam.

Inexpensive base-load nuclear power is also seen as the key ingredient for production of hydrogen, the so-called *"ideal fuel."* A recent report by the government-sponsored Ontario Hydrogen Energy Task Force suggests that electrolytic hydrogen could be competitive, first, in replacing hydrogen from natural gas in refineries and industrial plants, and second, as a transportation fuel, within a ten to twenty year span. The Ontario Ministry of Energy has awarded a $6.2 million contract to the Urban Transportation Development Corporation for development of hydrogen storage and fuel systems.

AECRC is also investigating the more distant prospect of using CANDU reactors to provide the large blocks of power which will be required for recovery of bitumen from oil sands in western Canada.

Petroleum & Permafrost

AECL is in partnership with another federal Crown corporation, Petro-Canada, which is the national petroleum company, now taking a major role in developing new technology for the recovery of *"unconventional oils"* – hydrocarbons too heavy to be produced from conventional oil wells.

To date, Petro-Canada's largest project for the exploitation of frontier resources is the Arctic Pilot Project, a one billion dollar venture. Principal areas of development are the designs of a barge-mounted gas liquification plant and a storage barge which can be floated into the eastern Arctic and then grounded, and design of an LNG carrier which will be able to negotiate the Arctic ice pack at least ten months of the year. Upon completion, the project is expected to deliver 400 million cubic feet of gas per day to the east coast of Canada.

Petro-Canada's role in development of new technology in the recovery of unconventional oils is aimed at the recovery of bitumen in situ from the 90% of the Athabasca and other oil sands deposits which are too deeply buried to be exploitable by the surface mining techniques employed by Suncor and Syncrude Canada Ltd., the owners of the two existing oil sands plants. It is carried out in cooperation with the Alberta Oil Sands

Technology and Research Authority, a provincial body that coordinates oil sands R & D.

Canada's total recoverable oil is placed at 275 billion barrels, making its reserves more than three times larger than those of Saudi Arabia.

Hydro Power & Biomass

In addition to immense reserves of uranium and petroleum, Canada has within its boundaries roughly a third of the world's fresh-water supply, much of it laced across the rolling topography of the precambrian shield.

Only within the last decade has the James Bay watershed in the Province of Quebec begun to be harnessed for hydro electric power generation. The project is enormous and still far from complete. Work has been underway since 1972 building several power plants and reservoirs on La Grande Riviere, the most important tributary of James Bay. Late in 1981, the largest fresh water lake in Quebec was born – the Caniapiscu Reservoir, a five-year undertaking in a remote northern environment. It has been the work of 20,000 men and women, deployed over the largest jobsite in the world. Work will continue on Phase I of this project until 1985.

To minimize energy losses, the power generated by this complex will be distributed to Quebec consumer centres in the south through six 735 kV lines, in three separate corridors. IREQ (Hydro-Québec's research institute), has facilities in its high-power laboratory for studying present-day systems rated at 800 kV and future systems rated up to 1500 kV.

According to Dr. Lionel Boulet, Director of the institute, *"IREQ's research, development, demonstration and test programs are in keeping with a pattern of technological development designed to carry Hydro-Québec into the 1990's. These programs cover four main fields: transmission and distribution systems, including substations and apparatus; power system analysis and automation; materials research; and finally, electricity generation from new energy sources."*

"The main aim of the research into new energy sources is to develop reliable and economical techniques for generating electricity from the wind or biomass or by nuclear fusion. Work also continues on heat production using heat pumps and solar energy."

IREQ has also entered into an *"action concertee"* with Noranda and Electrolyzer Corporation in the joint operation of a plant to produce hydrogen by water electrolysis. (In the *"action concertee"* provision is made for effective intersectoral cooperation among universities, government laboratories and private industry.) The use of a recently developed separator material in an electrolytic cell shows considerable promise. A planned 100-MW plant will permit the production of hydrogen (and – as a by-product – heavy water) at a price which will become more attractive as the price of natural gas goes up.

In a related development, Noranda is negotiating a joint venture for the commercial gasification of bio-mass, mainly peat and woodwaste; methanol is the final product.

According to Alan Pope, Ontario Minister of Natural Resources, *"Ontario possesses the energy equivalent of 72 billion barrels of oil in its 24 million acres of peat land south of the permafrost region, which could be developed with current technology."*

FROM BELL TO BILL S-898

No matter how vast the land mass, Canada's natural resources are ultimately finite. The nation's future prosperity is clearly a function of its human resources, the determined harnessing of intellectual energies to meet the challenge of emerging technologies. In the fields of telecommunications and computer processing, Canadians are already breaking new ground.

Canada's great spaces, thinly scattered with centres of population, created a national need for high-quality telecommunications as old as the country. Alexander Graham Bell, it may be remembered, after leaving Scotland, spent much of his life in Canada and his *"telephone"* was as much a Canadian as a U.S. invention.

Nevertheless, in spite of the growth of the Bell Canada system, which now dominates telephone service in eastern Canada, the country depended upon U.S.-generated communications technology until a decade ago. *"Ten years ago Northern Telecom (then Bell Canada's wholly owned manufacturing subsidiary) was basically manufacturing products or systems designed by Western Electric or Bell Labs,"* says Walter Light, president of Northern Telecom Ltd. (NT) which is now, after two public stock issues, 55% owned by Bell Canada.

In 1971 NT's Ottawa research division became Bell-Northern Research Ltd., now 70% owned by NT and 30% owned by Bell Canada. BNR had more than a new name: it had a mandate to provide technology, product design, and long range planning for its two parents. The change began a decade of rapid growth for NT, now acknowledged as Canada's leading high-technology company. Its estimated $2.5 billion sales this year are projected to grow to $5 billion annually by 1985. Its 51 plants staffed by 35,000 employees include operations in Britain, Ireland, Brazil, Malaysia, and Turkey.

But the Company's major presence outside Canada has been in the U.S.A.. Starting with its first U.S. plant in 1972, Northern Telecom Inc. now has 12,000 employees in 16 plants in 13 states. Part of its growth is the result of penetration of the telecommunications market. For instance, NT Inc. is the major supplier of digital telecommunications equipment to such massive customers as the United States Department of Defense. The rest has been due to the acquisition of several high-technology American companies such as Sycor and Data 100.

Northern Telecom's sales have been growing faster in the U.S.A. than in Canada. By 1985, the company expects U.S. sales to be generating 60% of its total revenue, vs. the present 40%.

Part of this expansion is due to the opening up of the telecommunications market in North America, which started with the Carterfone decision in 1968, and is continuing on both sides of the border, with last year's decision by the Canadian Radio Television and Telecommunications Commission to allow interconnect, and this year's comprehensive U.S. telecommunications Bill S-898 before the Senate. A major role must also be accorded to NT's aggressive marketing to independent telephone companies.

But NT couldn't have exploited its market openings without an edge in technology: specifically, it has been able to achieve a world lead in the transition from analog to digital switching and transmission in telecommunications.

World Recognition

As evidence of world recognition of this lead, NT last year received the International Industrial Award of the Institut International de Promotion et de Prestige, of Geneva. In presenting it, M. Louis Terrenoire, former Information Minister of France, said *"Northern Telecom foresaw the convergence of computers and telecommunications based on the common technology of integrated circuits and software control. Because of its exceptional awareness of the evolution towards decentralization through increased use of remote data processing methods and communication between computers, Northern Telecom has become a leader among those who design, manufac-*

ture, and market integrated information networks. . ."

NT was the first Canadian company, and the second in North America, to receive this award.

To meet the challenge of designing complex intricate circuits with high reliability, BNR developed a computer-based design system called *"COPES"* which arranges individual components and large scale integrated circuits (LSIs) on printed circuit boards. COPES was further developed to design the LSIs themselves. These LSIs (by convention, groupings of over 250 separate components such as resistors and capacitors) are placed on paper-thin silicon chips only six mm. square.

One challenge for BNR's LSI designer was to produce a filter codec – an essential element in the *"line card"* connecting the subscriber to the network – on a single chip. The filter codec translates (or codes and decodes) the analog signals created by the telephone speaker into digital bits – and back at the receiver end of the circuit. It also filters the resulting signals for accurate switching and transmission. It incorporates an incredibly intricate array of circuitry (256 binary unit capacitors are only one small element in its design) on a standard-size silicon chip.

So quietly did NT introduce the new chip that 18 months later the giant Japanese electronics corporation, Hitachi, proudly announced that it had achieved a first in producing a single chip filter codec. By now NT has produced over two million filter codec chips – and is phasing down production; the filter codec is obsolescent. In September, NT announced that a successor to the filter codec has been developed: the E-99 line-card chip. This integrates the filter codec with the *"Access"* chip, which passes the voice signal to the codec, and the supervision circuit, which monitors the busy or idle state of the set, controls the ringing function, and provides automatic impedance matching and gain switching. As a final touch, it conserves energy by reducing power for its own internal circuits when it is not in operation.

Even more complex chips are on the BNR drawing board: VLSIs (very large integrated circuits) containing as many as 30,000 to 40,000 components, whose manufacture will require another technological leap.

BNR, the company's research arm, and the manufacturing market organization of NT *"are now so closely integrated it is difficult to separate them."* BNR is tightly managed, on a mission-oriented, case-funded basis for each development project, but at the same time it gets about 10% extra from NT as a *"capability fund,"* which allows R & D management to pursue more *"blue sky"* projects. The concept of the filter codec chip itself started out as a *"capability fund"* project.

Saskatchewan Fibre Optics

With the basic switchover to digital technology completed, the company's development teams are now pushing ahead with the next big revolution: optical fibre transmission. In 1980, NT won a $22 million contract from the government of Saskatchewan to provide voice, data, and cable TV services. Now being plowed in, it is the longest optical fibre cable system in the world. Optoelectronics offers greater information transmission capacity at potentially lower cost than comparable copper core cables; the Saskatchewan system uses cable that has 12 glass fibre strands which have a transmission speed of 45 million bits of information per second. Forecasters predict the optoelectronic market will grow from under $100 million in 1979 to about $2 billion by 1990.

While NT executives feel comfortable about their present leadership in technology, they are dogged by one nagging worry: that Canada's Restrictive Trade Practices Commission might force Bell Canada to divest itself of its 55% interest in NT's shares. The result, ironically, could be to transform NT into a U.S. company, since 70% of the other 45% of the company's shares are U.S. owned. According to President Walter Light, *"if you divest Northern Telecom from Bell, you could destroy a winner. In the long run, the electronics industry in Canada would go the way of the textile industry."*

"The impact of separation wouldn't be restricted to Northern Telecom," points out BNR Executive Vice-President, John Roth. *"In the Ottawa area, there are currently some 80 high technology companies set up around BNR. Some of these companies were started by ex-BNR or Northern Telecom employees . . . they have taken with them the skills they have learned, the good ideas, and started other companies. These ideas have moved into the rest of Canadian industry."*

Telidon

It is in many ways appropriate that Canada, where Marshall McLuhan developed his theories of communication, should have been the originator of *"Telidon,"* a two-way television communication system.

In papers published as early as 1965, Douglas Parkhill of Canada's Department of Communications foresaw the uses for such a system. Thirteen years later, in Ottawa in 1978, H.G. Bown, Director General of Information Technology for the Canadian government, demonstrated the use of a new digital coding language devised by the Federal Communications Research Centre which, combined with the *"View-data"* approach invented by Samuel Fedida in England, provided information via television in discrete *"pages."* Together the two innovations created the *"Telidon"* videotex system.

According to Bown, now vice-president, corporate development of Norpak Ltd. of Kanata, Ontario, four basic design principles guided the evolution of Telidon.

"Because advances are continuously being made in computer and communications technology, a communication system should be designed, as far as possible, in a manner independent of the technologies involved."

"Because the creation and maintenance of information stored in databases, application programs, and other source material comprise an investment far in excess of all the hardware costs in a system, the overall system design must accommodate the changing nature of the communications delivery and display presentation media without obsoleting older display terminals or database coding methods."

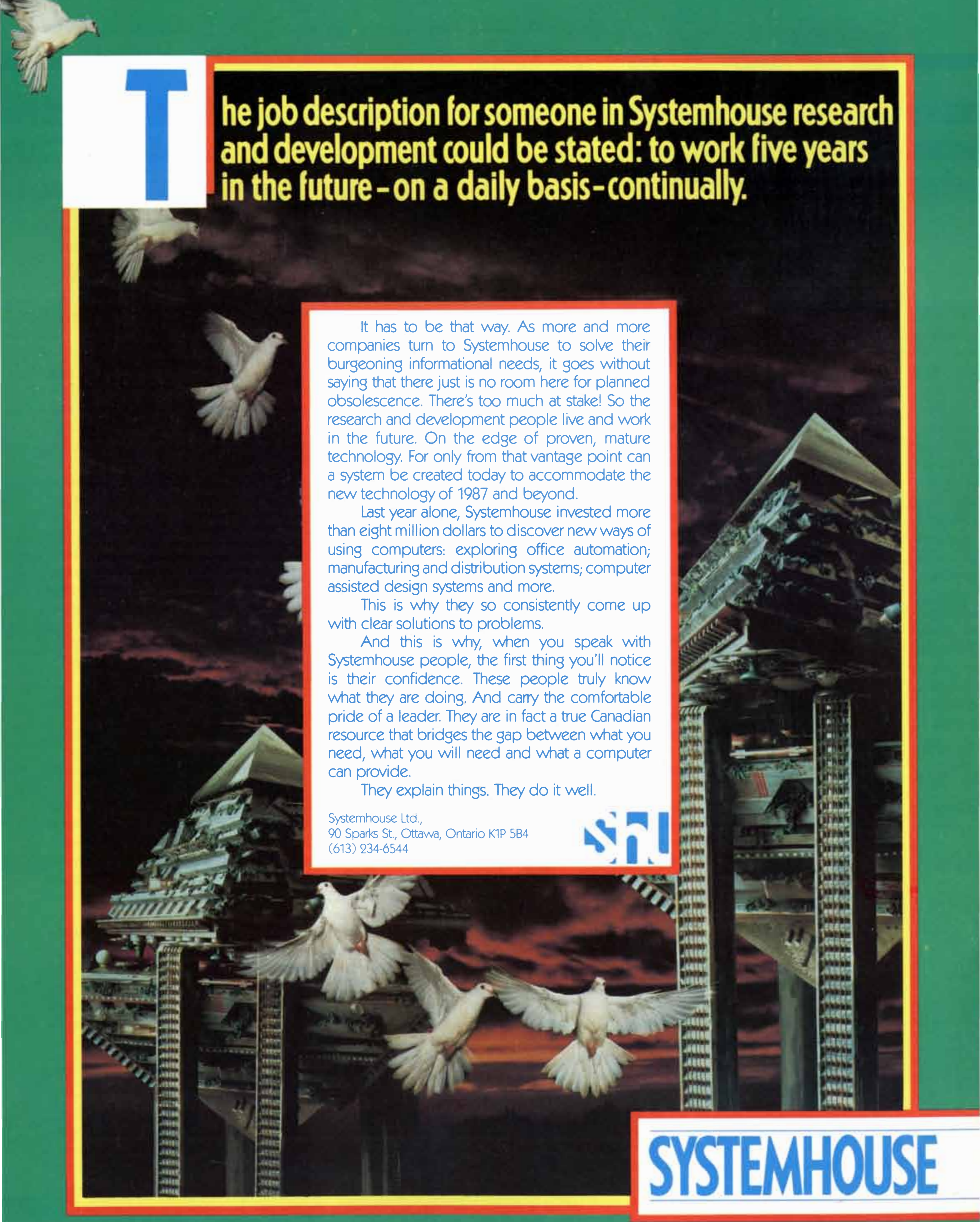
"Since the cost of computing hardware continues to drop faster than the cost of communications delivery systems one should maximize the amount of computing power in a terminal in order to reduce the amount of data which must be communicated."

"Fully interactive capability, including person-to-person, was deemed essential and such applications as information retrieval and selective sequential information broadcast were to be subsets of this overall requirement."

What evolved from these design principles was a flexible alpha-geometric coding scheme, in which visual objects are described in terms of abstract geometric parameters, such as line, arc, and polygon. This scheme provided the means for reducing a 300,000 byte photograph to a 300 byte coded message, which could be transmitted over a narrow bandwidth channel. The displayed objects can be translated, scaled, and even rotated without recoding.

The *"information provider"* in Telidon is a specialized terminal, either self-contained or backed up by a central computer for information storage. For high-volume applications, a stand-alone terminal justifies the cost of its built-in storage and retrieval circuits and programming, while for low-volume applications, the cost of a central computer facility is justified. (By next year some standard colour TV sets, built by Electrohome Ltd. and Panasonic in Canada, will be capable of receiving Telidon signals.)

At present, Norpak Ltd. and Northern Telecom both provide stand-alone, desk-type, self-contained terminals, at a cost in the $20,000 range. Another Ottawa firm, Hemton Ltd., which recently merged with


The job description for someone in Systemhouse research and development could be stated: to work five years in the future–on a daily basis–continually.
It has to be that way. As more and more companies turn to Systemhouse to solve their burgeoning informational needs, it goes without saying that there just is no room here for planned obsolescence. There's too much at stake! So the research and development people live and work in the future. On the edge of proven, mature technology. For only from that vantage point can a system be created today to accommodate the new technology of 1987 and beyond.
Last year alone, Systemhouse invested more than eight million dollars to discover new ways of using computers: exploring office automation; manufacturing and distribution systems; computer assisted design systems and more.
This is why they so consistently come up with clear solutions to problems.
And this is why, when you speak with Systemhouse people, the first thing you'll notice is their confidence. These people truly know what they are doing. And carry the comfortable pride of a leader. They are in fact a true Canadian resource that bridges the gap between what you need, what you will need and what a computer can provide.
They explain things. They do it well.
Systemhouse Ltd.,
90 Sparks St., Ottawa, Ontario K1P 5B4
(613) 234-6544
SYSTEMHOUSE

Norpak, has been marketing the EPS-1 projection system for the audio-visual industry. It has built-in memory for about 100 pages of data which can be loaded from an audio cassette recorder or telephone line, which can augment or replace the conventional slide or viewgraph presentation.

By next year, Norpak will be bringing out a self-contained local terminal costing only a quarter of present prices, and a Telidon decoder fed from a remote computer which will cost only $2,000.

David M. Carlisle, president of Infomart, a Toronto-based "electronic publishing" company using Telidon, anticipates the same dilemma facing Telidon developers as faced early radio and T.V. pioneers: how do you establish a brand-new medium? Which comes first, transmitting stations or receiving equipment?

Infomart's strategy is to address one market segment at a time in order to gain sufficient market acceptance to attract information providers.

The first commercial Telidon network in the world has been started in Manitoba, under the tradename *"Grassroots"* to serve the agricultural community with specialized information such as market trends, commodity prices, government programs, and weather forecasts, at a cost to the user of $49 per month. It is distributed by cable TV to its present 250 subscribers by a joint undertaking of Infomart and the Manitoba Telephone System. Manitoba Tel predicts the sale of a total of 2,500 terminals by the end of this year. A similar network under the same name is now being installed in the San Joaquin Valley of California, run jointly by Infomart and the *Bakersfield Californian* newspaper.

In another move to take "Telidon out of the lab and into offices and homes," Ontario Minister of Industry and Tourism Larry Grossman announced that two thousand Telidon terminals would be installed in libraries, public buildings, and hotel lobbies in Metropolitan Toronto, supplied via central computer with a 50,000 page information bank of data on tourism, restaurants, entertainment, and shopping, for use by the general public. *"We estimate that this program will stimulate the sale of 75,000 videotex terminals in our domestic market alone by 1983."* says Grossman.

Distribution agreements for Telidon terminals and information systems are in final stages with Siemens AG in West Germany and a system has been sold to the government of Venezuela.

In the near future, various presentation and communication enhancements to the Telidon system are anticipated. The earliest development will likely be the forwarding of simple textual communications between users, stored in a *"mailbox"* facility in the host computer. This will be followed by direct terminal-to-terminal textual communications, and by sharing simultaneous transmission of visuals from the host computer.

The ultimate development will be the design of a *"Common Visual Space,"* permitting a group of terminals to communicate and modify visuals between themselves. A business conference on new product design, for instance, could be conducted by a Telidon conference, with an image of the package being manipulated on the screen during the audio exchange between members of the marketing team, who might be at opposite sides of the continent.

In May of last year, AT & T announced its specifications for presentation level protocol (PLP). Telidon manufactuers, confident that they have come closest to meeting these norms, are now awaiting Canada's own standards, Standard 709, from the Department of Communications. *"Telidon is already a second generation videotex system,"* according to Cameron Miller, director of the federal Telidon Marketing Secretariat, *"because of its alpha-geometric coding scheme. Further enhancements will have more to do with the marketplace than technology."*

On Computing

The development of Telidon owes as much to Canada's computer industry as it does to telecommunications.

Almost from the beginning of the computer era, Canadian electronics manufacturers have been involved in the technology of electronic data processing, although none has become a major mainframe supplier. According to Hugh O'Rourke, director of research for Evans Research Corp., Toronto financial analysts specializing in EDP, *"the real hope for the rest of the decade and beyond – at least based on what appears to have worked up to the present – is specialization. In both services and hardware, nine out of ten firms that increased market share in 1979 have an identifiable specialty. Conversely, nine of ten firms that had a decreased market share are either generalists or have multiple specialties."*

While the computer industry in Canada is large, most of the manufacturing is done by subsidiaries of the major U.S. computer manufacturers. Of the 15 top firms, only three are Canadian owned, and one other (Philips Electronics) is controlled by Netherlands interests. Further, of the three Canadian firms, only one, AES Ltd., is a manufacturer, specializing in word processing systems. The other two, Canada Systems Group and Datacrown Ltd., are computer service suppliers.

O'Rourke sees the revolution in microprocessor chips, however, now making possible a Canadian hardware industry. *"I can see all sorts of microprocessor-based equipment, including intelligent terminals, that Canadian firms could manufacture,"* he says.

A number of Canadian companies have established themselves in the hardware business; though, as O'Rourke points out, the most successful have identifiable specialties. A company that has capitalized on terminals is Cybernex Ltd., which makes video display terminals and components such as keyboards and peripheral equipment.

Cybernex sells most of its terminals to large communications organizations, such as Bell Canada, Xerox and various government departments. Emphasizing quality, dependability, and easy to service products, the company is aggressively pursuing the custom terminal market.

Cybernex has designed the LC3 micro computer to function both as a compact development machine and the target machine. The Cymon Operating System and Cybol High Level Language have been developed by Cybernex and are optimized for the LC3. The LC3 is based on a Motorola 6809 processor. This single board computer is suited for use as a controller or a multiplexer as well as a dedicated application computer. The objective of the design is to give more programming power.

Cybernex originated in 1974, when James Gadzala and Colin Turner developed a digitizing video terminal to act as a bridge between aerial photographic data and a computerized system for map production while working for the federal government. Gadzala and Turner teamed up with two businessmen, Bruce Douglas and David Londry, to start Cybernex, *"with precious little money."* It has doubled its sales to $5 million this year, and expects an increase to $8 million with its entry into the U.S. market.

Software specialists have become especially successful. In a field with considerable competition, Systemhouse Ltd. claims to be *"Canada's leading computer systems consulting organization."* It had revenues of $99 million in 1981 with projections to $250 million by 1983-84. Export sales to customers in the U.S. and elsewhere accounted for 25% of the company's sales in 1980-81.

While Systemhouse continues to stress its service capability through its Services Division it is committed to its own product development. According to J.R. Davies, President and Chief Operating Officer of Systemhouse, *"Nineteen-eighty-one saw the formalization of our Advanced Technology research program, the opening of a research and development centre in Ottawa, the acquisition* *of quality software products including the GMS-3000 manufacturing system, the TOM and CDS systems for financial management in the distribution marketplace and Holguin's CEADS/CADD system. We have also completed the development of our Hospital Financial Management System and our DEC-based business system as well as the refinement of our computerized mapping system to produce Automap II. These major steps have been taken while continuing to expand our professional services division at a 50% per year rate."*

In 1981, the computer industry in Canada earned about $3.3 billion (Cdn.) in gross revenue, and industry analysts expect this will grow to $6 billion by 1990. The great bulk of this goes to hardware manufacturers: some 65%, with 25% spent on software and 10% on other services.

In Conclusion

While the technologies for energy generation and electronics development figure prominently in Canada's future, they are by no means the only examples of Canadian technological excellence. Canada was the third country in the world after the U.S.S.R. and the U.S. to launch an earth satellite. Having a population of approximately twenty-four million, occupying the second largest land mass in the world, Canada remains committed to aerospace R & D.

Much of the R & D throughout Canada is linked to its geography. Robert Long of the Canadian Advanced Technology Association would like to see further emphasis on these distinct technological needs and suggests that in the long run their solutions will provide the greatest return to the country.

"To attempt to solve strategic Canadian problems using foreign technological solutions will reduce the effectiveness of Canadian managers and the competitive position of Canadian business to the lowest common denominator of the foreign technology. How can Canadians, for example, effectively manage the resources of a northern latitude country with American visible light satellite technology (Landsat) when, as a northern latitude country, most of our landmass and coastline is in darkness or fog covered for so much of the time? Canadian synthetic aperture radar aboard a Canadian surveillance satellite would not only allow us to manage our Arctic with its oil tanker traffic but the chances are that we could sell realtime data to the United States for use in Alaska."

Long sums up Canada's relative world position this way, *"We have a highly educated labour force; we have one of the highest savings rates in the world – hence the potential for a self-sufficient pool of risk capital, and we are the envy of the world in natural resources. Canadian companies have a beachhead with the best in the world in a host of critical technologies ranging from telecommunications to nucleotide (DNA/RNA) synthesis chemistry and automation."*

Advertisement

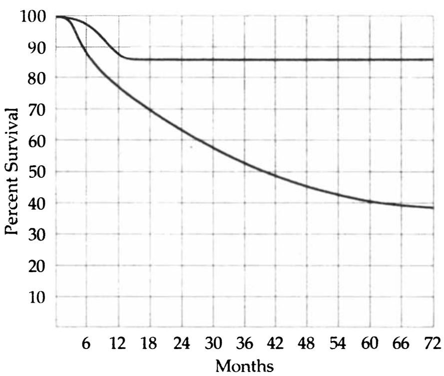
Medicine is always changing, now there's a text that changes right along with it.

SCIENTIFIC AMERICAN *Medicine.*

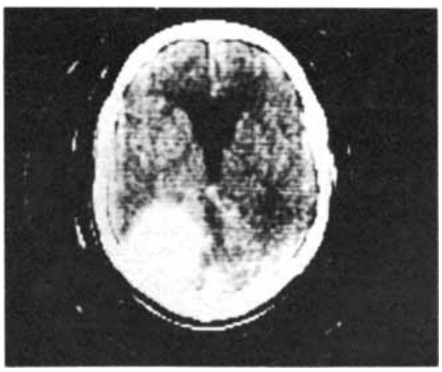
The answers that you get from SCIENTIFIC AMERICAN *Medicine* are current, because each month new chapters replace the old. You will know where to look for the answers, because this advanced text follows the familiar structure of medical knowledge, changing as medicine changes.

Comprehensive. Leading authorities from Harvard and Stanford cover the full range of medicine for you the seasoned clinician.

Complimentary CME. At no extra cost, we offer a CME program through which you can earn 32 top credits* per year.



Top curve represents survival in chronic hepatitis patients treated with steroids.



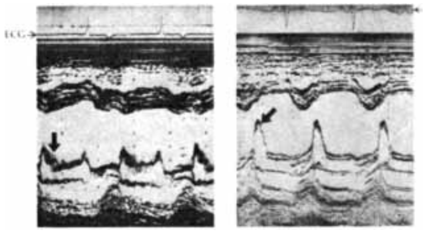
CT scan reveals large meningioma and mild hydrocephalus.

Trial Offer. We invite you to try SCIENTIFIC AMERICAN *Medicine* for two months with our compliments. Send us the coupon and you will receive the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, you will owe us $220 for the first 12 months (current renewal is $170); otherwise, return the volumes; there will be no charge.

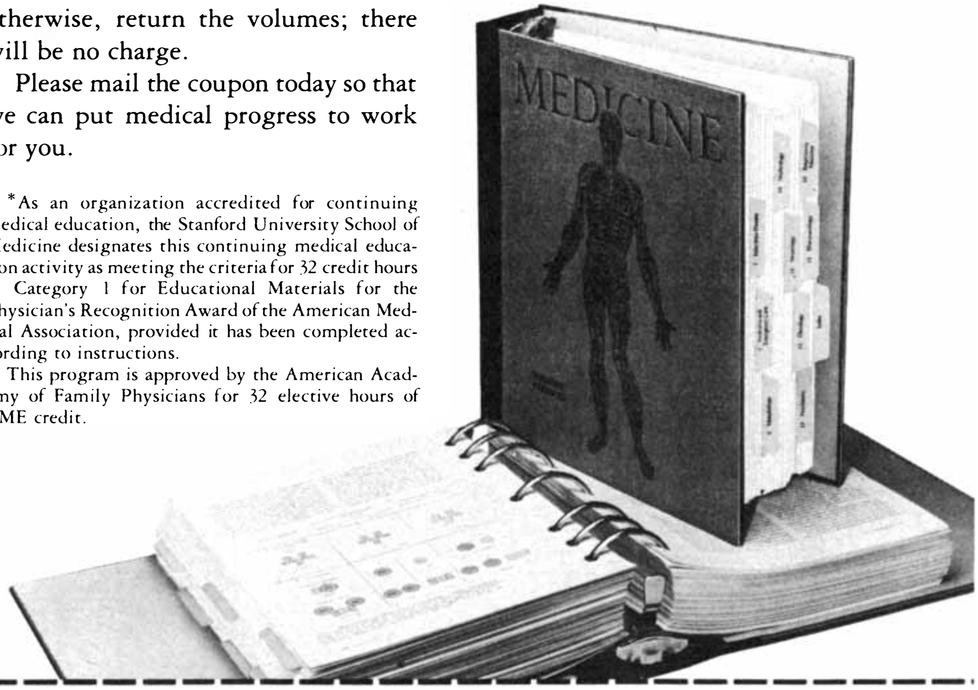
Please mail the coupon today so that we can put medical progress to work for you.

*As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association, provided it has been completed according to instructions.

This program is approved by the American Academy of Family Physicians for 32 elective hours of CME credit.



Echocardiograms from patients with aortic regurgitation.



SCIENTIFIC AMERICAN MEDICINE

415 Madison Avenue, New York, New York 10017

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine.* On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the annual subscription of $220 for SCIENTIFIC AMERICAN *Medicine* is tax deductible. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a *complete refund.*

- ☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*
- ☐ I shall also enroll in the CME Program
- ☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $220*
- ☐ Please bill me.

*Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York

Name ______________________ MD Specialty ______________

Address ____

City ______________________ State ______________ Zip ______________

Signature ____

Subscribers outside the U.S. and possessions will be charged extra for shipping and handling; updates will be sent by surface routes unless airmail delivery is requested. Please allow 6-8 weeks for delivery. M1

Phil and Steve Mahre, World Cup Race at Aspen

You have what it takes to win. Help support the U.S. Ski Team.

The U.S. Ski Team is on a winning streak. They need your help to keep it going.

World Cup victories by Phil Mahre and Tamara McKinney highlighted the 1981 season. The Team's wins provide high hopes for the 1982 World Championships.

Our Alpine and Nordic Teams are not subsidized by the government. These fine athletes devote years of their lives training to win. They depend on you, the American public, to support their efforts. Please send your tax-deductible donations to the U.S. Ski Educational Foundation, Box 100M, Park City, Utah 84060.

Please help.

US SKI TEAM

Titan

The largest moon of Saturn is the only moon in the solar system with a substantial atmosphere. The chemistry of the atmosphere may resemble that of the earth's atmosphere before life arose

by Tobias Owen

On November 12, 1980, the spacecraft *Voyager 1* passed within 7,000 kilometers of Titan, the largest moon of Saturn. It was the closest encounter between *Voyager 1* or *Voyager 2* and any planet or moon, but it was made at a cost. The encounter meant that *Voyager 1* could not be redirected by the gravitation of Saturn so that it would travel onward through the solar system to pass near Uranus and Neptune. The sacrifice seemed appropriate; Titan was known to be the only moon in the solar system that has a substantial atmosphere. Moreover, its reddish color, which is unique among Saturn's moons, suggested that the chemistry of Titan's atmosphere might be producing colored compounds. Because of the closeness of the encounter the instruments on board the spacecraft would be able to perform at their best.

The results of the encounter show how wise the sacrifice was. Titan turns out to be the only known body in the solar system besides the earth whose surface is at least partially covered by a liquid. On Titan the liquid is methane. Moreover, the Voyager instruments showed that the atmosphere of Titan is denser than the atmosphere of the earth. This denser atmosphere has retained conditions much like those that probably existed on all the planets soon after they formed. Specifically, the atmosphere of Titan has carbon, nitrogen and hydrogen but lacks molecular oxygen. Under these conditions the chemical reactions proceeding in Titan's atmosphere today may well be giving rise to some of the organic molecules that are thought to have been precursors to life on the earth.

The Presence of an Atmosphere

Christiaan Huygens discovered Titan in the spring of 1655, the year in which he also proposed that Saturn has rings. The body was named almost two centuries later, when Sir John Herschel assigned names to the seven moons of Saturn that were known at the time. The name Titan was well chosen. Herschel knew only that Titan was the brightest moon of Saturn; since then it has proved to be also the largest. Indeed, it is larger than the planet Mercury. For a time it was thought to be the largest moon in the solar system. Measurements made by *Voyager 1* show that it does not quite hold the record. The earlier measurements had been inflated by the thickness of Titan's atmosphere. The solid body of Titan has a radius of 2,575 kilometers. Jupiter's moon Ganymede is larger: its radius is 2,640 kilometers.

The first hint that Titan has an atmosphere came from observations of the body that the Catalan astronomer José Comas Solá published in 1908. Solá reported that the tiny disk of Titan he could see through his telescope was darker at its limb, or periphery, than it was at its center. The reason, he proposed, was the presence of an atmosphere. Specifically, the sunlight reflected toward the earth by Titan's limb must pass through more of Titan's atmosphere than the sunlight reflected by the center of Titan's disk. Thus the light from the limb is attenuated in greater measure by absorption in Titan's atmosphere.

It is hard to know whether Solá really observed that the limb of Titan was darkened. The descriptions he gave of patchy clouds on the giant moons of Jupiter are known to be mistaken. Nevertheless, his observations seem to have led Sir James Jeans to include Titan and the giant moons of Jupiter in his theoretical study of the escape of atmospheres from the bodies of the solar system. In 1916 Jeans concluded that Titan has probably retained an atmosphere in spite of its small size and weak gravity compared with, say, the earth because of its low temperature. Titan's distance from the sun combined with any reasonable estimate of its reflectivity (and loss of solar heating on that account) leads to the prediction that its surface and its atmosphere should have a temperature of between 60 and 100 degrees Kelvin. For such a range Jeans's work shows that a gaseous substance whose molecular weight is 16 or greater should not have escaped from Titan over the history of the solar system.

Several substances satisfy Jeans's limit on weight. One of them is ammonia (NH_3), whose molecular weight is 17. In the 1930's Rupert Wildt of the University of Göttingen identified it as a component of the atmosphere of Jupiter. Wildt had found that the spectrum of the sunlight reflected from Jupiter in the infrared showed the absorption of the radiation at wavelengths characteristic of ammonia molecules. By a similar method Theodore Dunham, Jr., of the Mount Wilson Observatory detected ammonia on Saturn. At the temperature assumed for Titan, however, ammonia would be a solid; it cannot be a substantial part of an atmosphere. Other substances that satisfy the limit are argon, neon and molecular nitrogen (N_2). All of them would have had an appreciable concentration in the mixture of gases and dust that condensed to form the solar system. The problem is that they are hard to detect spectroscopically. None of them absorbs much radiation in the infrared.

Still another substance that satisfies the limit is methane (CH_4), whose molecular weight is 16. Unlike argon, neon and molecular nitrogen it has a strong set of absorption bands in the infrared, and unlike ammonia it is gaseous at the temperature predicted for Titan. In 1932 Wildt identified methane in spectra of Jupiter, Saturn, Uranus and Neptune. Then in 1944 Gerard P. Kuiper of the University of Chicago identified it in the spectrum of Titan. His discovery constituted the first strong evidence that Titan indeed has an atmosphere. By comparing the spectrum of Titan with laboratory spectra of methane at low pressures Kuiper deduced that the absorption of sunlight by the gas along a vertical path through Titan's atmosphere is equivalent to the absorption of such radiation by a column of methane 200 meters long at a pressure of one earth atmosphere and a temperature of 273 degrees K. (Such conditions are called standard temperature and pressure, or STP.) For the purpose of comparison a vertical column through the earth's atmosphere amounts to a column eight kilometers long at STP.

Some problems with the early understanding of Titan's atmosphere began to materialize in 1965, when Frank J. Low of the University of Arizona deduced from the brightness of the radiation Titan emits at an infrared wavelength of 10 micrometers that the temperature of the body is 165 degrees K., a temperature nearly twice as high as the temperature one would attribute to simple solar heating of Titan's surface and lower atmosphere. Low's finding went largely unnoticed for seven years. Then a number of investigators began to find other surprises. For one thing, the calculations of Titan's temperature based on measurements of its brightness at various infrared wavelengths failed to yield a consistent value. Titan's "brightness" at radio wavelengths also yielded inconsistencies. The radio brightness in one observation implied a surface temperature of 200 degrees K.

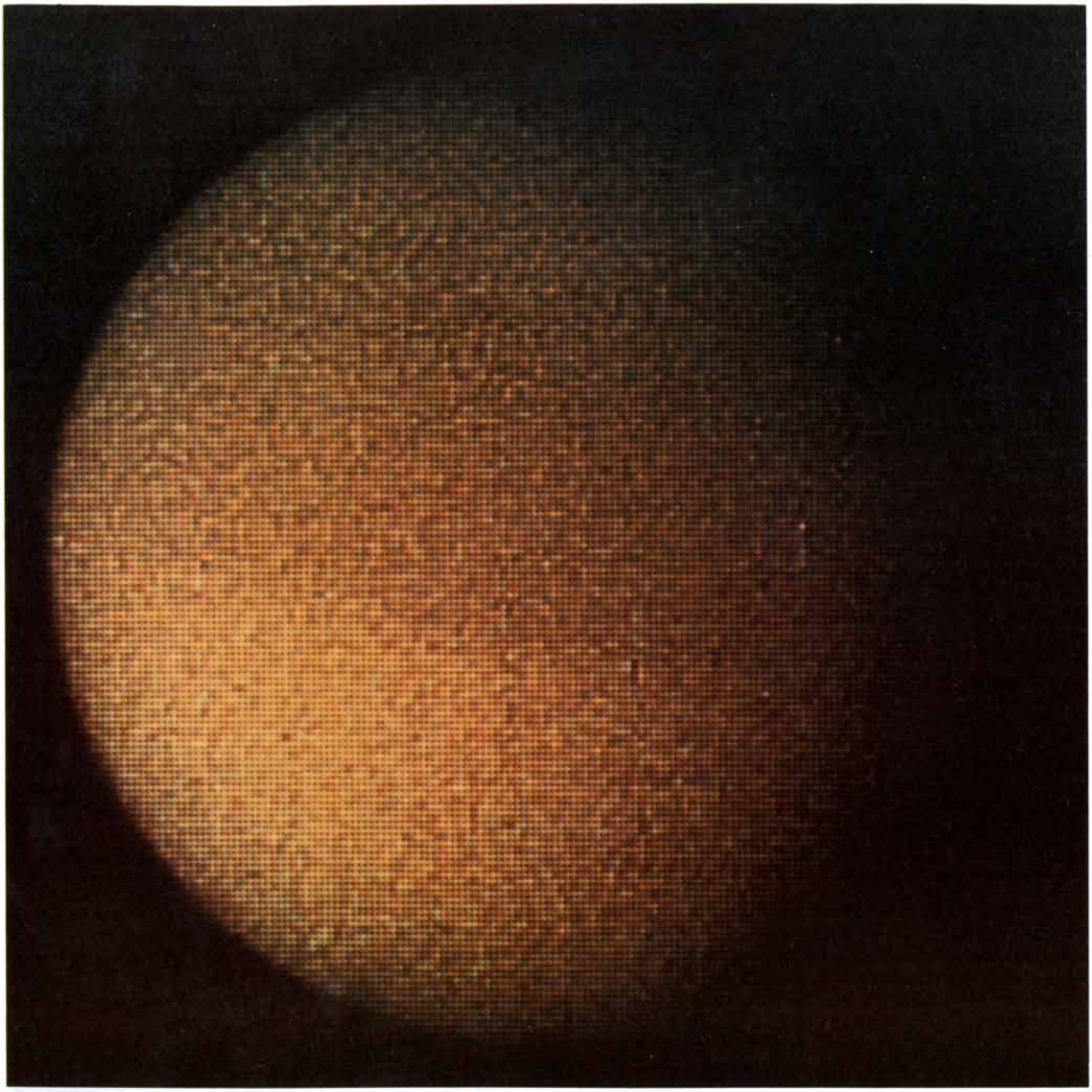
Further still, the light reflected from Titan at the small angles that prevail between the sun, Titan and the earth turned out to have a positive polarization: the vector representing the maximum strength of the electric component of the electromagnetic field of the light was perpendicular to the plane in which the angle lay. This finding, made independently by Joseph F. Veverka of Cornell University and Benjamin H. Zellner of the University of Arizona, suggested that the light reflected from Titan came not from the solid surface of the body through a transparent atmosphere but from a deep, cloud-filled atmosphere. Mars, for example, has only a thin atmosphere, and the light it reflects toward the earth at small angles shows negative polarization: the electric-vec-

ATMOSPHERE OF TITAN is apparent in an image of the night side of the body made last August 25 by the spacecraft *Voyager 2.* The orange crescent forming the left part of Titan's limb represents the reflection of sunlight by an aerosol: a layer of particles suspended in Titan's atmosphere some 200 kilometers above the surface. The blue halo surrounding both the crescent and the unlit part of Titan's limb represents the scattering of sunlight at large angles by particles of haze suspended as high as 300 kilometers above the surface.

tor maximum is in the plane in which the angle lies.

Meanwhile Laurence M. Trafton of the University of Texas at Austin had surmised that the amount of methane on Titan must be considerably greater than the amount deduced by Kuiper or else some other gas must also be present in quantity. With the aid of an infrared image intensifier that had just become available Trafton had found that a methane absorption band near a wavelength of one micrometer in the spectrum of Titan was unexpectedly strong. The strength of the band could be due to methane in surprising abundance. Alternatively the collisions between methane molecules and molecules of an undetected gas could be perturbing the vibrational states of the methane molecules, so that the methane would be absorbing infrared radiation over a broad range of wavelengths. In either case the strength of the absorption band was related to both the abundance of the methane and the local atmospheric pressure.

Trafton further discovered that the stronger methane absorption bands in Titan's infrared spectrum have an appearance different from that of the same bands in spectra of Jupiter and Saturn. Titan's bands are both shallower (less intense) and broader. This finding suggested to Trafton that small particles are suspended in Titan's atmosphere in numbers and at altitudes different from those of the particles composing the hazes observed on Jupiter and Saturn. The particles scatter sunlight toward the earth; thus they brighten the absorption bands. They also scatter sunlight at angles in Titan's atmosphere. The in-



SUNLIT SIDE OF TITAN was photographed by *Voyager 1* on November 9, 1980, three days before the spacecraft passed within 7,000 kilometers of the body. In the resulting image the surface of Titan is concealed by the unbroken opacity of the aerosol layer. The southern hemisphere is brighter than the northern hemisphere, perhaps in response to a seasonal change in the rate at which aerosol particles are produced. It is spring in the northern hemisphere; thus the southern hemisphere has just passed through a summer seven years long.

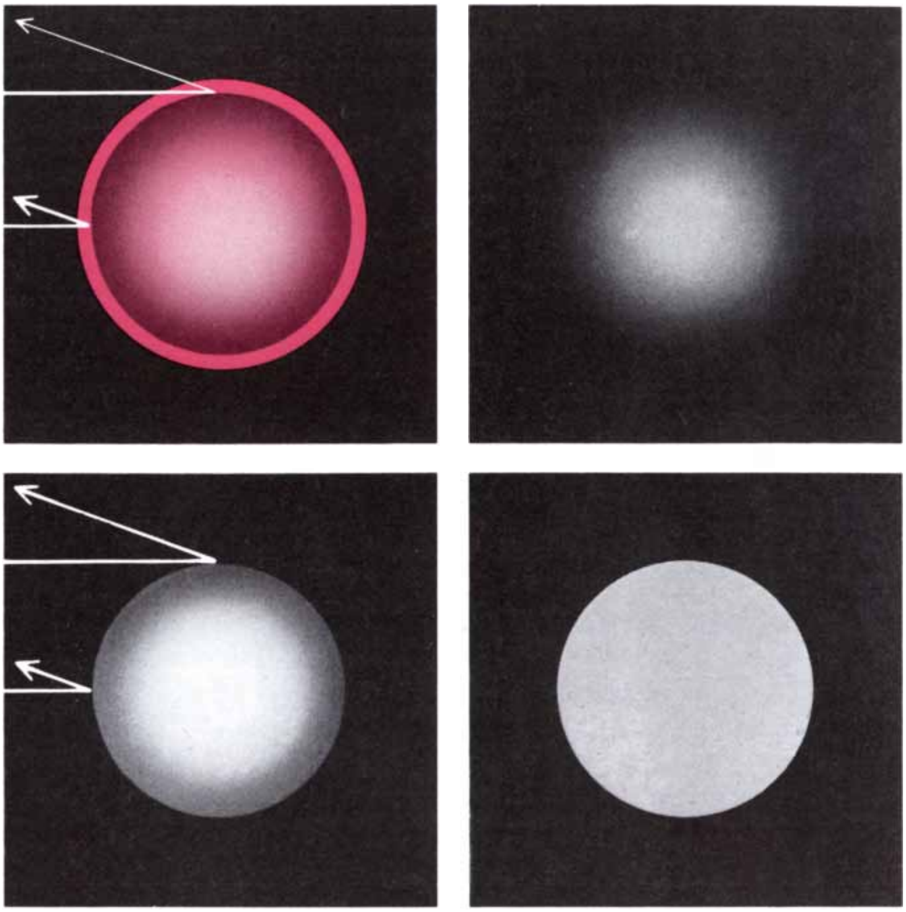
creased path length of a photon (a quantum of light) scattered in this way makes it more likely that the photon will encounter a methane molecule and be absorbed by it. The likelihood increases even for photons whose wavelength is somewhat different from the wavelength that defines the center of an absorption band. (According to quantum theory, such absorptions are unlikely but not impossible.) In this way the bands are broadened.

Models for Titan

At the State University of New York at Stony Brook, Barry L. Lutz, Robert D. Cess and I turned to a large set of laboratory spectra of methane made by Lutz in an effort to improve our understanding of the outer planets and Titan. We found that the absorption bands Wildt and Kuiper had studied are insensitive to pressure. Each of the bands consists of narrow absorption lines spaced so that any broadening of the individual lines by an increase in the pressure of the methane has no effect on the band overall. From the strength of the bands we deduced the abundance of methane on Titan. It was equivalent to a column of methane about 120 meters long at STP. Then from the relation between abundance and pressure that Trafton had proposed we deduced the atmospheric pressure that broadens the methane bands on Titan. (In this analysis we ignored the effects of the light scattered by particles.)

The result surprised us. We found that the pressure of Titan's atmosphere at the base of the zone of the atmosphere that reflects light is nearly 400 millibars, or almost half the pressure at sea level on the earth, which is approximately one bar. The methane by itself would give rise to a pressure of only one millibar. Trafton's second alternative seemed to be the right one: a gas other than methane is present in large quantity on Titan.

As we attempted to deduce the pressure of Titan's atmosphere other workers were attempting to account for the contradictory measurements of the temperature of the body. The first idea had been that the apparent high temperature of Titan, such as the 165 degrees K. reported by Low, results from the "greenhouse" effect. This concept assumes that the atmosphere of Titan is transparent to visible light but opaque to the infrared. Hence sunlight penetrates to the surface and heats it; then the surface gives off infrared radiation, which is trapped in the lower atmosphere. The concept had been compromised by the accumulating contradictory measurements, which suggested that the atmosphere has an unusual structure in which the upper layers are warmer than the surface. Thus the greenhouse concept had given way to a model championed by Robert E. Danielson and John J. Caldwell of Princeton University.



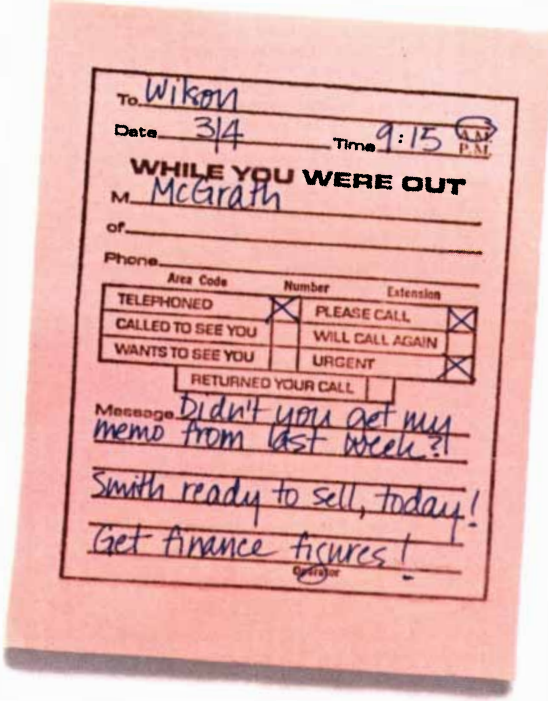
LIMB DARKENING causes the image of a moon or planet that has an atmosphere (*top*) to be darker at its limb, or periphery, than at its center. The darkening arises because the light reflected from the limb traverses a longer path through the body's atmosphere than the light reflected from the center. It is therefore attenuated by atmospheric absorption to a greater degree. In contrast the image of a planet or moon without an atmosphere (*bottom*) is a disk of more or less uniform brightness. In 1908 the Catalan astronomer José Comas Solá suggested that Titan must have an atmosphere because its image in his telescope showed darkening toward the limb.

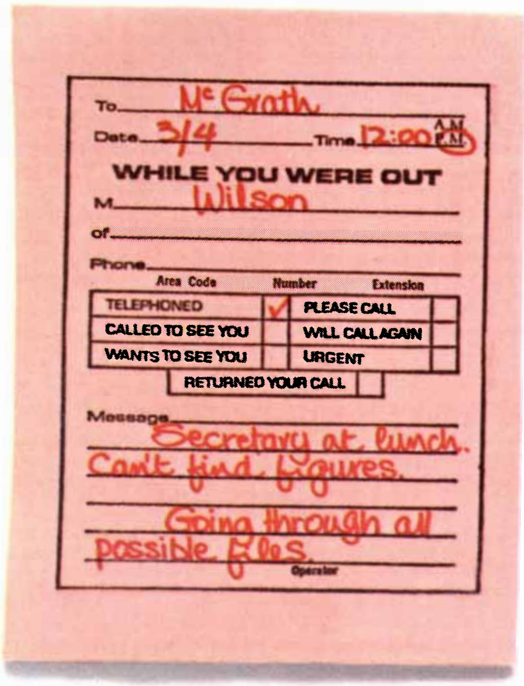
In that model the infrared brightness of Titan is ascribed to the emission of infrared radiation by an inversion in the atmosphere: a layer in which the temperature of the atmosphere increases with altitude instead of decreasing. Specifically, in the model the atmosphere is heated because sunlight is absorbed by particles that make up an aerosol (a suspension of particles) high in the atmosphere. The infrared radiation detected on the earth is emitted by methane and also by gases such as ethane (C_2H_6) and acetylene (C_2H_2), the products of light-induced chemical reactions in which methane molecules are destroyed by ultraviolet radiation from the sun. According to a later elaboration of the model by Caldwell, the atmosphere of Titan is 90 percent methane. Its surface pressure is 20 millibars and its surface temperature is 86 degrees K., a value in agreement with the prediction that would be made on the basis of simple solar heating of a body at Titan's distance from the sun. The evidence favoring a higher surface pressure than Caldwell contemplated could be rejected on the grounds that the scattering of sunlight by the aerosol particles could give the methane absorption bands the same appearance they would get from the pressure of a gas other than methane.

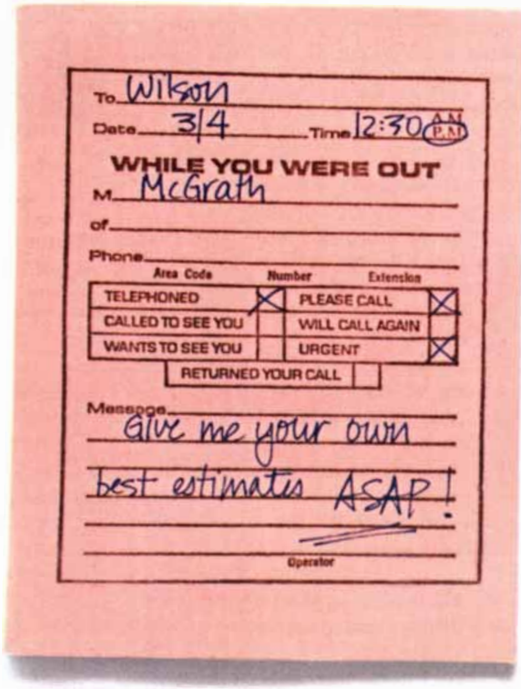
An alternative to this model was one proposed by Donald M. Hunten of the University of Arizona. Hunten pointed out that the light-induced dissociation of ammonia molecules in Titan's atmosphere could lead to the accumulation of molecular nitrogen; the hydrogen liberated by the dissociation would rapidly escape from the body. Since nitrogen is transparent to visible light and to infrared radiation it could not be detected spectroscopically from the earth. Hunten demonstrated, however, that if Titan's atmosphere includes enough molecular nitrogen to contribute 20 bars to the pressure at the surface, the increased pressure could lead to the increased absorption of infrared by nitrogen itself through collision-induced absorption. In this way a greenhouse effect could heat the surface of Titan to the temperature of 200 degrees K. inferred from one of the existing radio measurements.

Hunten's model included the aerosol layer and the corresponding tempera-

The life and death of a business deal.





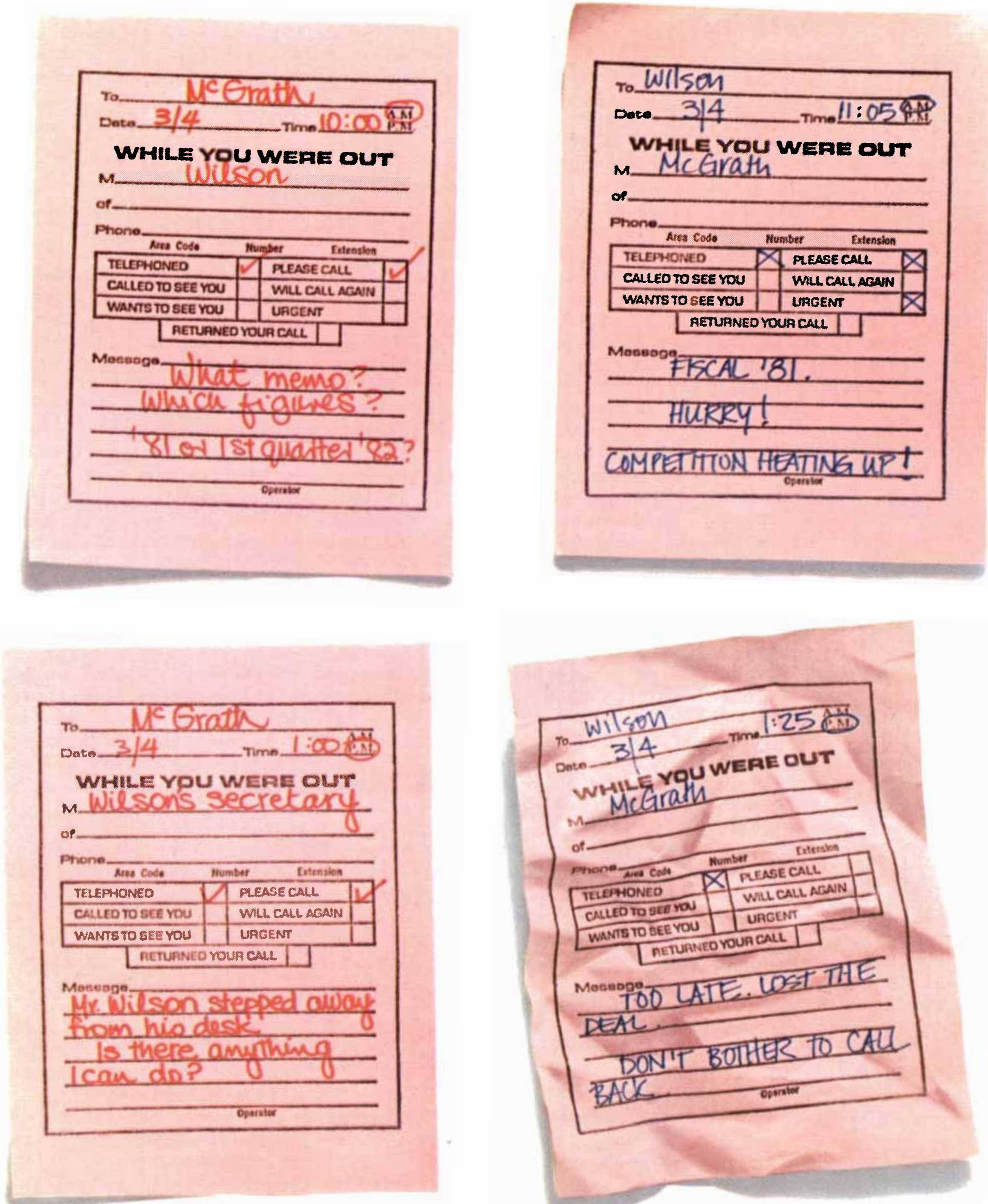


We know that half of all business phone calls never reach the intended party.

The average memo takes a full three days simply to go from one mailroom to another.

IBM office systems can help.

Our computers, word processors, printers and copiers form a system, office-wide or world-wide, that can manage the information you need to manage your business.


To McGrath
Date 3/4 Time 10:00 A.M.
WHILE YOU WERE OUT
M Wilson
of
Phone
Area Code Number Extension
TELEPHONED
PLEASE CALL
CALLED TO SEE YOU
WILL CALL AGAIN
WANTS TO SEE YOU
URGENT
RETURNED YOUR CALL
Message What memo? Which figures? '81 or 1st quarter '82?
Operator
To Wilson
Date 3/4 Time 11:05 A.M.
WHILE YOU WERE OUT
M McGrath
of
Phone
Area Code Number Extension
TELEPHONED
PLEASE CALL
CALLED TO SEE YOU
WILL CALL AGAIN
WANTS TO SEE YOU
URGENT
RETURNED YOUR CALL
Message FISCAL '81. HURRY! COMPETITION HEATING UP!
Operator
To McGrath
Date 3/4 Time 1:00 P.M.
WHILE YOU WERE OUT
M Wilson's secretary
of
Phone
Area Code Number Extension
TELEPHONED
PLEASE CALL
CALLED TO SEE YOU
WILL CALL AGAIN
WANTS TO SEE YOU
URGENT
RETURNED YOUR CALL
Message Mr. Wilson stepped away from his desk. Is there anything I can do?
Operator
To Wilson
Date 3/4 Time 1:25 P.M.
WHILE YOU WERE OUT
M McGrath
of
Phone
Area Code Number Extension
TELEPHONED
PLEASE CALL
CALLED TO SEE YOU
WILL CALL AGAIN
WANTS TO SEE YOU
URGENT
RETURNED YOUR CALL
Message TOO LATE. LOST THE DEAL. DON'T BOTHER TO CALL BACK.
Operator

ture inversion that Danielson and Caldwell's model had placed in the upper atmosphere. Other than that the models disagreed disturbingly. To summarize, Danielson and Caldwell proposed that the surface of Titan has a temperature of 86 degrees K. and that the atmosphere is 90 percent methane; Hunten proposed a surface temperature of 200 degrees and an atmosphere of 90 percent nitrogen. Worst of all, Danielson and Caldwell proposed a surface pressure of 20 millibars; Hunten proposed 20 bars, a figure 1,000 times greater.

It seemed possible that this worst discrepancy could be reduced, and perhaps even eliminated, by better measurements of the brightness of Titan at radio wavelengths. The point of making such measurements is that most of the predicted constituents of an atmosphere in the outer solar system, including methane and nitrogen, are transparent at radio wavelengths; therefore the measurements would probably represent the thermal emission from Titan's surface and not a layer in the atmosphere. A problem encountered in such measurements is that Titan's weak emission is swamped by the radiation emitted by Saturn. To diminish the problem Walter J. Jaffe of the National Radio Astronomy Observatory, working with Caldwell and me, employed the Very Large Array of radio telescopes in New Mexico. We made observations of Titan at the radio wavelengths of 1.3, two and six centimeters. The Very Large Array has sufficient angular resolution to separate Titan's emission from Saturn's. We found that Titan has a surface temperature of 87 degrees K. plus or minus nine degrees. Our finding allowed Danielson and Caldwell's model. It also allowed Hunten's model if that model is modified to have nitrogen contribute a maximum of two bars to the surface pressure. We had reduced the discrepancy, but only by a factor of 10.

Exploration by Voyager

That was how matters stood in the fall of 1980 as *Voyager 1* approached the Saturn system. Titan was known to be a large moon with an atmosphere at least three times denser than that of Mars (where the average surface pressure is seven millibars). The postulated presence of ethane, acetylene and a high-altitude aerosol suggested an active photochemistry. Methane had been detected spectroscopically; the presence of nitrogen had been proposed. Thomas W. Scattergood and I suggested the red color of Titan was indirect evidence that nitrogen was indeed present. Scattergood had attempted to make colored compounds by bombarding various mixtures of gases with energetic protons. His intent was to simulate the bombardment of Titan's atmosphere by charged subatomic particles trapped in the magnetic field surrounding Saturn. His efforts were unsuccessful if the gas was methane alone. If the gas was a mixture of methane and nitrogen, however, the bombardment led to the production of a reddish-brown material. But how much nitrogen was present on Titan? And what other gases were there?

The first results of the arrival of *Voyager 1* near Titan were images of the body. They were rather disappointing. Some investigators had hoped to see breaks in the aerosol layer that would allow a glimpse of Titan's surface. Instead the images showed a moon that resembled a fuzzy, seamless tennis ball. The aerosol was ubiquitous and opaque. The only markings visible in it were a dark north-polar hood and an abrupt change in reflectivity at the equator, so that the southern hemisphere was distinctly brighter than the northern. Titan was also shown to be surrounded by a high-altitude layer of haze about 100 kilometers above the top of the aerosol layer.

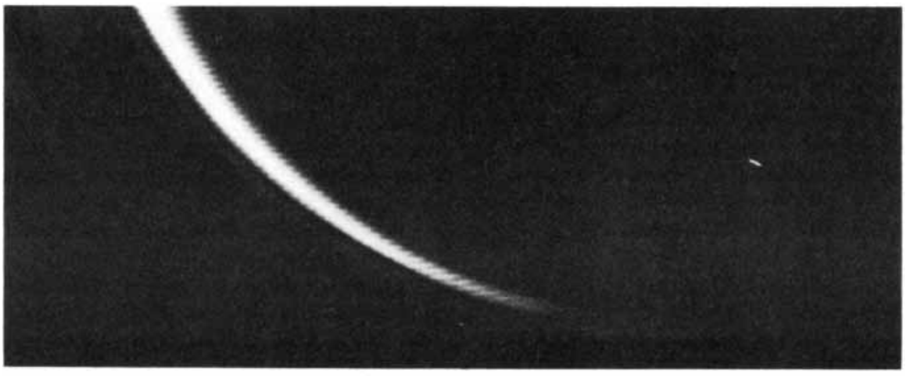
The difference in reflectivity between the northern and southern hemispheres has received two tentative explanations. On a hypothesis developed primarily by Lawrence A. Sromovsky and Verner E. Suomi of the University of Wisconsin at Madison the difference is a manifestation of the seasonal change in the solar heating of the aerosol layer caused by the 26-degree inclination of Titan's axis of rotation with respect to the plane of the solar system. If the hypothesis is correct, the southern hemisphere should alternate with the northern hemisphere in being the brighter one. Saturn and its moons take 30 years to circle the sun; hence the alternation should come once every 15 years.

On another hypothesis, developed by G. Wesley Lockwood of the Lowell Observatory, the difference results from a modulation in the production of aerosol particles due to the differing rates of arrival of the high-energy subatomic particles of the solar "wind." Over the past eight years Lockwood has been recording small changes in the net brightness of Titan. For much of that time the number of spots on the sun has been increasing, and with this increase the solar wind has been intensifying. Continued observations of Titan throughout the next few years, in the waning part of the sunspot cycle, should help to test the two hypotheses. They both may be correct.

After the images of Titan came other data from *Voyager 1*. For one thing, the ultraviolet spectrometer on board the spacecraft revealed to a group led by Lyle Broadfoot of the University of Southern California the presence of nitrogen: the instrument detected peaks in the ultraviolet spectrum of Titan due to the emission of ultraviolet radiation by nitrogen molecules, ionized nitrogen atoms and un-ionized nitrogen atoms. The spectrum gave no suggestion of carbon monoxide, argon or neon (other substances that emit in the ultraviolet). On the other hand, the presence of methane and other hydrocarbons was suggested as the spectrometer monitored the absorption of the light from a star by Titan's atmosphere. It was in essence an occultation experiment in which the limb of Titan, and therefore Titan's atmosphere, intervened between the star and the spectrometer.

In a second occultation experiment the spacecraft itself was the source of the radiation. Here the radio signals beamed toward the earth by *Voyager 1's* transmitters were attenuated by refraction in Titan's atmosphere. The attenuation grew as the spacecraft disappeared behind the limb of Titan and the density of gas traversed by the beam increased. The result of the experiment was a profile of density with respect to altitude in Titan's atmosphere. The profile of density, in turn, yields a profile of $T/\bar{\mu}$ with respect to altitude, where T is temperature and $\bar{\mu}$ is the mean molecular weight of the atmosphere. The experiment had been designed by G. Leonard Tyler of Stanford University to allow for either of the two prevailing models of Titan's atmosphere: Danielson and Caldwell's or Hunten's.

In the end it favored Hunten's. In particular, Von R. Eshleman of Stan-



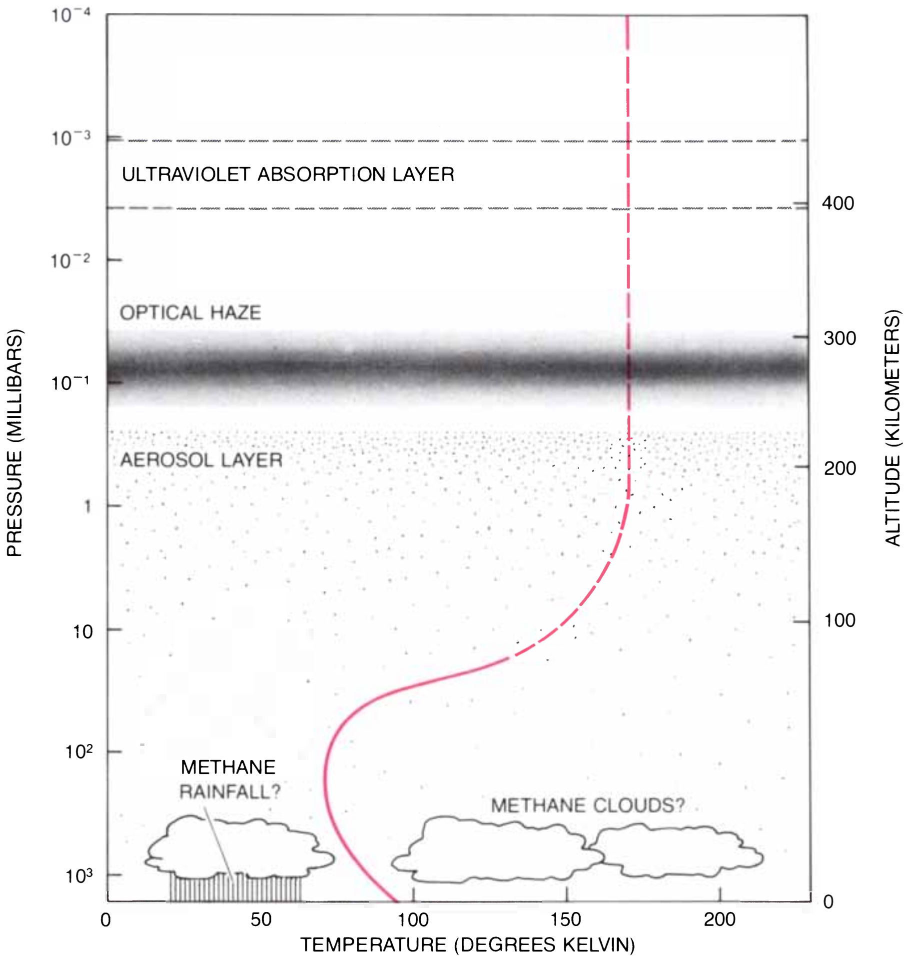
DISTINCT LAYER formed by Titan's high-altitude haze particles is shown in a photograph of Titan's south pole made by *Voyager 2*. The haze layer is concentric with the brightest feature in the photograph: the southern cusp of the sunlit crescent of Titan's opaque aerosol layer.

ford pointed out that the profile of T/μ produced by the experiment closely matches the predictions of Hunten's model for an atmosphere rich in nitrogen except that the profile implies a surface pressure lower than Hunten's original value of 20 bars. The group responsible for the occultation experiment is continuing to analyze its data. Gunnar F. Lindal of the Jet Propulsion Laboratory of the California Institute of Technology has given the group's most recent report of a surface pressure of 1.5 bars (plus or minus .1) and a surface temperature close to 94 degrees K. Lindal has also presented a value of 2,575 kilometers (plus or minus two) for the radius of Titan as derived from the occultation.

The structure of Titan's atmosphere was also analyzed by an infrared spectrometer group led by Rudolf A. Hanel of the Goddard Space Flight Center of the National Aeronautics and Space Administration. The best mutual fit of the occultation data and the infrared data indicates that the mean molecular weight of Titan's atmosphere is 28.6; hence the atmosphere must include an appreciable amount of a gas heavier than nitrogen. (The molecular weight of nitrogen is 28.0.) Robert E. Samuelson of the Goddard Space Flight Center suggests the gas is argon. He and his colleagues reason that argon is relatively abundant in the universe and that it is gaseous at Titan's temperature. Moreover, argon (like nitrogen) is transparent in the visible and the near infrared, so that it escapes detection by infrared spectroscopy.

The fact that argon was not detected by the ultraviolet spectrometer on *Voyager 1* means only that its abundance in the upper atmosphere of Titan is less than 6 percent. In the upper atmosphere one expects the lighter gases to dominate. Hence the limit of 6 percent in the upper atmosphere of Titan does not rule out the net abundance of about 12 percent required to give the atmosphere a mean molecular weight of 28.6. By setting $\bar{\mu}$ equal to 28.6 in the profile of $T/\bar{\mu}$ one finds a surface temperature of 95 degrees K., plus or minus two degrees.

The data transmitted to the earth from the vicinity of Titan by *Voyager 1* have thus revealed a moon in the outer solar system whose atmosphere, like the earth's, is rich in nitrogen. Furthermore, the data show that this body's atmosphere has a surface pressure greater than that of the earth's. The surface pressure of a body's atmosphere represents both the quantity of gas in the atmosphere and the degree to which it is compressed by the gravitation of the body. The gravity at the surface of Titan is only .14 times as strong as the gravity at the surface of the earth. Remarkably, therefore, the surface pressure on Titan (1.5 bars, according to the Voyager data) means that the atmosphere of Titan has about 10 times as much gas per unit area of the surface of the body as the atmosphere of the earth.



CROSS SECTION OF TITAN'S ATMOSPHERE includes two layers whose presence was discovered by *Voyager 1*. They are a layer transparent to visible light in which ultraviolet radiation is absorbed and below it the layer of high-altitude haze. Below the haze lies the layer of aerosol particles. It is presumed that the particles have been aggregating into larger particles and falling to the surface over the history of the solar system. Methane clouds and methane rainfall are shown above the surface; they are unconfirmed but likely. The curve showing temperature v. pressure (*color*) is based on an experiment in which the atmosphere of Titan intervened between the earth and the radio signals transmitted by *Voyager 1*. According to the data amassed by means of this occultation (along with Voyager data from infrared spectroscopy), the surface temperature on Titan is about 95 degrees Kelvin and the surface pressure is 1,500 millibars (1.5 bars). The average sea-level pressure on the earth is slightly more than one bar.

The infrared spectrometer on *Voyager 1* also recorded emission bands due to several gaseous substances whose presence in Titan's atmosphere had not been previously established. The first of these to be identified was hydrogen cyanide (HCN), a substance that may have been part of the chemical reactions that led to the synthesis of compounds such as adenine on the earth three billion years ago. Adenine is a constituent of DNA; hence it is essential to life on the earth. In subsequent studies Virgil G. Kunde and William Maguire of the Goddard Space Flight Center and their colleagues compared the infrared spectra of Titan with laboratory spectra they had made for that purpose. The comparisons rapidly led to the identification of six further substances in Titan's atmosphere. They include hydrocarbons such as propane (C_3H_8) and nitrogenous compounds such as cyanoacetylene (HC_3N). Darrell F. Strobel of the Naval Research Laboratory has shown that these substances can arise from reactions involving methane and nitrogen that are driven by the bombardment of Titan's atmosphere by ultraviolet photons from the sun and by high-energy electrons trapped in Saturn's magnetic field.

The Surface of Titan

It now seems clear that the early speculation regarding the origin of Titan's aerosol is essentially correct. The molecular fragments and compounds produced by the impact of ultraviolet photons and of high-energy electrons form polymers, or molecular chains. In that way they come to be suspended as solid particles in the atmosphere. By studying the manner in which the aerosol particles reflect sunlight, James B. Pollack and Kathy Rages of the Ames Research Center of NASA could show that the variation in the brightness of Titan with the change in the angle between the sun and Titan and each of the Voyager spacecraft could be explained if the aerosol particles high in the atmosphere have a mean radius of .5 micrometer.

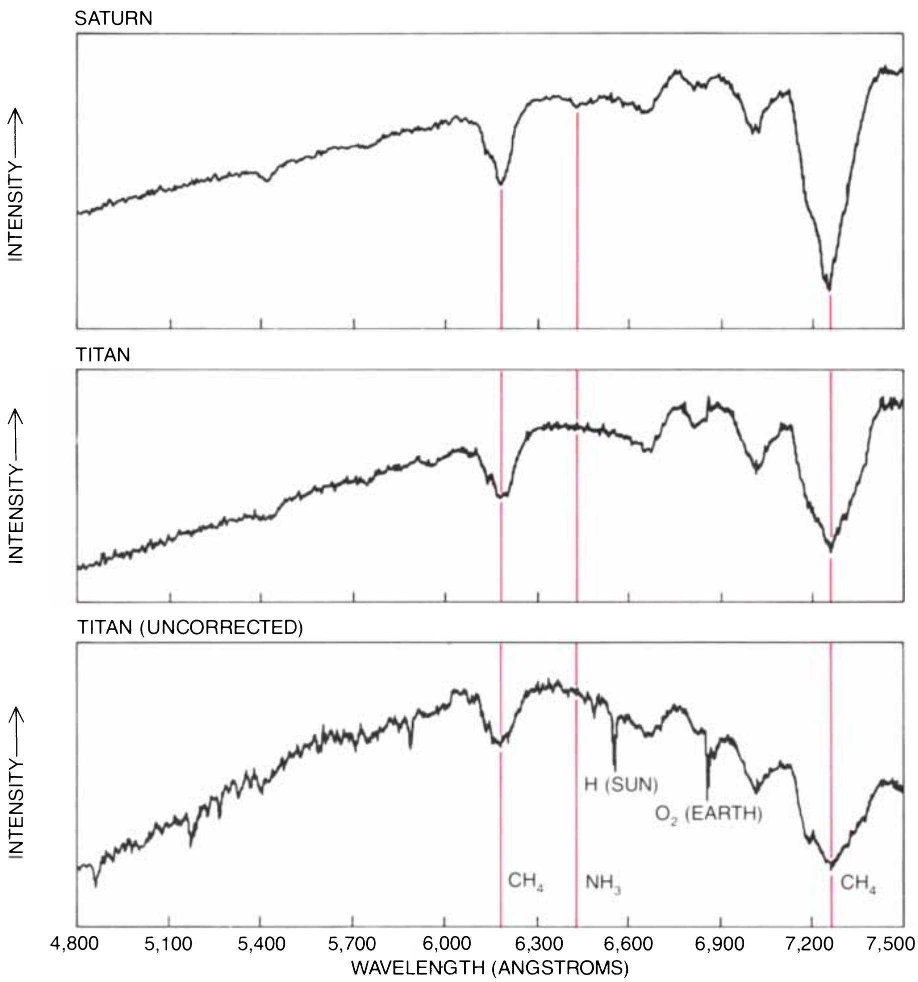
It can be hypothesized that these particles slowly settle out of suspension and that as they sink they collide and aggregate. The aggregates fall faster. Hence the atmosphere steadily loses the carbonaceous and nitrogenous molecules produced by processes high above the surface of Titan. Strobel has estimated that over the age of the solar system a quantity of hydrocarbons that would amount to a layer .1 to .5 kilometer deep has been deposited on the surface of Titan. Along with it has come a deposit of nitrogenous compounds that would amount to a layer some tens of meters deep.

What is the nature of the surface onto which this manna from heaven (as Carl Sagan of Cornell University likes to call the stuff) is falling? Refined analyses of the data from *Voyager 1* place the surface temperature of Titan at 94 degrees K., plus or minus one degree. Moreover, measurements made by the infrared spectrometer on *Voyager 1* suggest that the surface temperature varies by no more than three degrees between the equator and the poles. The reason is the ubiquity of the dense, light-absorbing atmosphere. These values of the surface temperature allow the presence of liquid methane. Indeed, they make it quite possible that Titan is covered by a global liquid ocean of what we on the earth call natural gas.

Methane, therefore, may play the same role on Titan that water plays on the earth. At the surface of Titan the methane is a liquid. In the lower atmosphere it is a gas. Perhaps the lower atmosphere of Titan includes methane clouds, and perhaps the lower atmosphere occasionally becomes saturated with methane in one place or another, so that a methane rain results. At last, it seems, we have found a world besides the earth where a large quantity of some compound is a liquid at the surface. It is too bad for the potential astronaut visitor that the temperature there remains so close to 94 degrees K., or −179 degrees Celsius.

What would it be like to sit in a boat on a methane sea on Titan? The horizontal visibility would be quite good. According to O. Brian Toon of the Ames Research Center, the large aerosol particles dropping out of the atmosphere would be few and far between. The visibility would diminish, of course, in a methane rainstorm. On the other hand, the light would be quite dim. Saturn is nearly 10 times farther than the earth from the sun. That alone diminishes the arrival of sunlight per unit area by a factor of 100. The weakened light would be further attenuated by the aerosol layer and by any methane clouds that happened to be above. It is difficult to estimate how extreme the net attenuation would be. Some of Toon's models predict that the view from the boat would be about as bright as a moonlit night on the earth, even at noon on Titan. Navigation would be hard to manage. The sun and the stars would not be visible. In addition a compass would be useless, because the Voyager spacecraft detected no magnetic field from Titan.

What kind of boat would be appropriate? Probably not a sailboat, because the winds at the surface of Titan are likely to be weak. The reason is the lack of pronounced differences in temperature from place to place. Such differences power the winds on the earth. Could one rely, then, on an outboard motor? Here there is a curious contrast. On the earth the oxidant for an internal combustion engine is freely available in the atmosphere. The fuel, however, is comparatively scarce. On Titan a boat would be afloat on a sea of fuel; it is the oxidant that would be scarce. Perhaps one could get it by drilling for water ice in the interior of Titan and extracting the oxygen from it. Or perhaps outcrops of water ice covered by a layer of hydrocarbons and nitrogenous polymers from the aerosol layer form continents on Titan.



SPECTRA OF SATURN AND TITAN were made by Robert G. Danehy and the author with the aid of the 2.7-meter telescope at the McDonald Observatory in Texas. Blue wavelengths of visible light are at the left of the scale; deep-red wavelengths are at the right. Several absorption bands appear. The one at 6,190 angstrom units and the one near 7,200 angstroms are among those that led to the identification of gaseous methane (CH_4) on Titan. A less prominent band near 6,450 angstroms reveals gaseous ammonia (NH_3) on Saturn but not on Titan. In general the stronger absorption bands in spectra of Titan are shallower and broader than those in spectra of Jupiter and Saturn. The reason is thought to be the scattering of light by Titan's aerosol particles. The top two spectra shown here have been divided by a spectrum of the light reflected from the earth's moon. In this way the spectral lines normally present in sunlight have been reduced to a value of unity and kept from being mistaken for spectral lines of Titan. The bottom spectrum of Titan has not been similarly corrected. It thus shows some spectral lines caused by the absorption of light in the atmosphere of the sun or the atmosphere of the earth.

How the Atmosphere Formed

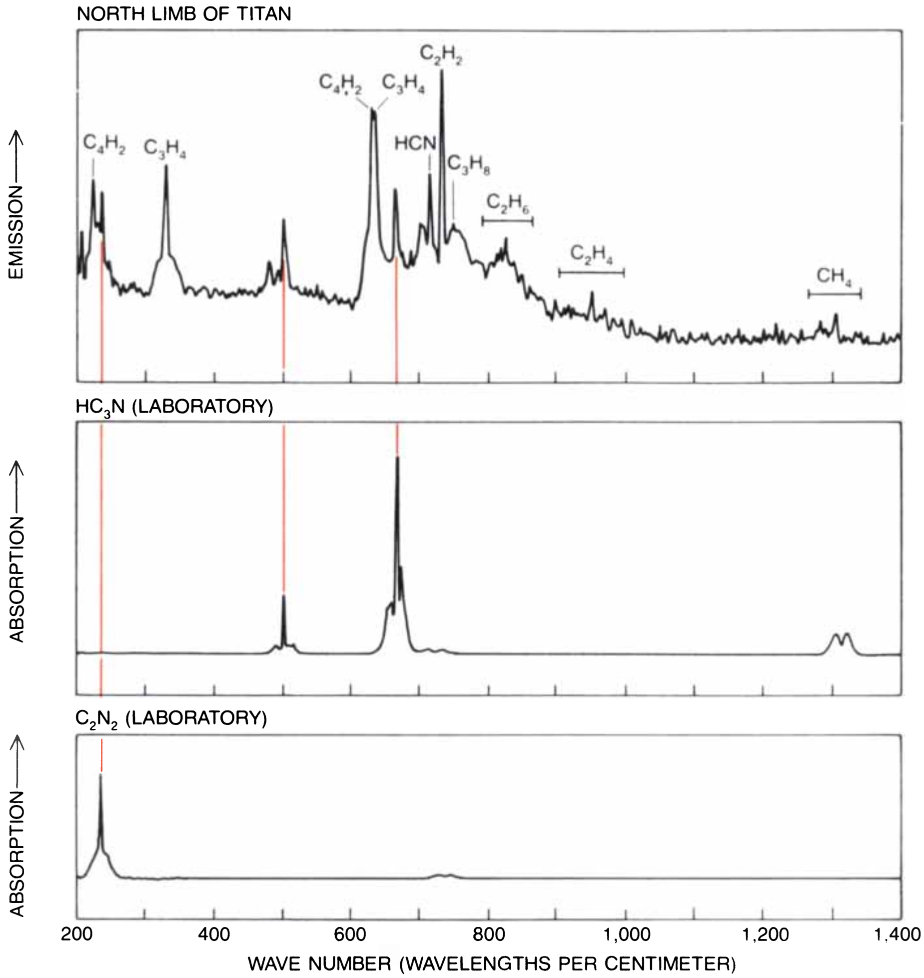
In the wake of the Voyager missions we must try to understand how the curious atmosphere of Titan evolved. In particular we are faced with a body from which hydrogen escaped, as Jeans could have deduced 65 years ago. In this respect Titan resembles the inner planets of the solar system. Why then has

Titan failed to develop an atmosphere like that of Mars or Venus, an atmosphere rich in carbon dioxide? The reason is that oxygen is unavailable: it is trapped in water ice inside the solid moon. The unique combination of Titan's size and Titan's temperature has allowed the atmosphere of Titan to evolve and yet remain a reducing one.

It is generally accepted that Titan formed in a proto-Saturnian nebula, an isolated part of the cloud of dust and gas that became the solar system. Then too it seems reasonably certain that Titan formed with Saturn, Saturn's rings and Saturn's other moons some 4.5 billion years ago. The density of Titan measured today (1.9 grams per cubic centimeter) indicates that it consists of approximately 52 percent rock and 48 percent ices. The proportions represent a slight enrichment in rock compared with the composition of the solar system overall. The ice, however, would have been crucial for the subsequent evolution of Titan's atmosphere, because the ice would have trapped gases from the proto-Saturnian nebula to a far greater extent than the rock would have.

Twenty years ago Stanley L. Miller of the University of California at San Diego predicted that the icy moons of Saturn should include methane hydrate ($CH_4 \cdot 7H_2O$), that is, methane trapped in water ice. The known presence of methane in Titan's atmosphere supported his idea. The newly discovered presence of several additional gases suggests that they too were trapped as hydrates. In order to predict with assurance which substances really were trapped, one must know the values of temperature and pressure for which a given substance and its hydrate are in equilibrium. (At equilibrium the rate at which molecules or atoms of a substance escape from the hydrate equals the rate at which they are trapped, so that the quantity of the hydrate does not diminish and the hydrate is stable.) It is thought that the proto-Saturnian nebula's temperature did not fall much below 60 degrees K. At that temperature the equilibrium pressure for nitrogen molecules or argon atoms and their respective hydrates is less than 10^{-7} bar.

A pressure of 10^{-7} bar is lower than the pressure that nitrogen or argon is likely to have contributed to the proto-Saturnian nebula; therefore these gases should have been trapped in ices. On the other hand, the equilibrium pressure for neon atoms and their hydrate at 60 degrees K. is nearly 40 bars. Neon has a high cosmic abundance; hence the absence of a detectable amount of it in Titan's atmosphere means two things. First, neon could not be trapped as a hydrate; second, it was not trapped as a gas. That is, the gravitation of Titan as it was forming was not strong enough to trap neon directly from the proto-Saturnian nebula. (The atomic weight of the most abundant isotope of neon is 20, or four more than the upper limit set by Jeans's theory for escape from Titan at its present mass.) The absence of neon tends to confirm that Titan's atmosphere formed after the body itself accreted and that the atmosphere formed from gases trapped as hydrates.



SPECTRA MADE BY VOYAGER 1 in the infrared part of the electromagnetic spectrum allow the identification of several gases on Titan other than methane. In this case the identification of cyanoacetylene (HC_3N) and cyanogen (C_2N_2) is demonstrated by a comparison of absorption spectra made in the laboratory with an emission spectrum of Titan made by the spacecraft. The comparison is valid because the molecules of a given gas absorb and emit radiation at the same set of characteristic wavelengths. Spectral features of several other gases identified by similar comparisons are labeled in the spectrum of Titan. For all three spectra the horizontal scale represents wave number, or waves per centimeter. A wave number of 200 corresponds to a wavelength of 500,000 angstroms; a wave number of 1,400 corresponds to a wavelength of about 71,000 angstroms. The spectroscopy was done by a group led by Rudolf Hanel of the Goddard Space Flight Center of the National Aeronautics and Space Administration.

How did the gases escape from the hydrates and get to the surface of Titan? In the first place the release of gravitational potential energy in the form of heat as Titan accreted would have been sufficient to vaporize a fraction of the ices in the body. Later the decay of radioactive nuclei inside Titan would have become the main source of heat inside it. According to models proposed by Mark Lupo and John S. Lewis of the Massachusetts Institute of Technology, the radioactive heating may have been sufficient to create a zone of liquid water deep in the mantle of the body. Gases could escape from the liquid.

Plainly there are ways in which gases once trapped in Titan's ices could escape and form an atmosphere. One sees evidence of such escape on other Saturnian moons. The cracks on Dione rimmed by material brighter than the surrounding terrain are the most conspicuous example. Dione was simply too small to retain an atmosphere. Other Saturnian moons show signs of fresh surfaces. The material that resurfaced the moons may have been driven upward to the surface partly by the internal pressure of gases escaping from hydrates.

The nitrogen in Titan's atmosphere calls for further discussion. In the events I have been describing it is assumed that the nitrogen in Titan's atmosphere today was incorporated into the accreting Titan as a hydrate. This assumption requires in turn that the dominant form of nitrogen in the proto-Saturnian nebula was molecular (N_2), which may not have been the case. To be sure, Ronald G. Prinn of M.I.T. has joined with Lewis in suggesting that N_2 was the stable form of nitrogen in the incipient solar system. Prinn and M. Bruce Fegley, Jr., of M.I.T. point out, however, that the increase of temperature near the incipient Jupiter and Saturn could have allowed ammonia (NH_3) to form. If it did and if Titan trapped it as a hydrate instead of trapping N_2, the subsequent history of Titan must have been substantially different.

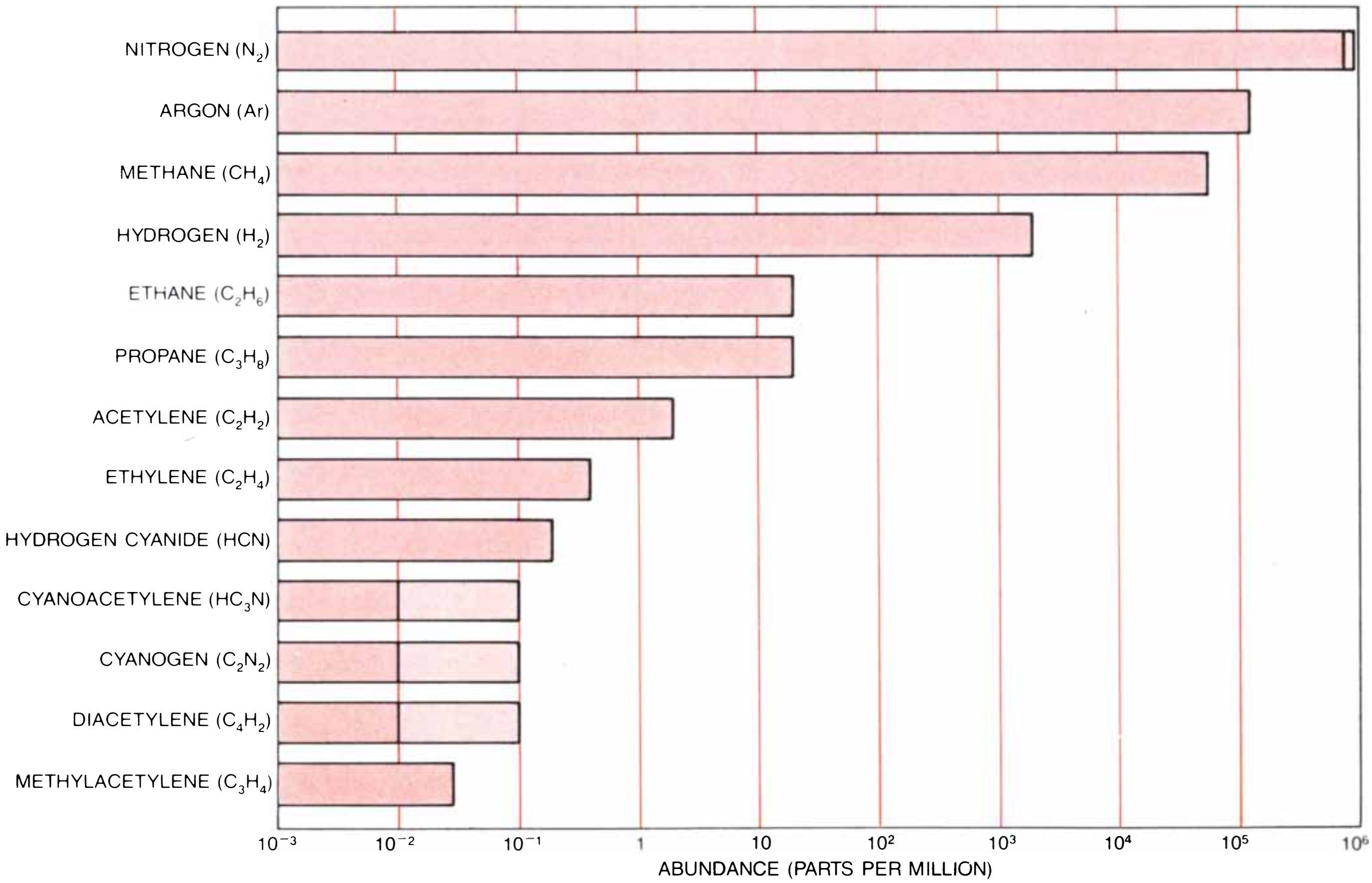
In particular, calculations made by Sushil K. Atreya and his colleagues at the University of Michigan show that one must postulate a "warm" epoch early in the history of Titan in which the surface temperature exceeded 150 degrees K. Throughout this epoch ammonia would have escaped from the interior and into the atmosphere, where it would have been broken up by solar ultraviolet photons. In this way the atmosphere would have lost its ammonia and gained the N_2 it has today. A temperature of 150 degrees K. is not out of the question; in principle a greenhouse effect created by hydrogen and ammonia in Titan's early atmosphere could have produced it. Still, the warm epoch is a complication that is avoided if nitrogen was trapped by Titan as a hydrate of N_2. Perhaps a combination of the two processes was involved, so that less of a constraint need be placed on the early surface temperature.

The argon in Titan's atmosphere also merits discussion. If it makes up 12 percent of the atmosphere, it must have come from trapped hydrates. The only competing possibility is that it came from radioactive decay. More than 99 percent of the argon in the atmosphere of the earth was made in just that way: it arose from the decay of the radioactive isotope potassium 40 into argon 40. Even if the rocky material in Titan had the same proportion of potassium as the rocky material in the earth, however, the current content of argon 40 in Titan's atmosphere would be about 70 parts per million, far from the predicted 12 percent, or 120,000 parts per million. The argon in Titan's atmosphere must indeed be primordial argon trapped as a hydrate from the proto-Saturnian nebula. It must therefore be argon 36 with an admixture of 20 percent argon 38.

The equilibrium pressures of nitrogen and argon with respect to their hydrates are sufficiently similar and sufficiently low to suggest that the proportions in which they become trapped as hydrates should be about the same as the proportions in which they later escape from those hydrates. A test of the presence of nitrogen instead of ammonia in the proto-Saturnian nebula begins, therefore, with an estimate of the amount of nitrogen that has escaped from the solid body of Titan. The estimate must include the nitrogen content of Titan's atmosphere today and also take account of how much nitrogen has escaped into space and how much has been incorporated into aerosol particles over the history of the body. It follows from Strobel's values for these two modes of depletion that the total amount of nitrogen that has entered Titan's atmosphere is about 1.7 times the atmosphere's current content of nitrogen, or 140 percent of the total content of Titan's atmosphere.

The amount of argon that has escaped from the solid body of Titan is easier to calculate. Argon is inert (it will not form chemical compounds), and it is too heavy to escape into space in appreciable quantity. The amount of argon that has escaped from the solid body is simply the 12 percent of the atmosphere. The ratio of nitrogen to argon is therefore 11.7. The ratio of nitrogen to argon in the cloud of matter that became the solar system was quite close to that: it was 11.

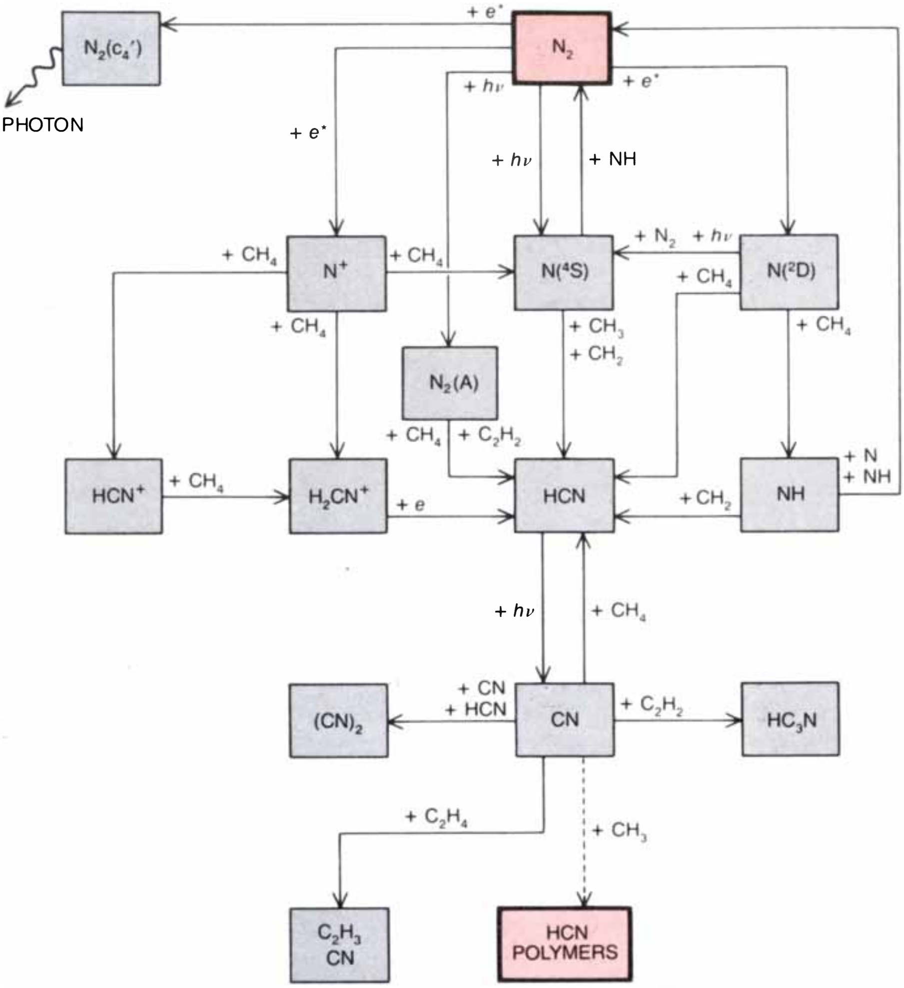
What can Titan tell us about the primitive earth? Opinion is shifting away



GASES IN TITAN'S ATMOSPHERE are now thought to vary in abundance from molecular nitrogen (set at 82 to 94 percent of the atmosphere, or 820,000 to 940,000 parts per million, as a result of Voyager data) to trace quantities of hydrocarbons such as methylacetylene and nitrogenous substances such as cyanogen. Several further trace constituents of Titan's atmosphere may remain to be discovered. The chart indicates some 12 percent (120,000 parts per million) of the inert gas argon. This 12 percent is required to raise the mean molecular weight of the gases that make up Titan's atmosphere to the value of 28.6 that tentatively emerges from the Voyager data.

from the view that the earth began with a highly reducing atmosphere consisting mostly of methane, ammonia, hydrogen and water vapor. The newer view is that the atmosphere was mildly reducing at first and that it consisted mostly of carbon dioxide, molecular nitrogen and carbon monoxide with no more than 10 percent of hydrogen. It is firmly agreed, however, that no free oxygen was present, because an oxidizing environment makes it extremely difficult for organic molecules to form. In an oxidizing environment the carbon is quickly locked into molecules no more complex than carbon dioxide.

Investigators working in laboratories on the earth have subjected gaseous mixtures of methane, ammonia, hydrogen and water vapor to repeated electric discharges. They have succeeded in producing, among other organic molecules, some of those that are constituents of DNA. What they have not been able to produce is something like DNA itself, or more broadly a molecule that can replicate itself by acting as a template for the assembly of a duplicate from the simpler molecules that are available. The problem seems to be a lack, in Andrew Marvell's phrase, of "world enough, and time." A test tube is simply too small and years may be too short for the advent of a self-replicating molecule to be likely.



CHEMISTRY FOR NITROGEN in Titan's atmosphere is proposed by Darrell F. Strobel of the Naval Research Laboratory. Here the main reactants and reaction products are shown; hydrogen atoms and molecules, among others, are omitted. The initial reactions involve the breaking of the N–N bond in a nitrogen molecule high in the atmosphere. The energy to break the bond is provided by the impact of solar ultraviolet radiation (*hν*) or of a high-energy electron (*e) trapped near Titan by Saturn's magnetic field. (The latter process supplies Titan's atmosphere with about 10 times as much energy as the former.) The final reactions include the formation of the reddish aerosol particles observed in Titan's atmosphere. The final steps are not fully known because the chemical composition of the particles themselves is not yet established. It is, however, known that hydrogen cyanide (HCN) can form a reddish-brown polymer. In the illustration ^2D and ^4S signify excited atomic states, A and c_4' signify excited molecular states and *e* signifies an electron freed in the atmosphere by solar ultraviolet radiation.**

Titan as a Laboratory

In this respect Titan offers what amounts to an immense natural laboratory free of experimental bias and the possibility of accidental catalysis or inhibition of chemical reactions by the walls of reaction vessels or the grease in the stopcocks of a vacuum system. Moreover, it is a laboratory where the reaction products over the history of the solar system have been collected into aerosol particles that rest at low temperature at the surface of the body, where we may one day be able to examine them at our leisure. Indeed, we may one day resolve to conduct experiments there ourselves. Heating up a few acres of Titan would produce more of the "primordial soup" than our terrestrial experiments could make in several lifetimes.

For now, NASA is contemplating a mission for the 1990's that would send a probe into Titan's atmosphere while a parent spacecraft was in orbit around Saturn, employing radar to pierce the opacity of Titan's atmosphere and make a map of the surface. We may expect much sooner than that to learn more about Titan from further analysis of the Voyager data, from the powerful new spectrometers now available on the earth and particularly from the instruments that will go into orbit around the earth with the Space Telescope in 1985.

Meanwhile the Voyager spacecraft continue their missions of exploration. *Voyager 1* is on its way out of the solar system. If its transmitters continue to function we can expect it to tell us the position of the boundary between the solar wind and the tenuous gas of the interstellar medium. The news should come within the next 10 years. *Voyager 2* is on its way to Uranus, which it will reach in January, 1986, and Neptune, which it will reach in August, 1989. The mechanism that points the cameras on the spacecraft suffered some damage when the spacecraft passed through Saturn's rings. Engineers at the Jet Propulsion Laboratory are confident, however, that *Voyager 2* will transmit valuable data from Uranus and Neptune if there are no further equipment failures.

Still, the Voyagers were not originally intended to travel to planets more distant than Saturn, and their instruments are far from optimum for that purpose. The change in plans was made in the mid-1970's when it became apparent that no specific missions to Uranus and Neptune would be undertaken for at least a decade. In the wake of the success of the Voyager program it is ironic that we find ourselves with no follow-on missions other than one that may be canceled: the Galileo project, which would send a probe into Jupiter's atmosphere to make measurements of such things as composition, pressure and temperature. We have already lost our chance to send a mission to Halley's comet. In effect we are retreating to the early 17th century, a time when the most distant known planet was Saturn and it was not understood that comets traverse the solar system. This is no time for such a retreat. There is much we need to learn about the worlds around us before we can fully understand the history of the solar system and that of the earth itself.

The Chemistry of Flames

Research in the field of combustion chemistry focuses mainly on the intermediate substances created as hydrocarbon fuels burn to produce carbon dioxide, water and trace pollutants

by William C. Gardiner, Jr.

The key discovery in the scientific effort to understand fire was made in 1774, when Antoine Laurent Lavoisier recognized that the apparent disappearance of matter in flames is an illusion. Instead, Lavoisier showed, an invisible component of the air (which he later named oxygen) reacts chemically with matter at high temperature, yielding heat and a variety of combustion products. Today this concept defines a fuel: a substance that can participate in an exothermic (heat-releasing) chemical reaction with oxygen.

Combustion science has taken on added significance in recent years, owing to increased public awareness of the finite supply of comparatively inexpensive fossil fuels and of the injurious effects of some combustion products on the environment and human health. With the aid of modern laboratory techniques it is possible to detect not only the end products of combustion processes but also many substances that appear transiently in the course of burning. As a result fire has come to be understood chemically as an intricate network of molecular events. The practical objective of this work remains what it was for prehistoric man: to learn how to burn the cheapest available fuels as efficiently, intensely and cleanly as possible.

Some fuels, such as wood or camel dung, are too complex to deal with in detail at the molecular level. Others, such as aluminum and ammonia, require more energy to produce than they liberate when they burn, and so they are normally considered impractical as fuels. Combustion scientists are primarily concerned with the fuels most often burned as energy sources: coal, petroleum products and natural gas. Such fuels are not pure chemical substances. Coal, which is mainly composed of elemental carbon, also has a variety of other components, both combustible and noncombustible. Petroleum fuels and natural gas are mixtures of hydrocarbons (compounds consisting of hydrogen and carbon in a roughly two-to-one ratio) with traces of other substances. In addition to these common fuels combustion chemists study special-purpose fuels such as acetylene for welding torches, powdered metals for fireworks and hydrogen for rocket engines.

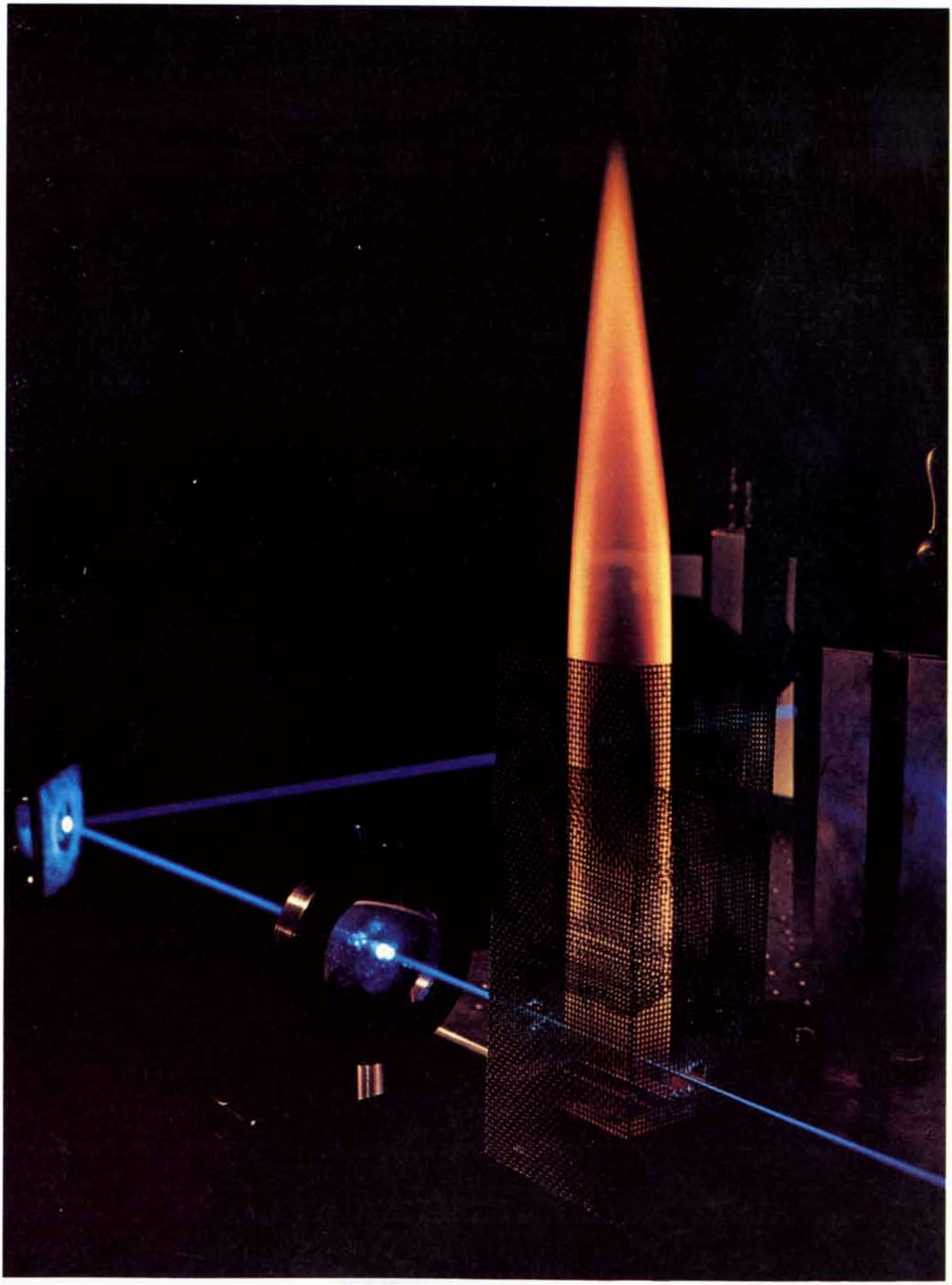
Although fuels have the ability to burn, they rarely do so spontaneously. A mixture of acetylene and oxygen, for example, could be stored for hundreds of years with no perceptible reaction. If the mixture were to be exposed to a flame or a spark, however, it would detonate in a microsecond. The layman's explanation is simple: A fire has to be lighted first, and it will then sustain itself until either the fuel or the oxygen runs out or until it is extinguished. The task of combustion science is to put the basic notions of ignition, propagation and extinction into precise chemical and physical terms. Doing so requires more details than are provided by merely identifying the overall chemical reaction involved in a fire.

The first combustion chemists believed temperature change alone would explain the phenomena of ignition and propagation. They reasoned that the chemical reactions involved in combustion are simply too slow to be perceived at normal temperatures; if the ignition temperature is reached in one place, however, the rate at which the ensuing chemical reaction releases heat becomes high enough to raise the temperature in the adjacent region to the ignition temperature too. This interpretation, which lumps all the chemistry into a single heat-releasing process whose rate depends on temperature, begs the question of why such a temperature dependence exists. Nevertheless, when the question is framed in quantitative terms, it does provide models that agree with experiments and are useful to engineers for designing combustion devices. Such models explain how burning rates depend on temperature, but not why. That is because they do not deal with the molecular events underlying the combustion process.

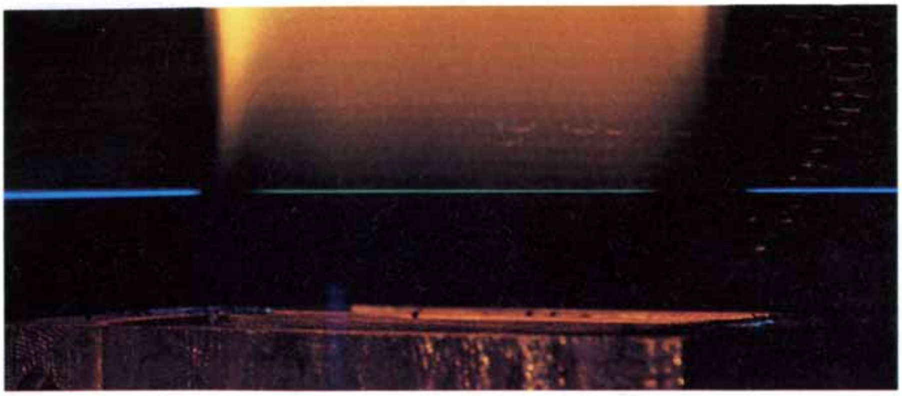
Before proceeding further some definitions are needed. In combustion science the term fire is generally applied to any accidental or deliberate combustion event in the "real" world (that is, the world outside the laboratory). In such an event, say a forest fire or a furnace fire, air is supplied and combustion products and heat are removed by means of large-scale, unsteady flows. Flames are the subunits of fire. They in turn can be broken down into two classes: premixed flames, in which the fuel is mixed with oxygen prior to combustion, and diffusion flames, in which the fuel and the oxygen meet in the combustion zone.

Gasoline and air (which is about 20 percent oxygen) mix in the carburetor of a conventional internal-combustion engine and burn later (after compression and spark ignition in the cylinder) as a premixed flame. Vaporized wax and air meet above a candlewick to burn as a diffusion flame. Premixed flames and diffusion flames can be either turbulent or laminar, depending on the flow rates of the substances involved. Turbulent flow increases the burning rate and so is advantageous when fast combustion is desired. Laminar, or smooth, flow is easier to describe mathematically and is usually better suited for laboratory experiments. In all the flames mentioned so far the fuel is in the form of a vapor before combustion begins. If it is not, then one is dealing with a more complicated physical situation: either heterogeneous combustion, such as takes place at the surface of a burning piece of coal, or a combined vaporization-combustion process, such as takes place in a diesel engine or a jet-aircraft engine.

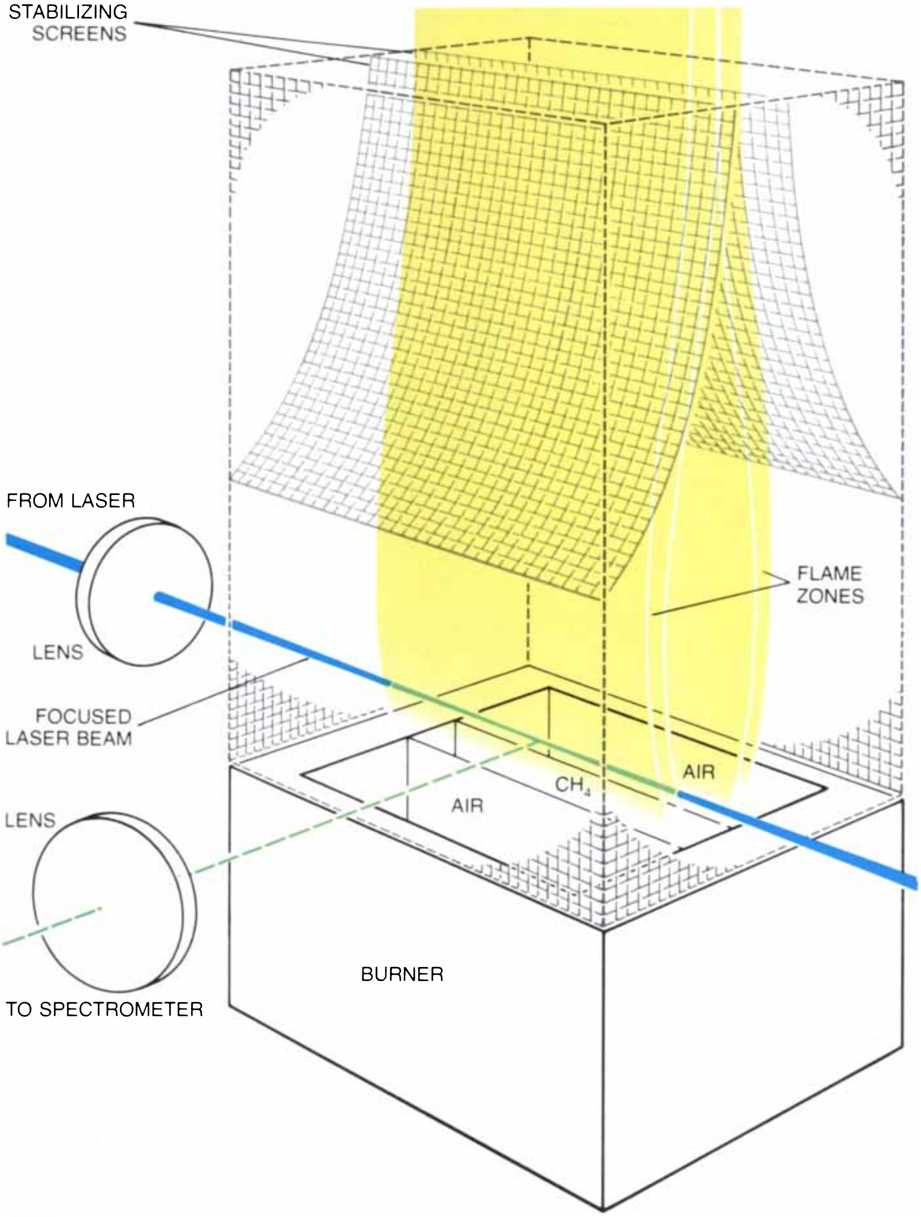
Of all these types of flames only the laminar premixed flame offers the possibility of conducting a quantitative study of chemical processes in spite of concurrent physical processes, in particular the diffusive movement of heat and matter. Even the comparatively simple reaction zone of a Bunsen-burner flame is not optimal for studying combustion chemistry, however; the flow pattern around such a flame is two-dimensional, which is one dimension more than can be dealt with in simple flow equations. Moreover, the reaction zone is very thin, so



LAMINAR DIFFUSION FLAME is stabilized with the aid of screens in this demonstration of the laser-probe technique, photographed by Fritz Goro at the Fire Research Center of the National Bureau of Standards. In a flame of this type the fuel and the oxidizer meet in the combustion zone; here the fuel is methane and the oxidizer is air. A closeup view of the flame, showing the green fluorescence excited by the blue laser beam from large organic molecules that may be soot precursors, is shown at the top of the next page.



THIN GREEN LINE of fluorescing gas molecules excited by the passage of a blue laser beam through the reaction zone of a methane-air diffusion flame was captured by Goro in this close-up photograph of the flame. The front stabilizing screen has been removed; the rear one is dimly visible through the flame. The brass flame holder is approximately four centimeters long.



FLUORESCENCE FROM FLAME can be analyzed by means of a spectrometer, as is shown in this schematic diagram of the experimental setup seen in the photograph on the preceding page. The methane fuel flows out of a slot in the middle of the burner; the fuel outlet is flanked by two laminar sheets of air flowing from channels along the sides of the burner. Combustion takes place in the two reaction zones at the interface of the fuel and the air. The laser beam is focused on the fuel side of reaction zone on near side of burner. The experiments on soot formation at the Fire Research Center are being done by W. Gary Mallard and Kermit C. Smyth.

that it is difficult to do chemical analyses within the flame. For laboratory studies of the structure of flames a one-dimensional flow pattern can be generated by replacing the tube of a Bunsen burner with a large, flat flame holder fitted with a porous plug. The reaction zone can then be stretched by reducing the pressure of the fuel-oxidizer mixture to less than atmospheric pressure.

In low-pressure premixed flames the course of the chemical reactions can be followed with suitable probes of the composition and the temperature of the gas in the combustion zone. Intrusive probes, such as quartz gas-sampling tubes and thermocouples, provide spatially resolved information about composition and temperature at the cost of disturbing the flow, often significantly and in a way that is difficult to describe mathematically. Optical probes, such as laser beams, have the advantage that the flame itself is undisturbed and the disadvantage that the path of the light beam is bent by density gradients in the flame. A variety of flames have been probed by a variety of methods since the late 1950's, when the technology of generating stable, low-pressure flames was developed. These continuing studies have provided a chemical description of the changes that take place in flames.

The variation in concentration of the molecular components in a flame as a function of distance from the burner can be represented by a set of graphs known as flame profiles. The mathematical analysis of such profiles, pioneered by Joseph O. Hirschfelder and Charles F. Curtiss of the University of Wisconsin at Madison in the 1950's, enables one to separate changes in molecular concentrations into two classes: changes resulting from diffusion and changes resulting from chemical reactions. One can then derive net chemical-reaction rates for each substance found in the flame; the resulting set of reaction-rate profiles serves as the basic chemical description of the flame. Even before combustion chemists developed the techniques for getting reaction profiles reliably, however, they knew that no set of flame profiles, however complete, would suffice to identify the molecular events responsible for the profiles. They knew already that the diversity of molecular events was far greater than the number of flame profiles they could ever hope to compile.

The molecules of most fuels have far too many atoms for combustion to proceed as a concerted event. Imagine the tangle that would arise if the eight carbon atoms and the 18 hydrogen atoms of an octane molecule (C_8H_{18}) were to disengage from one another and combine all at once with the surrounding molecules of diatomic oxygen (O_2), forming the new chemical bonds of car-

FLAME	CHEMICAL REACTION	TEMPERATURE (DEGREES KELVIN)	ENERGY RELEASE (JOULES PER GRAM)
HYDROGEN-OXYGEN	$2 H_2 + O_2 \rightarrow 2 H_2O$	3,100	24,000
METHANE-OXYGEN	$CH_4 + 2 O_2 \rightarrow CO_2 + 2 H_2O$	3,000	10,000
METHANE-AIR	$CH_4 + 2 O_2 \rightarrow CO_2 + 2 H_2O$	2,200	2,700
OCTANE-OXYGEN	$2 C_8H_{18} + 25 O_2 \rightarrow 16 CO_2 + 18 H_2O$	3,100	9,900
ACETYLENE-OXYGEN	$2 C_2H_2 + 5 O_2 \rightarrow 4 CO_2 + 2 H_2O$	3,300	11,800
CYANOGEN-OXYGEN	$C_2N_2 + O_2 \rightarrow 2 CO + N_2$	4,800	6,300
PRODUCER GAS–AIR	$2 CO + 4 H_2 + 3 O_2 \rightarrow 2 CO_2 + 4 H_2O$	2,400	4,100
METHYLHYDRAZINE – NITROGEN TETROXIDE	$CH_4N_2 + N_2O_4 \rightarrow 2 H_2O + CO_2 + 2 N_2$	3,000	7,500

FLAME CONDITIONS depend on both the fuel and the oxidizer. The temperatures listed for the chemical reactions in this table refer to flames burning at atmospheric pressure. At higher pressures the dissociation of the combustion products would be suppressed, leading to higher flame temperatures. The value given for the energy released by each reaction refers to the amount of energy liberated per gram of fuel and oxidizer together (including the inert components in the case of air) after the product gases have cooled back to ambient temperature. Hydrogen and cyanogen flames are important in the laboratory for high-temperature spectroscopic studies. Producer gas is a comparatively cheap fuel mixture prepared from coal and water. Methylhydrazine is burned with nitrogen tetroxide as the oxidizer in the attitude-control engines of the space shuttle; unlike most fuel-oxidizer combinations, these two substances ignite instantly on contact.

bon dioxide (CO_2) and water (H_2O). No fuel burns that way. Instead the breakdown of fuel molecules and the formation of combustion products proceed in long sequences of steps; each step involves only a small rearrangement of chemical bonds.

A step of this type is termed an elementary reaction; the various molecules created along the way are called reactive intermediates, and the set of all elementary reactions that together account for the net chemical transformation is the reaction mechanism. The equation describing the overall chemical reaction that takes place in a flame gives no hint of what the individual molecular changes are; the equations describing elementary reactions, however, do represent real chemical events at the molecular level. Only when all the important elementary reactions are known can the path from fuel to combustion products be accurately described in terms of the rearrangements of atoms in molecules.

Discovering the elementary reactions is the key to understanding how combustion proceeds. In order to get this key, however, it is not enough to identify particular elementary reactions as possible chemical events. One must also know what the probability is that the molecules concerned will indeed react when their paths cross, since most of their encounters do not lead to a chemical reaction. That probability turns out to depend on the relative speed with which the molecules collide. Usually the more violent the collision, the more likely it is to lead to a chemical reaction. Since both the frequency and the speed of molecular encounters depend on the temperature of the gas, the reaction probability increases, often quite sharply, with increasing temperature.

The study of reaction probabilities and their dependence on temperature falls into the area of chemical kinetics, a field in which it is the custom to express such probabilities in terms of the temperature-dependent proportionality constants called rate coefficients. The reaction rate is the rate coefficient multiplied by the concentration of each of the reactant molecules. Since the temperature-dependent values of the rate coefficients of most flame reactions have not been measured for the temperature range of flames, extrapolations are usually required. Fortunately the theory of rate coefficients has been thoroughly tested and is (with minor reservations) reliable enough for accurate extrapolations to be made.

The goals of combustion chemistry are therefore to identify the elementary reactions of a flame, to determine the rate coefficient of each reaction as a function of temperature and to assemble this information into flame models that provide testable predictions. To arrive at such predictions one employs computer programs in which the equations describing the diffusive and reactive rates of change in concentration of all the molecules in the flame are combined with the equations describing the flow of the reacting gases.

In searching for the elementary reactions of flames combustion chemists knew in advance the kinds of reaction they would have to identify. Early in this century it was recognized that the intermediate substances participating in most chemical reactions are present in extremely small amounts that are usually undetectable. This is particularly true when the intermediates are highly reactive compared with the other molecules present and when they enter readily into elementary reactions that produce other reactive intermediates. An example that proves to be very important in flame chemistry is the elementary reaction that converts carbon monoxide into carbon dioxide: $CO + OH \rightarrow CO_2 + H$. The reactive intermediate entering into the elementary reaction, the hydroxyl radical (OH), yields another reactive intermediate, atomic hydrogen (H), with the result that there is no change in the number of intermediates. (The term radical is used by chemists to denote reactive fragments of molecules.)

The hydrogen atom subsequently participates in other elementary reactions, some of which may in turn yield a hydroxyl radical, making it possible for the reaction with carbon monoxide to recur. Because of the repetition of the same molecular events this kind of reaction mechanism is called a chain reaction; the elementary reactions are called chain-propagating steps, and the atoms and radicals that participate in the propagation of the chain are called chain centers. A very small number of chain centers can lead to a large amount of chemical reaction.

Chain-propagating steps alone are insufficient to account for flames. It was recognized in the 1930's that chain-initiating, chain-terminating and chain-branching steps are also necessary. The 1956 Nobel prize in chemistry was awarded jointly to Cyril N. Hinshelwood of the University of Oxford and Nikolai N. Semenov of the Institute of Chemical Physics in Moscow for their leading roles in showing how the main features of oxidation reactions could be explained in terms of branched-chain reaction mechanisms. (At the time of their research chemists had only provisional ideas about what the actual elementary reactions might be and almost no information at all about the values of the rate coefficients.)

A chain-initiating step is an elementary reaction in which one or two normally stable molecules react to form one or two chain centers. An example is the reaction $C_3H_8 \rightarrow C_2H_5 + CH_3$, in which the stable molecule propane (C_3H_8) decomposes spontaneously, when it is raised to a high energy by collisions with other molecules in a hot gas, to form the two radicals C_2H_5 and CH_3, each of which can in turn serve as a chain center.

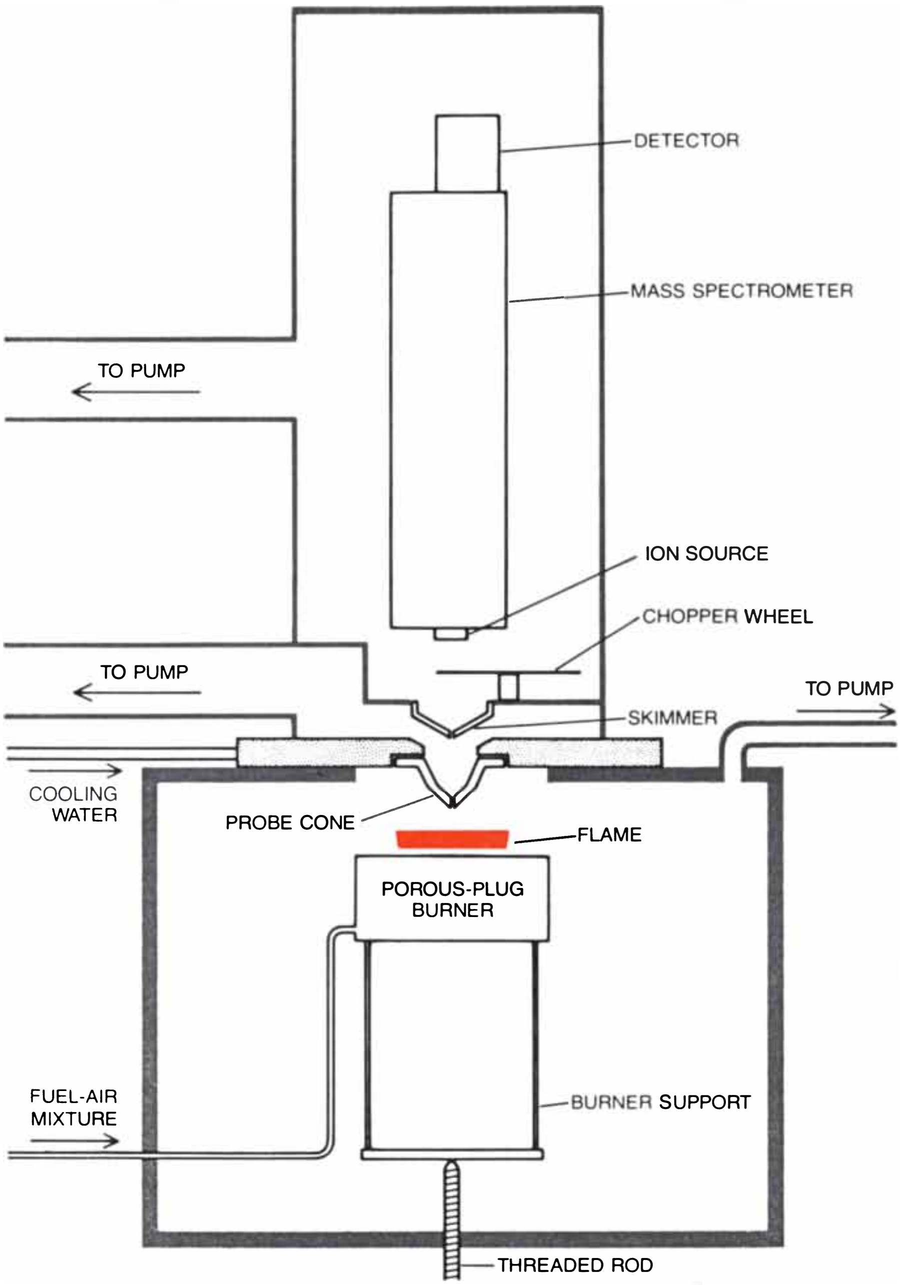
A chain-terminating step is an elementary reaction that has the opposite effect. For example, in the reaction $H + OH + N_2 \rightarrow H_2O + N_2$ two chain centers, a hydrogen atom and a hydroxyl radical, simultaneously encounter the chemically stable molecule nitrogen (N_2) to form another chemically stable molecule: water. Here the nitrogen molecule acts as a kind of chaperon, absorbing some of the energy that is liberated when the new chemical bond of the water molecule is formed.

The distinctive chemical feature of flames, however, is the participation of chain-branching steps: elementary reactions in which the number of chain centers increases from one to two or perhaps even three. The most important chain-branching reaction is $H + O_2 \rightarrow O + OH$, in which atomic hydrogen reacts with molecular oxygen to yield atomic oxygen and a hydroxyl radical.

This much basic theory has been known for half a century. To develop it into a detailed description of flame chemistry, however, two more factors were needed. The first was a set of practical mathematical procedures to find the combined effects of all the elementary reactions on observable chemical and physical properties. Expressing the physical processes of diffusion, heat conduction, chemical reaction and flow in mathematical form was only the start; the resulting simultaneous differential equations present two difficulties, one obvious and the other subtle.

The obvious difficulty is that an exact description of diffusion and heat conduction requires that one provide the values for certain parameters (called multicomponent transport coefficients) that cannot be measured in the laboratory. Approximate descriptions of them can be made only by relying on assumptions of questionable accuracy. Fortunately recent computer-modeling research has shown that approximations made in describing diffusion and heat conduction have only minor effects on the results.

The subtle difficulty is that the differential equations have a property called stiffness, which means that some of the concentration variables are held to nearly constant values by very rapid reactions, while others change. The effect of this factor is to require intolerably long computer time if the set of simultaneous differential equations is to be solved numerically on a digital computer by standard methods. Special computer techniques for dealing with stiff differential equations were developed in the 1970's, most of them based on pioneering research by Charles W. Gear of the University of Illinois at Urbana. The advent of stiffly stable numerical techniques made it feasible to model one-dimensional flames, taking detailed account of the flame chemistry, on large digital computers.



SAMPLING PROBE is used to follow the progress of combustion reactions in flames burning at low pressure. The rapid expansion of the gas entering the cone-shaped probe has a cooling effect that "freezes" the flame reactions. Accordingly the gas sample taken through the probe can be analyzed outside the reaction zone by conventional analytic methods, such as the mass spectrometer employed in this apparatus, devised by Joan C. Biordi, Charles P. Lazzara and John Papp at the U.S. Bureau of Mines Research Center in Pittsburgh. For obtaining flame profiles of substances too reactive to survive such a probing process, optical methods are required.

A computer simulation is only as good as its input data, in this case the rate coefficients for all the elementary reactions. Since flames themselves are too complex to define these coefficients, one must find them in less complex environments. Hence the second requirement was to measure the rate coefficients of the elementary reactions experimentally. How?

In spite of the world's reliance on gas-phase combustion for energy, scientists studying gas reactions (gas kineticists) epitomize "small science": small groups (sometimes as small as one person working alone) with small laboratories and small budgets, scattered over the globe, meeting one another personally mostly at the fringes of scientific gatherings whose main purpose requires some input of basic gas kinetics. When the theory of flame reactions was first being developed, gas kineticists studied oxidation reactions in glass bulbs, where the reactions could be made to proceed slowly by maintaining lower tempera-

SCIENCE/SCOPE

The first carbon-dioxide laser rangefinder developed in the U.S. for tactical military use offers several advantages over existing solid-state lasers to improve the first-round accuracy of tank gunners. The new laser, being developed at Hughes for the Army's M1 main battle tank, will penetrate battlefield smoke and dust much better. Because the laser is harmless to the human eye and requires minimal safety restrictions to be operated, gunners will have more training time than they do with the solid-state unit and will become more proficient. The Army is evaluating an advanced development model.

A new adaptive radar, using technology that could be applied in the future to many different weapon control systems, has completed feasibility tests. The radar, called FLEXAR (Flexible Adaptive Radar), uses a multimode transmitter and a programmable signal processor that are now in production, plus a new lightweight, low-cost electronically-scanned antenna. The antenna rotates once each second while the beam electronically scans up and down and back and forth. Waveforms are selected automatically to match the environment. Such flexibility enables the radar to adapt its waveform beamwidth and scan rate as needed to acquire and track targets. Hughes developed FLEXAR for the U.S. Navy.

Besides taking pictures of clouds every 30 minutes, a new satellite provides meteorologists with other important information. The GOES-5 spacecraft relays data from more than 1,500 stations that monitor sea ice conditions and water and snow distribution in remote areas, providing flood warning, among other services. It also measures solar winds and detects solar flares and fluctuations in the earth's magnetic field. This data, besides being useful in weather predictions, is used in communications and electrical power distribution. GOES-5 is the second of three Geostationary Operational Environmental Satellites built by Hughes for the National Oceanic and Atmospheric Administration.

Hughes Industrial Electronics Group offers the advantages and opportunities of a small company backed by the resources of a $2-billion company. Our facilities are in the Southern California communities of Carlsbad, Irvine, Newport Beach, Torrance, and Sylmar. Our programs incorporate 34 different technologies. They include silicon and GaAs semiconductor technologies, fiber optics, microwave and millimeter-wave communications, microprocessors, lasers, and solar cells. Send resume to B.E. Price, Hughes Industrial Electronics Group, Professional Employment, Dept. SSA, P.O. Box 2999, Torrance, CA 90509. Equal opportunity employer.

Better and timelier weather forecasts will be possible when a microwave sensor is launched aboard a military satellite in the mid-1980s. The instrument will tell how hard rain is falling in a specific area rather than simply how much has fallen over a wide area within 24 hours. It also will determine wind speed, atmospheric water content, soil moisture, and sea ice conditions. Because the satellite will follow a low polar orbit, the sensor will gather important data on the little-studied polar regions and oceans. Hughes will soon deliver the prototype Special Sensor Microwave/Imager to the U.S. Air Force.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA 90230

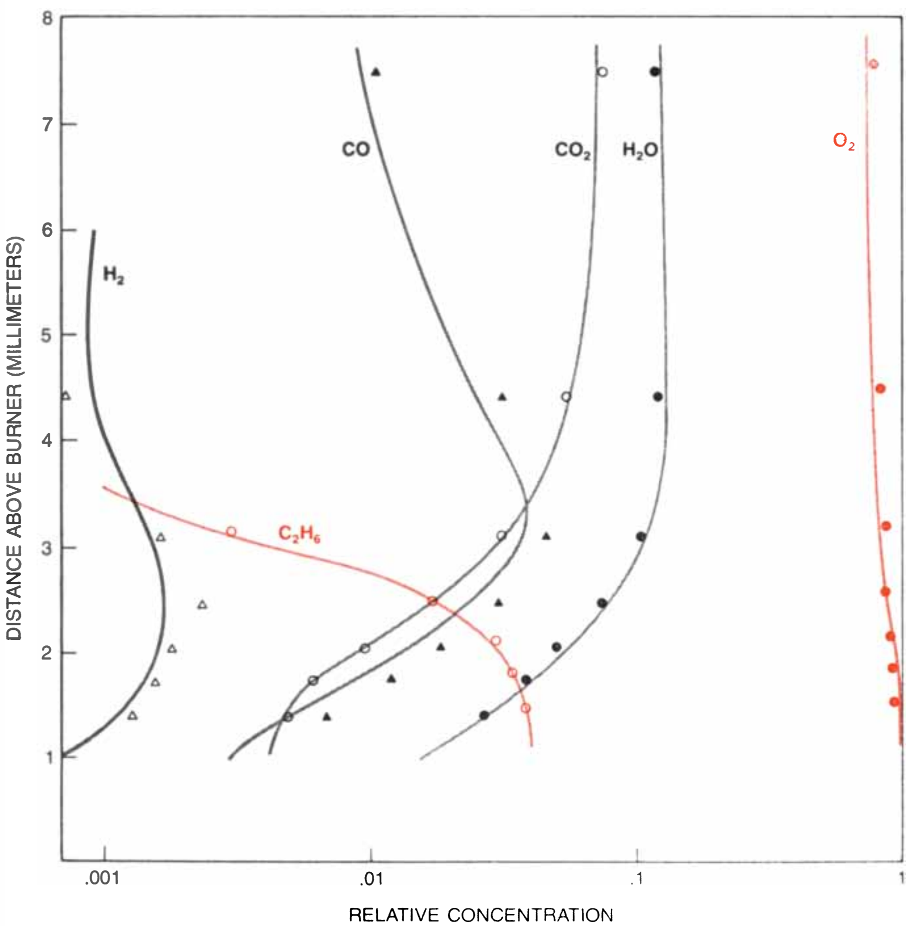
(213) 670-1515 EXTENSION 5964

tures and pressures than normally prevail in flames. As the result of long years of painstaking work it became possible to identify a large number of elementary reactions in this way and to deduce rate coefficients for them. Virtually all this effort, however, bypassed the central chain reactions, because they are too fast to be studied by such methods.

The development of sophisticated laboratory techniques has made it possible to study even the fastest elementary reactions. No single experimental technique and no single laboratory has dominated the field; the new knowledge of elementary combustion reactions was generated by many investigators, working with various techniques. To give the flavor of this research I shall describe three experiments based on three techniques that have been responsible for major contributions. The set of three elementary reactions studied in these experiments constitutes the main chain reaction of the hydrogen-oxygen flame.

A shock wave traversing a gas heats it rapidly. The temperature can jump thousands of degrees Kelvin in a billionth of a second. In a shock tube, a laboratory device in which shock waves of controlled strength are passed through a test gas, chemical reactions resulting from shock heating can be monitored by fast analytical methods. In 1966 David Gutman and Garry L. Schott of the Los Alamos Scientific Laboratory used a shock tube to generate shock-heated mixtures of hydrogen and oxygen that were highly diluted with the chemically inert gas argon in order to maintain a constant temperature during the reaction. The growth of the chain centers could be followed by fluorescence spectroscopy on a time scale of microseconds.

Gutman and Schott provided an analytical probe by adding to the test gas some carbon monoxide, which emits a blue luminescence with an intensity proportional to the concentration of oxygen atoms present. The rate at which the intensity of the luminescence increased was extrapolated to an O_2-to-H_2 ratio of zero, where the rate of chain branching would be governed exclusively by the rate of the reaction $H + O_2 \rightarrow OH + O$. The result was a determination over the temperature range from 1,100 to 1,700



COMPUTED FLAME PROFILES (*curves*) **can be compared with laboratory measurements (*data points*) for low-pressure laminar flames. The profiles shown here were computed by Jürgen Warnatz of the Technical University in Darmstadt for a reaction mechanism with 58 elementary reactions, taking into account the diffusion of all 20 substances involved in the overall reaction. The experimental data, for an ethane-oxygen flame burning at one-tenth atmospheric pressure, were obtained by Robert M. Fristrom, William H. Avery and C. Grunfelder of Johns Hopkins University. Laser analyses of similar flames by James H. Bechtel and his co-workers at the General Motors Research Laboratory have confirmed the mass-spectrometer profiles.**

degrees K. of the rate coefficient of the elementary reaction responsible for chain branching in hydrogen-oxygen flames.

Reactive atoms can be made at low temperatures by decomposing molecules in electric discharges. Fast reactions of atoms produced in this way can be studied by mixing them with other gases and flowing the mixture rapidly past an observation station downstream. In 1968 Arthur A. Westenberg and Newman deHaas of the Johns Hopkins University Applied Physics Laboratory made a small amount of atomic oxygen in a microwave discharge of molecular oxygen diluted with the inert gas helium. The mixture of atomic oxygen and helium was added to a stream of hydrogen flowing down a tube maintained at a desired reaction temperature. Downstream from the mixing point the loss of oxygen atoms due to reaction with hydrogen molecules was measured with an electron-spin-resonance spectrometer. The result was a determination over the temperature range from 500 to 900 degrees K. of the rate coefficient of the second elementary reaction of hydrogen-oxygen flames, namely $O + H_2 \rightarrow OH + H$.

Another way to decompose stable molecules is with ultraviolet radiation. In 1980 Frank P. Tully and Akkihebbal R. Ravishankara of the Georgia Institute of Technology developed an apparatus in which a small amount of water vapor in hydrogen at a desired reaction temperature could be irradiated with an intense flash of ultraviolet radiation. The hydroxyl radicals created by the flash could then react with the hydrogen molecules in the third elementary reaction of the hydrogen-oxygen flame, namely $OH + H_2 \rightarrow H_2O + H$. The decay in the concentration of OH was followed by ultraviolet analytical spectroscopy. The result was a determination of the rate coefficient of this reaction over the temperature range from 500 to 1,000 degrees K.

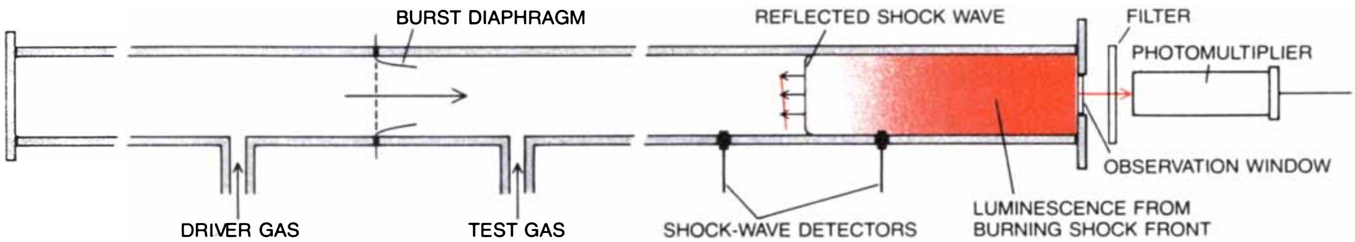
Dozens of experiments such as these have collectively provided chemists with a set of rate coefficients for the elementary reactions of the hydrogen-oxygen flame, thereby making it possible to model these flames accurately on a computer. In addition to the three elementary reactions discussed above some 20 more, mostly concerned with the roles of the secondary intermediates hydrogen peroxide (H_2O_2), the hydroperoxyl radical (HO_2) and ozone (O_3), can be included to construct models that describe even the minor details of the hydrogen-oxygen reaction.

Hydrogen-oxygen flames are scientifically interesting as prototype combustion processes, but their only application outside the laboratory is in rocket engines. Practical flames burn hydrocarbons. What is known about them?

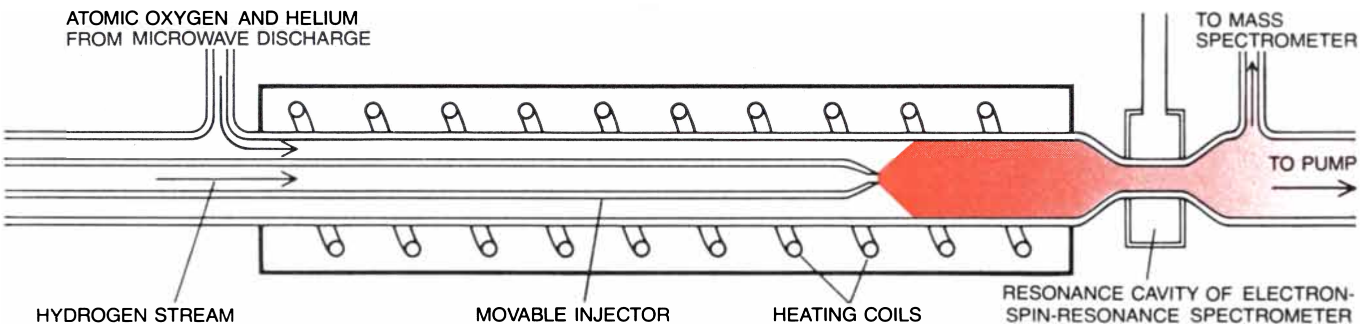
A current chemical description of a hydrocarbon flame would resemble a book with an almost complete introduction (the reactions of the fuel molecules themselves) a complete last chapter (the oxidation of H_2 and CO) and only fragments of everything in between. The number of chapters in the book would increase with the number of carbon atoms in the fuel molecules, except that the simplest fuels, those with one or two carbon atoms, share common reaction pathways and therefore have the same combustion mechanism. Chapter headings could be taken from the fuels and the known intermediates: CH_4, CH_3, CH_2, CH, C_2H_6, C_2H_5, C_2H_4, C_2H_3, C_2H_2, C_2H, CH_3O, CH_2O, CHO, C_2H_4O, C_2H_3O, C_2H_2O....

Trial models of hydrocarbon flames are made by combining the elementary reactions an investigator feels are important, employing both measured and estimated rate coefficients. They are then tested against a variety of experimental facts: flame-propagation speeds, shock-tube ignition measurements, flow-reactor molecular-concentration profiles and so on. Even for the smallest hydrocarbons, however, the list of elementary reactions considered important for one reason or another soon has more than 100 entries, and the only way to keep track of the process is with the aid of flow diagrams.

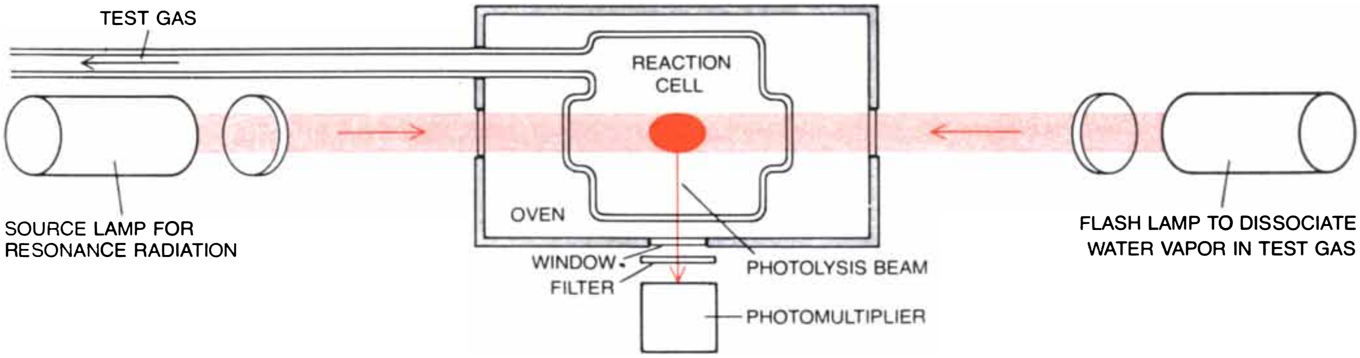
The details of hydrocarbon-flame models are still in dispute, but many general features are clear. First, most of the elementary reactions are types



SHOCK-TUBE EXPERIMENTS are carried out by heating a test gas with a shock wave passing down a long tube and then observing the ensuing reaction in the hot test gas. The apparatus shown was used by David Gutman and Garry L. Schott of the Los Alamos Scientific Laboratory in 1966 to determine the rate coefficient of the reaction $H + O_2 \rightarrow OH + O$. The test gas was a mixture of diatomic molecules—hydrogen (H_2), oxygen (O_2) and carbon monoxide (CO)—highly diluted in the inert gas argon. The shock wave was generated by bursting the thin metal diaphragm separating the high-pressure driver gas (in this case H_2) from the low-pressure test gas. The chemical reaction began when the shock wave was reflected from the end wall of the tube, heating the test gas in a billionth of a second to the reaction temperature. The intensity of the luminescence from the $CO + O$ reaction, which is proportional to the concentration of oxygen atoms in the test gas, was recorded by a photomultiplier tube through a window in the end wall of the shock tube. The rate of the chemical reaction, measured on a scale of microseconds, was deduced from the rate of increase in the intensity of the luminescence.

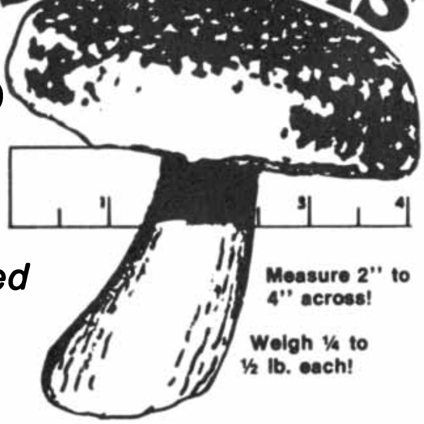


DISCHARGE-FLOW EXPERIMENTS make it possible to study fast reactions by streaming the reacting gas at high speed past a fixed observation point. In this experiment, done in 1968 by Arthur A. Westenberg and Newman deHaas of Johns Hopkins, atomic oxygen was generated in a microwave discharge and was mixed with molecular hydrogen flowing through a heated tube. The rate coefficient of the reaction $O + H_2 \rightarrow OH + H$ was determined by measuring the concentrations of the reactants and the products with the aid of an electron-spin-resonance spectrometer situated downstream from the mixing point. Measurements were also made by mass spectrometry.



FLASH-PHOTOLYSIS EXPERIMENTS are done by initiating the reaction with an intense flash of ultraviolet radiation. The apparatus depicted here was used by Frank P. Tully and Akkihebbal R. Ravishankara of the Georgia Institute of Technology in a 1980 study of the reaction $OH + H_2 \rightarrow H_2O + H$. Hydroxyl radicals (OH) were obtained by the photodissociation of a trace of water vapor inside the heated reaction cell by means of a secondary flash lamp (*right*). The reaction rate was determined by monitoring the intensity of the fluorescence from a small fraction of the hydroxyl radicals that had been excited to a radiating energy state by the primary flash lamp (*left*).

GIANT GOURMET MUSHROOMS LIKE THESE INDOORS YEAR 'ROUND'



the majestic Shiitake mushroom - revered by gourmets for its flavor, revered in the Orient as the Elixir-of-Life

by Hal Taub (At Lovin' Spoonful, I'm Chief Cook & Bottle Washer)

Like most Americans, I love mushrooms. Mushrooms on steak...in an omelet...stuffed mushrooms...as part of my salad...mushroom gravy—you name it, I eat 'em up.

When I can afford it!

With the price of mushrooms *mushrooming* to over $2 a pound for the small button type—when you can find them—and the fancy dried imported kind going for $20-$40 a pound and more, I think twice before indulging myself. Is it any wonder that I've even tried growing them myself? What I got for my labors was a handful of tiny buttons, along with some nasty comments and dirty looks from my neighbors. (Our American or "button" mushrooms require large amounts of manure to grow.)

A TASTE EXPERIENCE

Do I love my food? Let's just say that I very seldom miss a meal. Recently, on a business trip to California, I was treated to a business lunch at an absolutely delightful restaurant in Beverly Hills. Of course I ordered a mushroom salad. It was incredible! I had never tasted anything like it before. Not even the imported mushrooms came close. I can only describe the flavor as being somewhere between filet mignon and lobster! I not only ate my salad, I ordered two more to boot. In fact I almost OD'd on mushroom salad!

THEY ARE CALLED SHIITAKE'S

That was my introduction to the Shiitake Mushroom. Let me tell you, I did not leave that restaurant without learning their source. I discovered they were being raised in very limited supply by a Chinese American botanist, Dr. Henry Mee. I called Dr. Mee thinking I could tote a few pounds home with me. He most graciously invited me out to his facilities. I went to buy mushrooms, but instead, received an education.

ELIXIR OF LIFE

The first Shiitake spawned during the misty era of a hundred million years ago. Early Chinese sages attributed great powers to the Shiitake and it was often called the Emperor's Food. In ancient China and Japan, it was eaten by royalty to fend off old age, revered as an aphrodisiac and fought over by Japanese warriors who fiercely guarded the growing sites.

In their natural habitat, it takes two years to bring a Shiitake crop to harvest. They grow on oak logs in the remote mountain forests of Japan. After 25 years of study and labor Dr. Mee has developed a method that cuts the time down to 45 days. He now produces some 100 pounds daily, of which the entire crop is taken by a relatively small handful of gourmet restaurants and shops.

MORE THAN I BARGAINED FOR

Rather than sell me a few pounds of mushrooms, Dr. Mee suggested I grow my own. He had perfected his process to the point where he claimed anyone who could water a house plant could enjoy fresh, luscious "Shiitake" mushrooms. Simply stated, he simulates their natural habitat by producing a "log" fabricated of 100% sterilized organic plant material, with nothing artificial, and no chemicals added. The log is then innoculated with pure culture of the "Shiitake" mushroom spore...and then cured or "aged" to hasten fruition under home environment with the addition of nothing but water. When Dr. Mee said I would not require any manure or fertilizer of any kind, I decided to give it a try.

SIMPLE AS A.B.C.

The instructions were simple. Start by soaking the log in water for 24 hours. Then simply place the tree in its own wooden planter-stand and set on a foam rubber pad, which is supplied with each log and acts as a moisture "reservoir". After that just mist it once a day. And, unlike buttons, Shiitakes thrive in daylight.

INCREDIBLE RESULTS

In only 5 days I actually saw mushrooms start budding, 10 days later I picked my first giant Shiitake. One month later I had enough for not only myself but my friends as well. What's more, Dr. Mee has informed me that I can expect the log to keep producing for the next 10 to 12 months. If you're growing more mushrooms than you can use, simply store the tree in the frig (or outdoors if the weather is cool) and it will stop growing. When you want more mushrooms, just place it at room temperature and it will start producing.

NUTRITION

With a virtually unlimited supply of my favorite food I've become something of an expert. Mushrooms are nature's unique low cal fat-free food. One pound contains fewer calories than a single apple! Shiitake mushrooms have more than twice as much protein and fiber as common button mushrooms, almost 3 times the minerals. Calcium, Phosphorous and Iron are present in large quantities, as are high levels of B Vitamins and Vitamin D2.

A FEAST

All of this nutrition stuff is great. But the eating is even better! One ounce of Shiitake will equal the flavor of an entire pound of buttons. They are super meaty, super mushroomy in taste, succulent, heady and 100% edible from cap to stem.

One of two slices turn an ordinary pot roast into a gourmet delicacy...an ordinary salad into an extraordinary taste sensation...a gravy into a nectar for the gods. And spaghetti; let me tell you about my spaghetti. I serve it with a mushroom meat sauce that is truly memorable!

My guests insist that Julia Child had come in to cook for me!

AM I SELFISH?

My friends (and my wife) have accused me of being selfish. I admit to being somewhat of a miser when it came to sharing my mushroom crop. It seems every time I gave my friends a super-size Shiitake, they would come back and pester me for more. I finally had to ask Dr. Mee for additional mushroom logs to save my sanity and several valued friendships.

Dr. Mee is now producing a limited number of Shiitake Mushroom Log kits. We at the Lovin' Spoonful are fortunate in being chosen to introduce it to the general public. As a measure of our enthusiasm and our confidence we offer it to you at our risk. If the Shiitake is not every bit as delicious as we claim, or for any reason, you are not satisfied with the production of the log, return it for a full refund of purchase price. The cost is only $19.95, complete with everything you need to grow a bumper crop. We also include a selection of fabulous mushroom recipes. I guarantee it will be the most delicious investment you will have ever made.

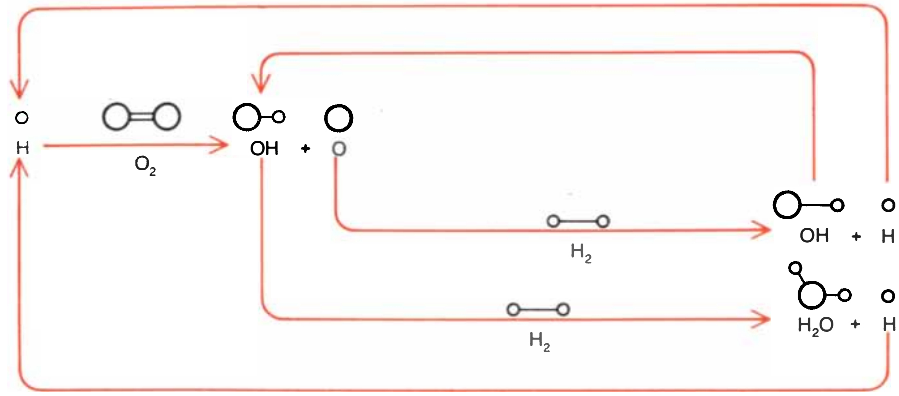
To order use our coupon. Even better, use your credit card and our handy toll free "800" number. Remember you are under no obligation. Order today!

ORDER TOLL-FREE
Call 800-228-2606 Operator 152
In Nebraska call 800-642-8777
Our telephone lines are open
24 hours a day, 7 days a week

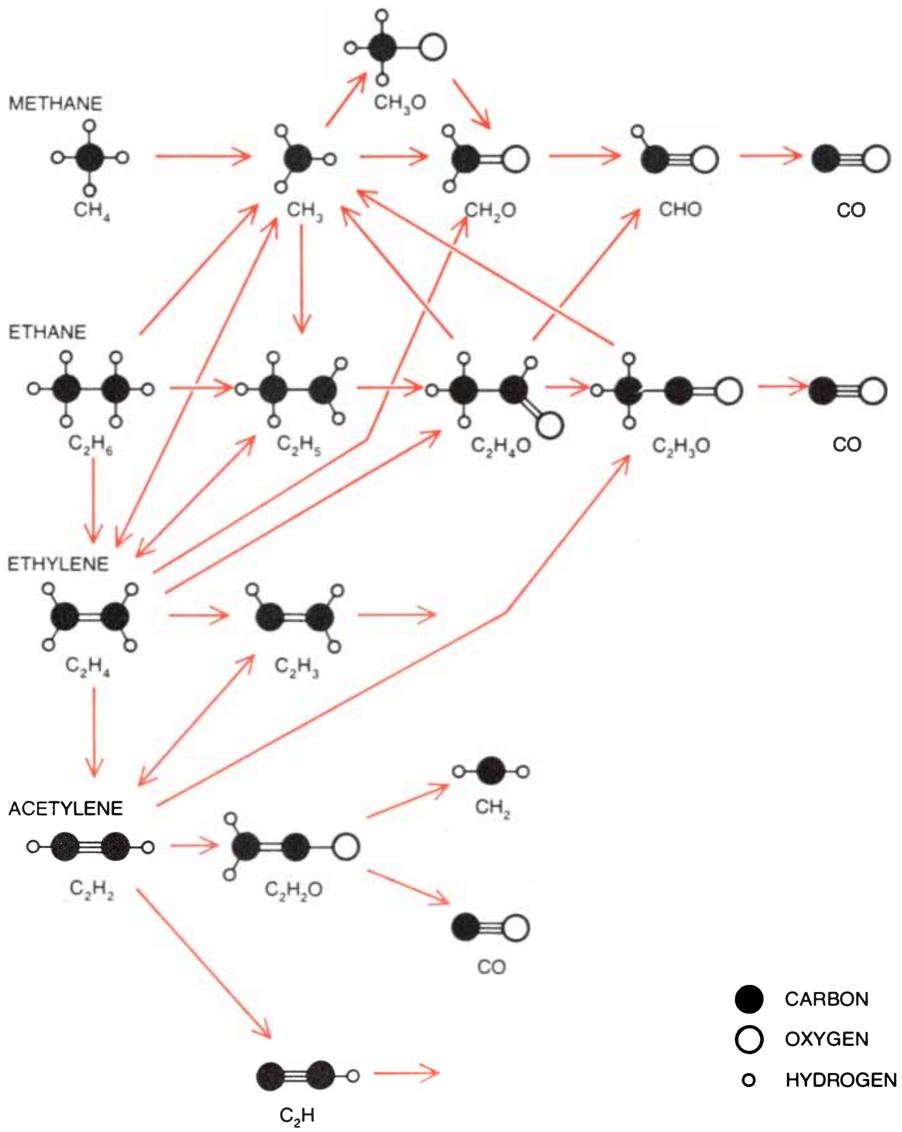
The Lovin' Spoonful™ Dept. 152
28 Durham Dr.,
Dix Hills, NY 11746

Okay, Hal, tasting is believing! Please send ________ Shiitake Mushroom Log(s) @ $19.95, plus $2.00 postage and handling. I understand that I am under no obligation and if for any reason I am not satisfied with my Shiitake Mushroom Log I can return it within 60 days for a full refund of purchase price.

USE YOUR CREDIT CARD
☐ American Express ☐ Diners Club ☐ Visa
Master Charge Exp. Date ________
Charge # ________________
Signature ________________
(Please Print):
NAME ________________
ADDRESS ________________
CITY ________________
STATE ________ ZIP ________
N.Y. residents add sales tax.
©The Durham Group, Inc. 1980



ELEMENTARY REACTIONS of a hydrogen flame involve the intermediate substances H, O and OH. This set of reactions exhibits a typical "chain branching" effect: for each hydrogen atom entering the sequence three hydrogen atoms are generated. This phenomenon is what gives combustion reactions their explosive speed. The rate coefficients of all these reactions have been measured. At low temperatures and high pressures there are other reactions that produce and consume hydrogen peroxide (H_2O_2), the hydroperoxyl radical (HO_2) and ozone (O_3).



HYDROCARBONS BURN by a complex network of combustion routes involving so many elementary reactions that a flow diagram is needed to keep track of the overall course of the reaction mechanism. This comparatively simple flow diagram shows the major transformations that take place in the combustion of the small hydrocarbons methane (CH_4), ethane (C_2H_6) and ethylene (C_2H_4). The intermediates O, H, H_2 and OH and the combustion products CO_2 and H_2O are omitted from the diagram for the sake of clarity. Also not shown are the further reactions of acetylene (C_2H_2) and the reactions of the intermediate hydrocarbons C_2H, C_2H_3 and CH_2, some of which lead to the formation of larger hydrocarbons and soot.

long known in chemical kinetics. Atom-transfer reactions such as $OH + C_2H_6 \rightarrow H_2O + C_2H_5$ constitute the main attack on hydrocarbon fuels. Larger fuel molecules, however, usually do not reach the reaction zone intact, because they are decomposed by heat faster than they diffuse. The commonest type of decomposition involves the loss of two carbon atoms at a time, as in the decomposition of octane by the reaction $C_8H_{18} \rightarrow C_2H_5 + C_6H_{13}$. Most intermediates with an odd number of hydrogen atoms, such as the C_2H_5 radical, quickly lose one of them to form intermediates in the class of hydrocarbons called olefins, in this case ethylene (C_2H_4). Others may lose hydrogen atoms to form intermediates in the class called ethynes; a typical product of this kind is acetylene (C_2H_2).

Olefins and ethynes have been found to undergo some kinds of elementary reactions in flames that are not known in traditional chemical kinetics. For example, Jürgen Warnatz of the Technical University in Darmstadt has recently concluded from modeling the speed of acetylene flames that CH_2 radicals that are formed when oxygen atoms react with acetylene by the reaction $O + C_2H_2 \rightarrow CH_2 + CO$ can fly to pieces in reacting with molecular oxygen by the reaction $CH_2 + O_2 \rightarrow CO_2 + H + H$. Such complex transformations of chemical bonds in single elementary reactions rarely occur except in flames.

Computer models for the combustion of methane (CH_4) already include more than 100 elementary reactions. Is there any hope, then, of constructing one for octane (C_8H_{18}), even assuming the availability of vast computer power? The answer is that there is no hope at all if the model is to take explicit account of all the intermediates: the number of elementary reactions required increases geometrically with the size of the fuel molecule and becomes excessive long before one reaches octane. If the octane-combustion model is only intended to account quantitatively for the central chemical events at the molecular level, however, present-day computers could handle the job by grouping the intermediates by their structure and the reactions by their type. Whether the rate coefficients for such a model can be established in the laboratory remains for the next generation of combustion chemists to discover.

Combustion adds pollutants to the atmosphere: oxides of nitrogen and sulfur, incompletely burned hydrocarbons and particulates such as soot. On a more or less pragmatic engineering basis the suppression of pollutant emissions can be achieved by adjusting the conditions of combustion, by selecting low-nitrogen and low-sulfur fuels, by pretreating fuels or by posttreating combustion gases. Scientifically the goal is


Dance has a new partner.

to discover what elementary reactions create or remove pollutants.

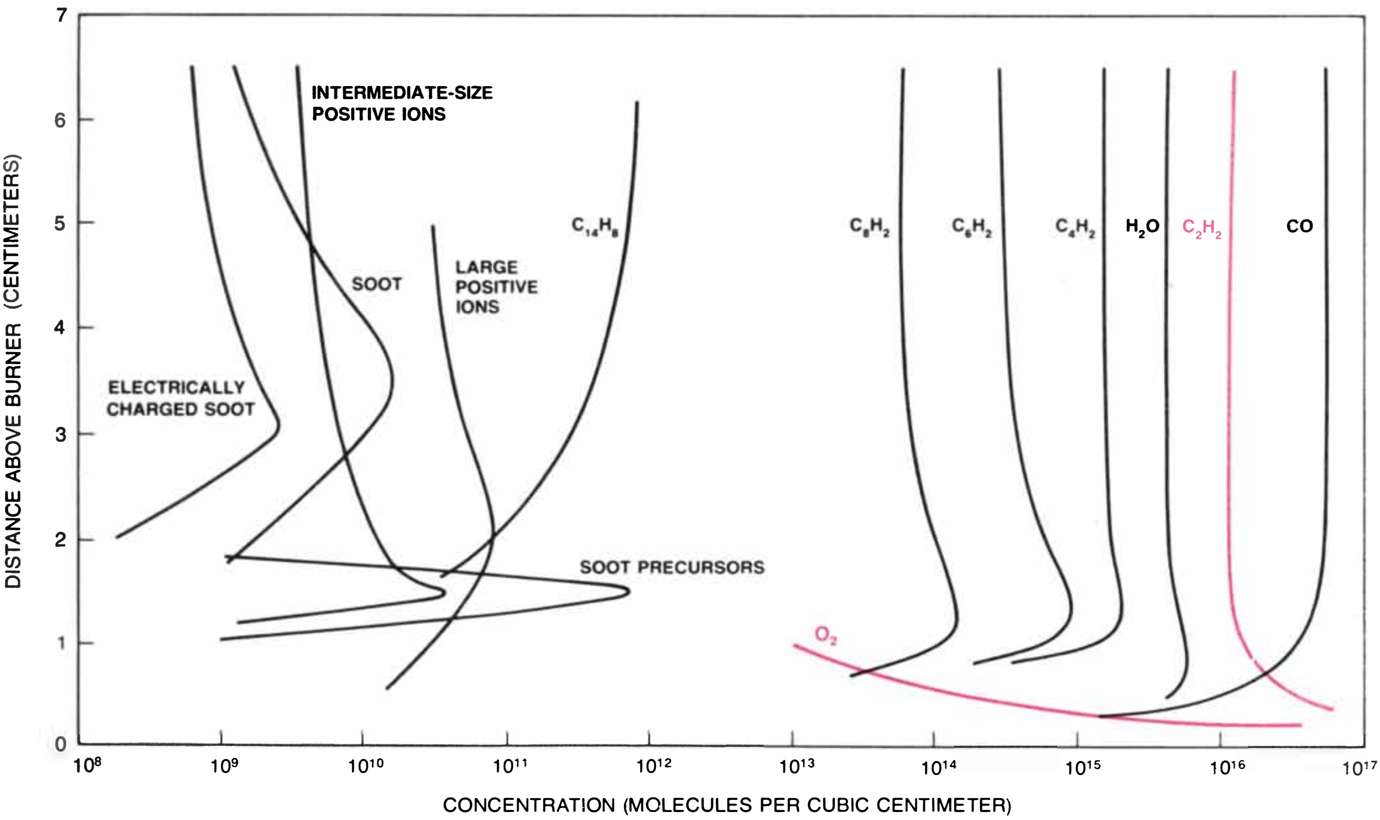
The best-understood aspect of combustion-generated pollution is the formation of oxides of nitrogen with the generic formula NO_x in very hot flames (at temperatures above 2,000 degrees K.). Nitric oxide (NO) is apparently the only oxide of nitrogen formed directly in flames; nitrogen dioxide (NO_2) arises later through slower reactions with atmospheric oxygen, which take place even at room temperature. In very hot flames molecular nitrogen from the air loses its inert character and is split by oxygen atoms in the reaction $O + N_2 \rightarrow NO + N$. Nitrogen atoms rapidly attack molecular oxygen to form nitric oxide and atomic oxygen in the reaction $N + O_2 \rightarrow NO + O$, thereby completing a two-reaction chain known as the Zeldovich mechanism (after Yakov B. Zeldovich of the Institute of Chemical Physics in Moscow, who first proposed it in 1947). Direct shock-tube measurements of the rate coefficient for the $O + N_2 \rightarrow NO + N$ reaction, reported by Jamie Monat, Ronald K. Hanson and Charles H. Kruger, Jr., of Stanford University in 1980, make it possible to model the production of nitric oxide accurately by this route.

At temperatures low enough to suppress the Zeldovich reactions other reactions still generate nitric oxide. Traces of it were shown by Charles Fenimore of the General Electric Research Laboratory to appear very early in flames, leading him to coin the term "prompt NO." What intermediate is energetic enough and plentiful enough then to attack N_2 is unclear. One suggestion is that it is the CH radical, perhaps by the reaction sequence: $O + C_2H_2 \rightarrow CH + CHO$, followed by $CH + N_2 \rightarrow HCN + N$, followed by $N + O_2 \rightarrow NO + O$. Hydrogen cyanide (HCN) is also thought to be an important intermediate when nitric oxide is formed from fuel nitrogen (the generic term for a variety of nitrogen-containing compounds found in small concentrations in petroleum). The reactions leading from hydrogen cyanide to nitric oxide are not yet known. Most of the nitric oxide generated early in flames never escapes into the atmosphere, however, but is converted into molecular nitrogen. How? No one knows. There is obviously ample room for further research on the formation and removal of NO_x compounds in flames.

Basic understanding of reactions involving nitric oxide has led to an ingenious method for removing nitric oxide from postcombustion gas. The method, which has been developed into a commercial process by Richard K. Lyon of the Exxon Research and Engineering Corporation, calls for the injection of ammonia (NH_3) into the postflame gas at a particular point in the cooling process. This leads quickly to the formation of NH_2 radicals, mainly by the reaction $NH_3 + OH \rightarrow H_2O + NH_2$. One of the elementary reactions of NH_2, characterized in 1972 by Manfred Gehring, Karlheinz Hoyermann, Helmut Schacke and Jürgen Wolfrum of the University of Göttingen, is $NH_2 + NO \rightarrow N_2 + H_2O$. In this subsequent step both of the nitrogen atoms in the reactant molecules end up in molecular nitrogen, effectively eliminating the nitric oxide.

A sulfur compound entering a flame quickly forms the stable sulfur dioxide molecule (SO_2), since SO, the sulfur analogue of NO, reacts directly with oxygen by the reaction $SO + O_2 \rightarrow SO_2 + O$, even at room temperature. Chemists studying sulfur reactions that form SO and SO_2 have the problem that sulfur in the gas phase generates a bewildering array of reactive and stable compounds, all of them acidic and all of them unpleasant additions to the atmosphere. There is thus little practical incentive to pursue the identification of specific elementary reactions of sulfur in flames. Pollution control for sulfur means removing all forms of the element prior to combustion or removing all the acidic oxides of sulfur later.

Finally there is soot, the formation and oxidation of which are currently the



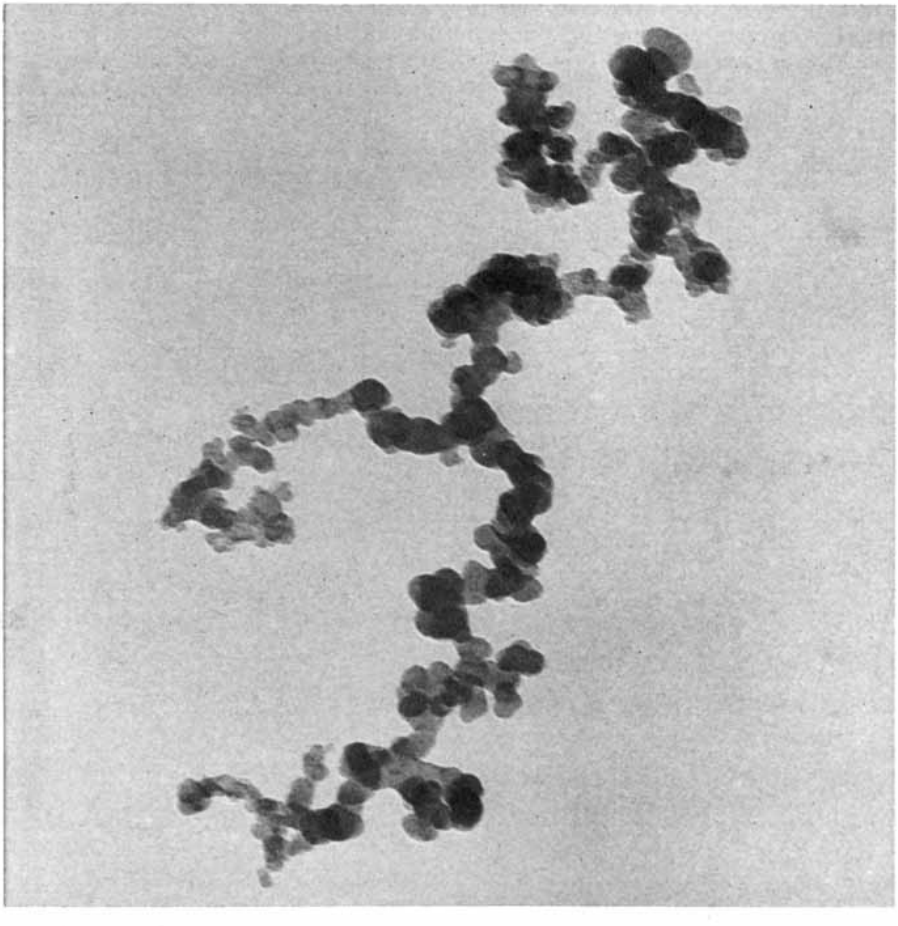
MOLECULAR COMPOSITION of a typical soot-producing laboratory flame varies as a function of distance above the burner. This set of logarithmic concentration profiles, for an acetylene-oxygen flame, combines the results of measurements made by Ulrich Bonne, Klaus H. Homann and Heinz-Georg Wagner of the University of Göttingen, Douglas B. Olson and Hartwell F. Calcote of the Aerochem Research Laboratories and Jack B. Howard and his co-workers at the Massachusetts Institute of Technology. The group of intermediates labeled "soot precursors" consists of a variety of hydrocarbons with between 20 and 50 carbon atoms. Unlike polyaromatic, or multiple-ring, hydrocarbon molecules (such as $C_{14}H_8$), whose flame profiles continue to rise, the soot precursors vanish when soot appears.

two most actively studied processes in combustion chemistry. At one time the main use of petroleum was for kerosene lamps, which require the intermediate formation of soot to provide luminescence and the subsequent oxidation of soot to avoid the blackening of the lamp chimney. Nowadays light can be produced more conveniently by electricity, so that the soot in most flames is just a nuisance. The factors that enhance the production of soot are well known: high fuel-to-air ratios, fuel containing compounds with low hydrogen-to-carbon ratios, erratic ignition and poor mixing of fuel and air. It appears that much of the future supply of liquid fuels will come from oil shale and liquefied coal, which are intrinsically low-hydrogen fuels. Can they be burned on a large scale without burying us all in soot?

Soot is formed in three stages. In all hydrocarbon flames there is a decrease in the hydrogen-to-carbon ratio of the hydrocarbon molecules and radicals as they are successively decomposed by heat and chemical attack. The last of them are typically molecules such as acetylene (C_2H_2) and fragments such as the ethynyl radical (C_2H). In a soot-producing flame such molecules and radicals are not oxidized; in the first stage of soot formation they mostly polymerize to form polyacetylenes and their radicals by reactions such as $C_2H + C_2H_2 \rightarrow C_4H_2 + H$, followed by $C_2H + C_4H_2 \rightarrow C_6H_2 + H$. Other large molecules characteristic of a soot-producing flame are a variety of polyaromatic compounds (in which the carbon atoms are arranged in multiple rings) and positive hydrocarbon ions.

After a characteristic delay time some of these molecules vanish and in the second stage a soot aerosol appears: spherical particles of carbon containing varying amounts of hydrogen and trapped hydrocarbons, depending on the flame. The aerosol particles give sooty flames their opacity and are therefore responsible for their yellow glow. The spheres cannot grow beyond about a tenth of a micron in diameter, because of the depletion of carbon in the flame gas, and so the larger particles that constitute black smoke form in the third stage by the agglomeration of the aerosol particles. Both the particles and their agglomerates anneal as long as they are hot, losing hydrogen and organic material and forming increasingly rigid structures that in some flames begin to resemble the layered structure of graphite. They can also oxidize readily, and in many flames, such as those of candles or well-run industrial furnaces, they vanish completely.

The evidence supporting this picture comes from two sources: analyses of the molecular and ionic composition of flame gases and related soot-forming systems prior to the appearance of a soot aerosol, and microscopic and chemical studies of soot particles recovered from the later stages of the process. The investigator of soot is therefore isolated from two key aspects of the flame: the steps by which hydrocarbon molecules, radicals and ions first nucleate to form an aerosol, and the competition among the annealing, agglomeration and oxidation of aerosol particles later in the flame. The largest molecules that can be detected in flames have some two or three dozen carbon atoms; the smallest particles able to exist as aerosols have perhaps a million. Agreement on a coherent picture of how the molecules become soot particles has not yet been reached. The aerosol particles must be a beehive of chemical activity, the chemical characterization of which may never be achieved.



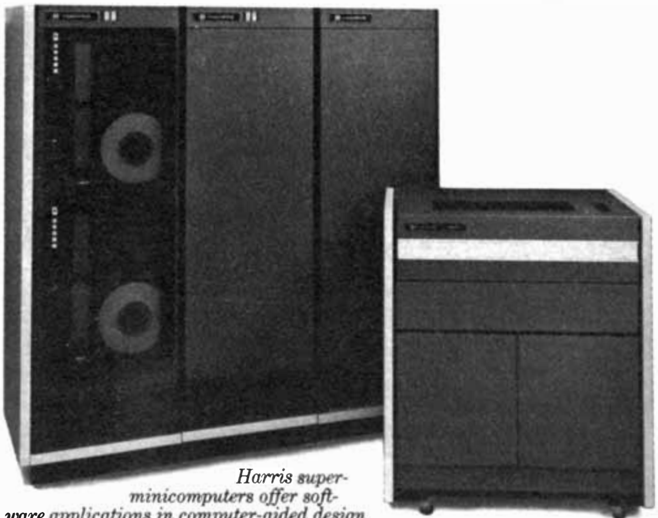
EXTENDED STRUCTURE OF SOOT is evident in a transmission electron micrograph made by Robin Stevenson of the General Motors Research Laboratories. Soot particles of this type are formed in a fuel-rich hydrocarbon flame by the agglomeration of an aerosol of minute carbon spheres containing varying amounts of hydrogen and trapped hydrocarbons. In flames producing a yellow glow but no black smoke the soot aerosol is oxidized by the flame gases faster than it agglomerates. This particular particle was collected on an amorphous carbon substrate from the exhaust of a diesel engine. The image was deliberately underfocused to enhance the contrast. The magnification of the micrograph is some 110,000 diameters.

Petroleum and natural gas will not continue to dominate the energy scene for many more years; alcohol is already in use as a fuel, and other "synfuels" made from coal, biomass, shale, tar sands and garbage may soon be common. Will they be so expensive that they become luxuries, or will noncombustible energy reservoirs such as batteries or fuel cells replace them?

There is good reason to think that liquid fuels will continue to be hydrocarbons and that they will continue to be burned much as they are now at least for mobile energy sources, even if they have to be synthesized. The reason is that no one has discovered any promising substitute for their unique combination of convenience and high energy density. Coal, charcoal and uranium are suitable for burning in stationary power plants, and substitutes for natural gas in space heating may become economical, but liquid hydrocarbons are likely to remain the optimal form of energy storage for powering automobiles and airplanes. The main difference between the petroleum liquids we burn and the synfuels our descendants will burn will probably be that the synfuels will have less hydrogen and hence a greater tendency to produce soot. What to do about that problem is going to be the next challenge of combustion engineers, and how to understand it will be the next challenge of combustion chemists.

HARRIS HAS A SOLUTION FOR YOUR COMPANY'S BIGGEST PROBLEM:

SOLUTIONS FOR YOUR COMPANY'S SMALLER PROBLEMS.



Harris super-minicomputers offer software applications in computer-aided design, simulation and some 300 other areas.

Most big companies are made up of a lot of smaller companies. They're called *departments*. And each has its own way of doing business.

It's the business of management to make sure that the efforts of each department work in concert with the rest. And that's a job Harris can make a whole lot easier.

We offer super-minicomputers for your R&D and technical people that are unmatched in power, performance and reliability.

The people who run your company's data center will appreciate our line of communications-oriented computer terminals (remote batch and interactive). It's the broadest in the industry.

Our data terminals can grow into sophisticated distributed data processing systems.

Your MIS executives will value our distributed data processing systems for their speed, power and unique capabilities. Like multi-host support and word processing options.

Your communications managers will recognize the difference in bottom-line telephone costs our PBX's can make.

We also provide intelligent telephones, a unique private tandem networking capability and business satellite earth stations that can further reduce your telecommunications costs. And that will surely make your accounting department happy.

Our PBX can accommodate from 20 to 1,000 lines.

The fact is, Harris manufactures products that can improve the productivity of people in every one of your departments. What's more, because they're designed for future expansion and compatibility, they can maximize *corporate* productivity, too.

And that's the biggest contribution we can make.

HARRIS

For a copy of the Harris Integrated Information Processing Systems brochure, send your business card or write on your company's letterhead to Jack C. Davis, Group Vice President, Department 3, Information Systems, Harris Corporation, 1025 West Nasa Boulevard, Melbourne, FL 32919. Or phone 305-727-9100.

Quasars as Probes of the Distant and Early Universe

The light from most of these enigmatic objects was emitted 15 billion years ago. Therefore they are a unique clue to how the universe looked when it was only a fourth its present age

by Patrick S. Osmer

Nineteen years after their discovery by Maarten Schmidt quasars are still one of the great enigmas of astronomy. Although their nature remains controversial, their description is not: quasars are starlike objects with a large red shift. The light emitted by them is strongly displaced toward the red end of the spectrum, signifying they are receding at a substantial fraction of the velocity of light. If quasars are as remote as their large red shifts would indicate, they must be much farther away than ordinary galaxies, whose own red shifts betoken the general expansion of the universe. A quasar can be 1,000 times brighter than an entire galaxy of 100 billion stars. The light from the most distant quasars started on its journey when the universe was only a fourth its present age and has taken 15 billion years to reach us.

It is not my intention in this article to discuss the difficult physics of the nature of quasars. Rather I want to concentrate on what quasars can tell us about the distant universe and its condition at an early epoch. I shall adopt the working hypothesis that quasars are the extremely bright nuclei of galaxies otherwise too distant to be observed. Around some faint, relatively nearby quasars one can detect evidence of a galaxy, but the luminous, distant quasars can be distinguished from stars only by their spectral properties. Distinguishing quasars from stars has presented astronomers with a major challenge. A single wide-field photograph of the sky made with a large Schmidt telescope (named for the optical system designed in the 1920's by Bernhard Schmidt) will show at least 200,000 stellar images, of which only a few hundred will be quasars.

The original quasars were found by optical astronomers who were trying to identify celestial radio sources. As radio techniques improved, the positions assigned to the radio sources became increasingly accurate and in many instances pointed to particular stellar images on photographic plates. Because normal stars could not be detected as sources of radio emission with the equipment then available the coincidence of a radio source with a stellar object was a good way to separate quasars from stars. As a result the majority of the quasars first identified were strong radio sources, hence quasi-stellar radio source, or quasar.

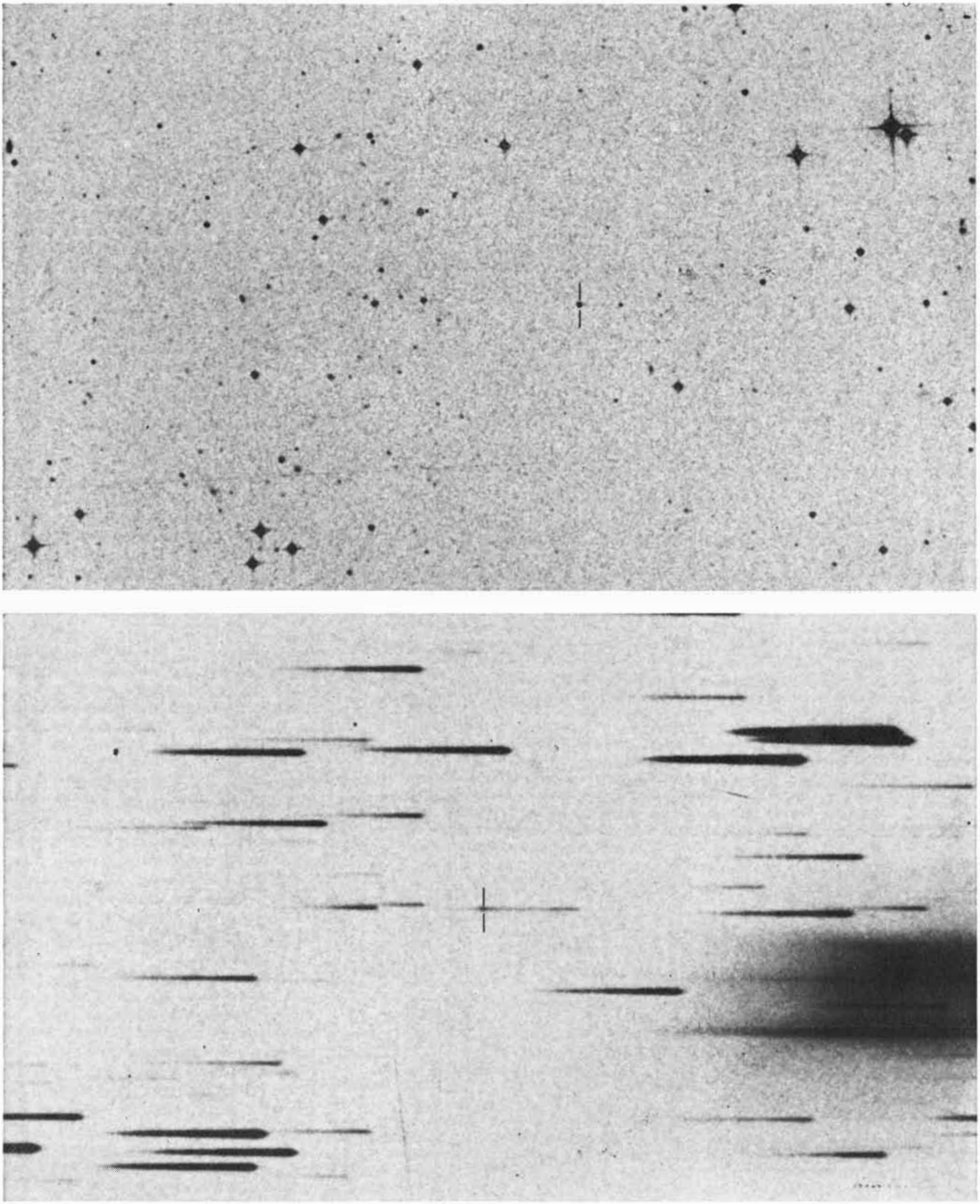
It was soon discovered, however, that a much larger population of quasars were weak radio emitters and therefore had escaped detection by radio telescopes. Allan R. Sandage, working with the 200-inch Hale telescope on Palomar Mountain, found that quasars emitted much more ultraviolet radiation than ordinary stars and therefore could be identified by comparing the stellar images on ultraviolet-sensitive photographic plates with the corresponding images on ordinary plates, which respond chiefly to light at the blue end of the visible spectrum. Sandage showed that quasars that are quiet at radio wavelengths greatly outnumber those with strong radio emission.

Recently a new optical method has been developed for discovering quasars of large red shift that is complementary to the method based on their ultraviolet brightness. The method developed fortuitously from a project Malcolm G. Smith and I began several years ago when we undertook a search for certain galaxies whose spectra exhibit strong emission lines rather than absorption lines, which are more characteristic. The search was conducted with the 60-centimeter Curtis Schmidt telescope at the Cerro Tololo Inter-American Observatory in Chile, in conjunction with two larger reflecting telescopes of 1.5 and four meters. Smith planned to photograph the spectra of faint objects over a wide field by placing a newly acquired thin prism over the entrance aperture of the Schmidt telescope.

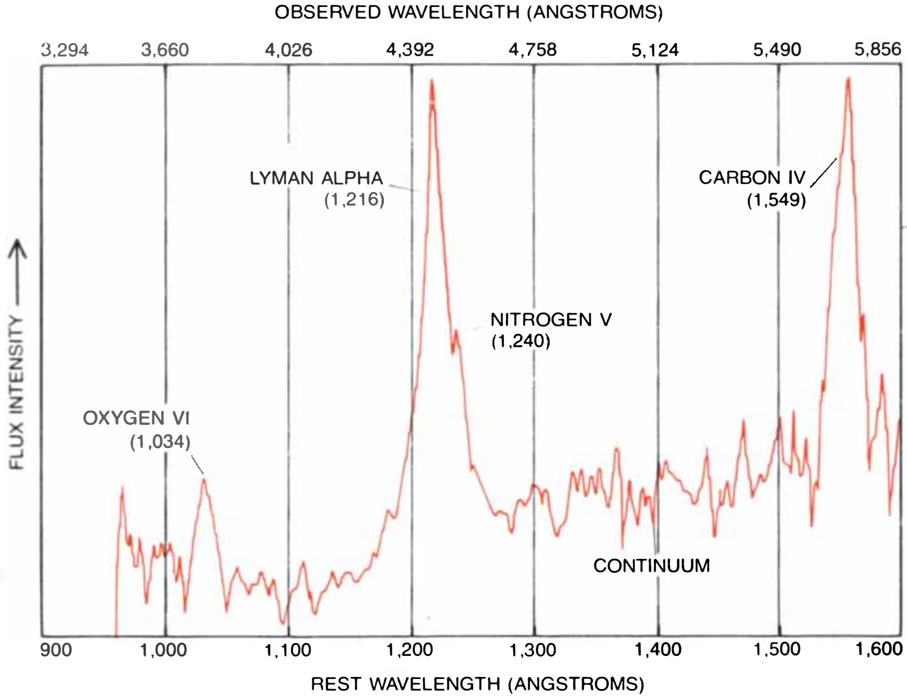
The method of spreading the normally pointlike images of stars into spectra by means of an objective prism has had a long and distinguished history in stellar astronomy. What was new about Smith's method was the combination of a prism of low dispersion (to conserve the light from faint objects) and newly available photographic plates with fine grain and high contrast. This combination, together with the exceptional observing conditions at Cerro Tololo, made it possible to record spectra of fainter objects than had been achieved with a telescope of the Curtis Schmidt's modest size. Once emission-line galaxies had been identified by this method their spectral features could be examined in detail in the high-dispersion spectra yielded by the larger telescopes.

On searching the wide-field Schmidt plates Smith noted that in addition to the emission-line galaxies we had hoped to find there were a few objects, possibly quasars, with emission lines at unexpected positions in their spectra. When we examined these objects more closely in the larger telescopes with a Vidicon (intensified television camera) spectrometer I had been adapting for the project, we learned that they were indeed quasars of large red shift. Smith's technique thus provided a direct and efficient method of distinguishing quasars from stars in wide-field plates. In addition the method was more efficient for distinguishing quasars with a large red shift than the one based on ultraviolet brightness.

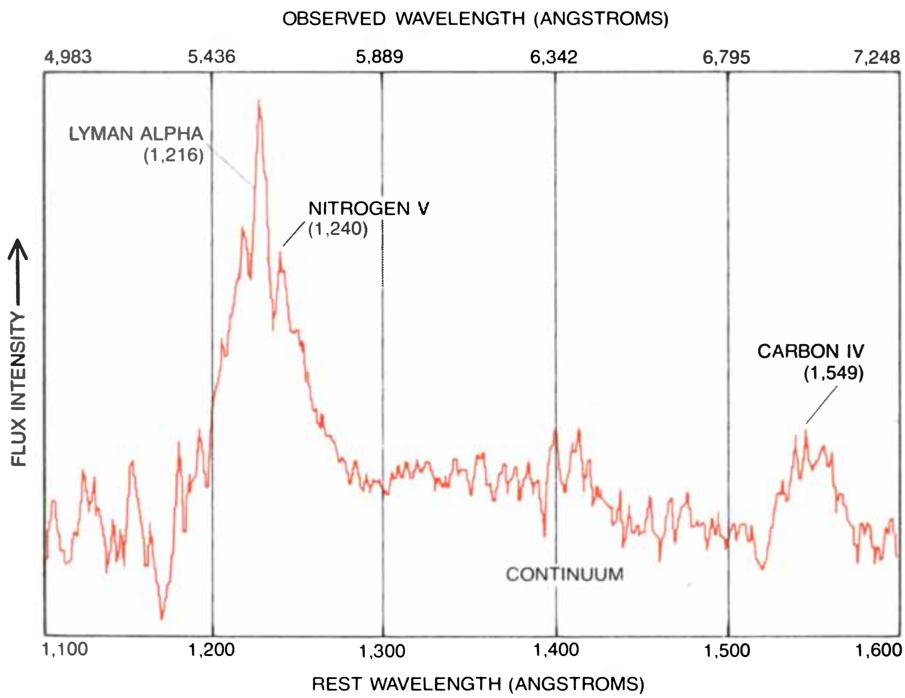
Smith subsequently surveyed a long strip of the southern sky with the Curtis Schmidt telescope. In addition he and Arthur A. Hoag of the Kitt Peak National Observatory extended the survey to fainter magnitudes in small fields of the Schmidt survey by putting a transmission diffraction grating (equivalent to a prism) at the prime focus of the Cerro Tololo four-meter telescope. I continued to examine the quasars uncovered in the two surveys with the



METHOD FOR FINDING QUASARS is illustrated by these two photographs made with the 60-centimeter Curtis Schmidt telescope at the Cerro Tololo Inter-American Observatory in Chile. In a single direct photograph of the sky quasars and stars are indistinguishable because both appear as points of light, as in the negative print at the top of the page. When a thin prism is placed over the aperture of the Schmidt telescope, the image of each object is dispersed into a spectrum, as in the negative print of the same part of the sky at the bottom of the page. In the bottom photograph the longer wavelengths at the red end of the spectrum are at the right and the shorter wavelengths at the violet end of the spectrum are at the left. The spectra of the stars are either devoid of features or exhibit white bands indicating that some of the stellar emission has been absorbed. The quasar, indicated by bars in both photographs, is conspicuous because of a strong emission line, the Lyman-alpha line of hydrogen, that has been shifted from its normal position at 1,216 angstrom units in the far-ultraviolet part of the spectrum to the threshold of the visible spectrum at 3,720 angstroms, corresponding to a red shift of 2.06. This particular quasar, QO 149–397, was one of the earliest discovered by Malcolm G. Smith with the prism on the Curtis Schmidt.



SPECTRA OF QUASARS disclose that the emitting atoms are more highly ionized (stripped of more electrons) than the atoms in nebulas around hot, young stars in our own galaxy. The emission lines are also strongly shifted toward the red end of the spectrum from their rest, or un-red-shifted, wavelength. This spectrum of the quasar QO 453–423 was made with the Vidicon camera system on the 1.5-meter telescope at Cerro Tololo. The strong Lyman-alpha line of ionized hydrogen has been red-shifted to 4,451 angstroms, in the blue part of the visible spectrum, from its rest wavelength of 1,216 angstroms in the far ultraviolet. This corresponds to a red shift of 2.66 (obtained by subtracting 1,216 from 4,451 and dividing by 1,216). The spectra show that emissions from oxygen VI, nitrogen V and carbon IV are comparably red-shifted. Roman numerals are one higher than the number of electrons stripped from the atoms.



LARGEST RED SHIFT of any known celestial object, 3.53, is that of quasar OQ 172. Its spectrum, made with the four-meter telescope at Cerro Tololo, shows the Lyman-alpha line to be shifted from a wavelength of 1,216 angstroms to a wavelength of 5,508 angstroms in the green part of the visible spectrum. The numerous absorption lines to the left of the Lyman-alpha line are caused by gas clouds lying in the line of sight between the quasar and the solar system.

more sensitive Vidicon spectrometer. The combined observations are the basis for the new results I shall describe below.

The spectra of quasars are quite unlike the spectra of all other astronomical objects. The large red shift brings into view regions of the far-ultraviolet spectrum never before recorded by ground-based telescopes. The red-shift value, often designated *Z*, is obtained by subtracting the rest wavelength (the non-red-shifted wavelength) of an emission line from its observed wavelength and dividing the difference by the rest wavelength. The strongest feature in quasar spectra is the Lyman-alpha line of atomic hydrogen, shifted by a factor of as much as 4.5 from its normal position at a wavelength of 1,216 angstrom units in the ultraviolet to about 5,500 angstroms, the wavelength of greenish-yellow light. In such a case *Z* would be 3.5 (5,500 minus 1,216 divided by 1,216). Emission lines of oxygen, nitrogen and carbon with normal ultraviolet wavelengths between 1,034 and 1,549 angstroms are also prominent in the spectra of quasars with large red shifts.

Moreover, in many instances the emission lines are wide, an indication that some of the gas surrounding the quasar is moving at velocities as high as 10,000 kilometers per second. The physical conditions deduced from the intensities of the various lines show that the gas is hotter than the gas in normal nebulas and that the central source in the quasar does not radiate at all like a normal star. For this discussion, however, the main point is that the Lyman-alpha line is the strongest feature in quasar spectra and therefore is the one most easily detected on the Schmidt plates made through an objective prism. For this reason the objective-prism method favors the detection of quasars with large red shifts.

If red shifts are interpreted as velocities of recession, the red shift needed to displace Lyman-alpha radiation to 3,648 angstroms in the near-ultraviolet region of the spectrum, equivalent to a red shift of 2, corresponds to 80 percent of the velocity of light. OQ 172, the quasar with the largest red shift known, 3.53, is evidently receding at 91 percent of the speed of light. On the same scale stars within galaxies have velocities of .1 percent of the speed of light, and nearby galaxies are moving away from us at no more than 1 percent of the speed of light.

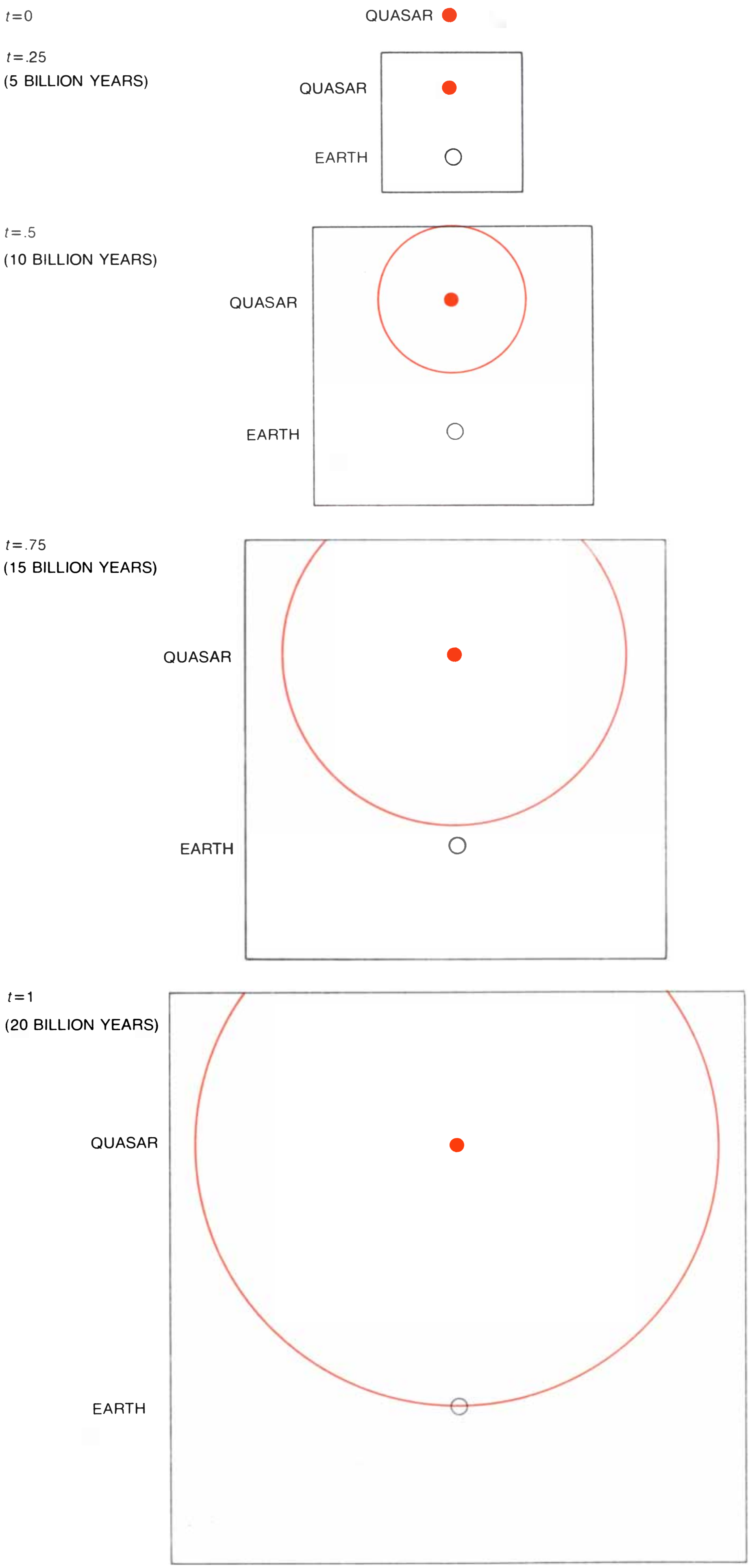
In more than a quarter of a century of measuring galactic red shifts with the 100-inch telescope on Mount Wilson and the 200-inch telescope on Palomar Mountain the largest value recorded by Milton L. Humason was a red shift of .2, obtained in 1949. Another 11 years elapsed before Rudolph Minkowski, also working with the 200-inch telescope, pushed the limit out to .46, a record that stood for many years. Within two years of Maarten Schmidt's discovery in 1963 of the first quasar, 3C 273, with a rather modest red shift of .158, the red-shift "barrier" of 2 was broken. The current maximum value of 3.53 was recorded in 1973.

If the quasar red shifts are interpreted as being due to the expansion of the universe, they are not nearby objects that happen to be moving at high velocity but are extremely distant objects. This is a direct consequence of Edwin P. Hubble's conclusion in 1929, based largely on Humason's observations, that galaxies are receding from us (and from one another) at velocities proportional to their distances. On the basis of the current scale of the universe a galaxy with a red shift of only .01 (or an apparent recessional velocity of 3,000 kilometers per second) will be 200 million light-years away. This is already a considerable distance; the nearest spiral galaxy, the Andromeda galaxy, is only about two million light-years away. Quasars are vastly more distant. The quasars with the largest red shifts approach the horizon, or limit, of the observable universe.

Measuring distances in astronomy has always been a problem. Only the nearest stars have truly been measured by trigonometry, as a surveyor measures distances on the earth; everything beyond those stars has been extrapolated. No one contends that red shifts reveal the absolute distances of the galaxies, but it is generally accepted that the red shifts are a good measure of relative distances.

The concept of an expanding universe implies that an object with a large red shift is observed as it existed long ago. The light now reaching us from a quasar with a red shift of 3 was emitted some 15 billion years ago. The light from a quasar with a red shift of only 1 has been traveling for 10 billion years, or half the age of the universe.

A further consequence of an expanding universe is that the red shift supplies



SCHEMATIC PICTURE of the universe at different epochs illustrates how quasars can supply clues to conditions existing in early epochs. The light now reaching the earth from a quasar with a red shift of 3 was emitted some 15 billion years ago, only five billion years after the hypothetical big bang that initiated the universe's expansion from a dense state (t = 0). At an age of five billion years the universe was only a fourth its present size. The colored circles show the distance covered by the quasar's radiation as the universe expands and the quasar and our galaxy move apart. The radiation finally reaches telescopes on the earth.

a direct measure of the expansion factor. When the light was emitted from a quasar with a red shift of 3, the universe was only a fourth as large as it is now, that is, everything was four times closer together. The distribution of quasar red shifts therefore supplies many clues to the structure and character of the early universe. I hasten to add that the details of the mathematical relations between the observed red shifts and the properties of the universe are somewhat complicated and certainly controversial, given the sweeping assumptions that are drawn from still limited observational data. Nevertheless, certain aspects of the inquiry are relatively independent of the fine points, and we shall concentrate our attention on them.

In the early years of quasar investigations many astronomers hoped the quasars would help to decide which of the many possible hypotheses about the evolution of the universe was correct. That hope was not fulfilled, but in the late 1960's Maarten Schmidt discovered a remarkable property of distant quasars: they are far more numerous in space at large distances than they are in our vicinity. At a red shift of 2, or about 13 billion years ago, their density was 1,000 times greater than it is now. Evidently whatever process gave rise to quasars must have been extremely active when the universe was young and has subsided to practically nothing today.

At the time of Maarten Schmidt's work on the spatial density of quasars the largest red shift known was 2.88. Enormous as this value was compared with what had been expected only a few years earlier, his result suggested that quasars with much greater red shifts should be plentiful. Why were quasars with red shifts of 3 not being detected? Both Schmidt and Sandage called attention to this point and suggested that quasar red shifts might have a limit.

The implications of such a limit were extremely important. The limit implied that for at least one type of object astronomers were seeing to the edge of the universe. It looked as if quasars had formed suddenly in a great burst of activity about 15 billion years ago. Such a concept, of course, would greatly influence ideas about the fundamental nature of quasars. If it is also assumed that quasars represent enormously energetic processes in the center of galaxies, the red-shift limit has important implications for the evolution of galaxies themselves.

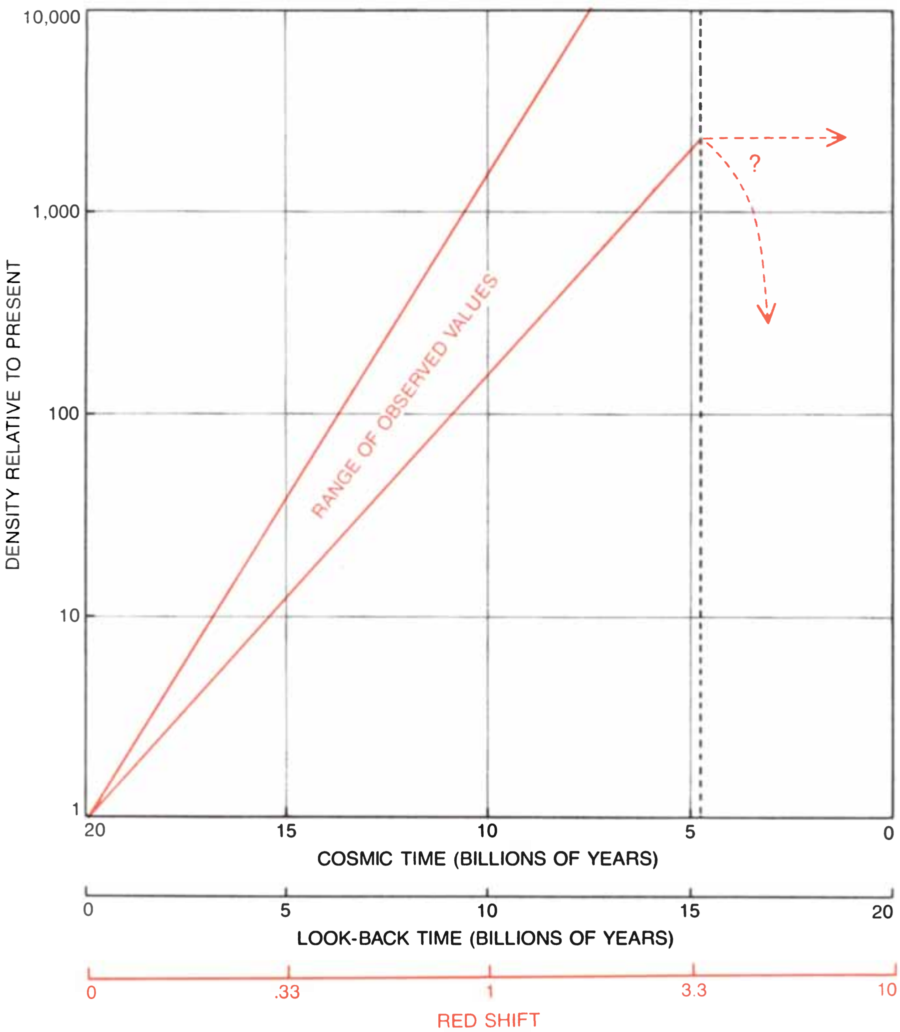
Naturally astronomers kept searching for quasars of ever greater red shift. An obvious concern was that the apparent limit might be the result of some selection effect in the methods of identification, even though the known quasars had been found by quite heterogeneous surveys based on diverse methods. The major point of contention was whether or not quasars with red shifts larger than 2.5 or 3 would continue to be brighter in the ultraviolet than ordinary stars are. The quasar with a red shift of 2.88 had no particular ultraviolet excess, and since its strongest emission line, the Lyman-alpha, is shifted into the middle of the visible spectrum, it and other objects like it would be hard to distinguish from normal stars.

Actually when the quasar with a red shift of 3.53 was found in 1973, it proved to be red in color. It was identified only because radio astronomers had determined its position with high accuracy. If it had been a radio-quiet quasar, it would not have attracted attention. Its discovery, combined with doubts about the completeness of earlier searches, kept the reality of a red-shift limit very much an open question.

The development of our objective-prism method at Cerro Tololo offered a new approach to the problem. Since the method was based solely on the detection of emission lines, it was completely independent of quasar color, ultraviolet or otherwise. Although the method inevitably had biases of its own, at least they were different from those of all other searches.

The objective-prism method was capable of detecting the Lyman-alpha line in quasars as far into the red end of the spectrum as photographic emulsions allow. The emulsion selected by Hoag and Smith for their deep survey with the four-meter telescope would respond to the Lyman-alpha line as far into the red as 6,900 angstroms, equivalent to a red shift of 4.7. Nevertheless, none of the 71 new quasars found in their survey was found to have a red shift greater than 3.45, which was close to, but not surpassing, the previous limit. Suggestive as this result was, Robert F. Carswell of the University of Cambridge and Smith subsequently showed that the Hoag-Smith survey did not have enough sensitivity for red shifts greater than 3.5 to require that their density actually decrease. Therefore a new survey was in order.

I made observations with the Cerro Tololo four-meter telescope in a program optimized to find quasars with red



DENSITY OF QUASARS 15 BILLION YEARS AGO was more than 1,000 times greater than it is today, according to calculations made by Maarten Schmidt, who discovered the first quasar in 1963. The absence of quasars with a red shift larger than 3.53 implies astronomers have probed to the earliest epoch of quasar formation. Some 1,500 quasars are now known.

CENTAURUS CLUSTER OF GALAXIES, one of many such clusters, is shown in this photograph made with the four-meter telescope at Cerro Tololo. The Centaurus cluster, which is about 225 million light-years away, harbors some 250 major galaxies, typically separated by about 700,000 light-years. An understanding of the clustering process should help to show how galaxies formed in the early universe.

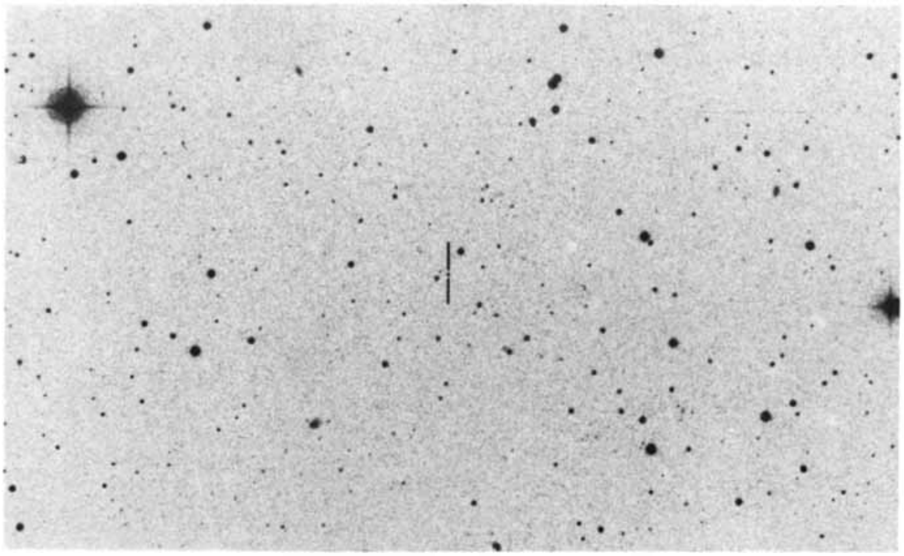
shifts between 3.7 and 4.7, if they existed. Working with a transmission grating of greater sensitivity in the 6,000-angstrom region in conjunction with a filter to eliminate the light of the night sky at wavelengths of less than 5,700 angstroms, I was able to detect quasars three times fainter than Hoag and Smith could in the crucial 5,700-to-6,900-angstrom region. I found 15 objects with emission lines, any one of which could have been beyond the 3.5 limit. Subsequent observations with the Vidicon spectrometer, however, showed that in no instance was the Lyman-alpha line the one detected on the photographic plate. In five objects the line turned out to be from carbon in quasars with red shifts between 2.8 and 3.4; all the remaining objects were either quasars with red shifts near 1 or galaxies with red shifts of about .2. Although these results were disappointing in one sense, they gave confidence that quasars with a red shift of 4 would have been seen if they were there. All the new quasars were discovered by the detection of lines that were weaker than Lyman-alpha, a complete reversal of the previous surveys with the objective-prism method.

My subsequent quantitative analysis showed that the new survey results could indeed be understood only if the space density of quasars beyond a red shift of 3.5 was a factor of three or more below the value seen at a red shift of 3. The limit does not preclude that a few quasars may eventually be found at greater distances, but it does show convincingly that we have seen the turnover point.

Supporting evidence in favor of the red-shift limit is provided by the failure of radio astronomers to find more distant quasars, even though they now have highly accurate positions that allow them to ignore the color of quasar candidates. Recently X-ray astronomers have entered the field with data collected by the X-ray satellite, the Einstein Observatory. Although the satellite's X-ray telescope readily imaged OQ 172, the most distant quasar, and several other quasars with red shifts of about 3.1, it found nothing to break the record. The next big advance may come with the Space Telescope, which is scheduled to go into earth orbit in 1985. It may tell us just how steep the cutoff is near the limit.

Our work at Cerro Tololo indicates that the density of quasars in space continues to be large up to red shifts of 3.2. The limit at 3.5 implies an abrupt change in the properties of the universe. One of the simplest explanations is that quasars suddenly formed about 15 billion years ago, which would certainly have been a remarkable occurrence in the evolution of the universe. Alternatively an absorbing screen of dust or some other kind of material may be present at a red shift of 3.5 and block our view of the more distant quasars. For the universe to be transparent on one side of the limit and opaque on the other would be equally remarkable. In either case further work on the subject is likely to bring important results in the next few years.

So far I have been discussing only the quasars' radial distribution, which historically is the way their study developed. Now that the Cerro Tololo quasar



MOST DISTANT OBJECT YET DISCOVERED, on the basis of its red shift of 3.53, is quasar OQ 172, which appears to be an undistinguished star in this photograph made with the 1.2-meter Schmidt telescope on Palomar Mountain. OQ 172 was discovered in 1973 at Lick Observatory by E. Joseph Wampler, Lloyd B. Robinson, Jack Baldwin and E. Margaret Burbidge.

samples have been completely investigated, the data base at large red shifts is substantial enough for a look at the quasars' three-dimensional distribution. By combining the red shifts with the quasars' positions in the sky one can construct a three-dimensional picture of the quasars in space. This is the only available information on the structure of the universe 13 to 15 billion years ago. The three-dimensional distribution can also yield valuable clues to the nature of the quasars themselves.

If one looks at the distribution of galaxies in photographs showing large areas of the sky, it is apparent that galaxies are anything but uniformly dispersed. They come in pairs, in small groups, in larger groups and in great clusters. Often there are blank regions with few galaxies or none. The study of galactic distribution is crucial to theories of how galaxies form. The current evidence indicates that in the very early stages of the universe matter was evenly distributed as a gas. How, then, did the condensations arise from which stars, galaxies and clusters of galaxies later developed in the course of the expansion of the universe?

Once a condensation has appeared with enough self-gravity to resist the expansion of the universe, one can imagine various ways it might collapse under its own weight, so to speak, to form a star or even an entire galaxy of stars. Most calculations show that once a body of gas begins to collapse the infall of matter to the center is swift and leads to a rapid buildup of density. If the infalling gas is of galactic magnitude, it is easy to imagine that a runaway collapse of matter might create a quasar in the center of the galaxy as part of the galaxy's evolution. This is one possible connection between quasars and the formation of galaxies. Another possibility is that quasars in their own right are powerful enough to influence formation processes on a scale considerably larger than that of a single galaxy. In that case the distribution of quasars might be different from the distribution of galaxies, which we can observe only at a time much later in the expansion of the universe.

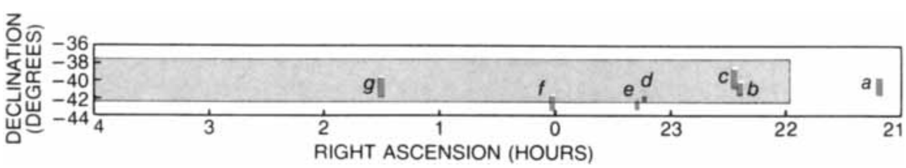
To gain a first impression of how the quasars are distributed let us look at the Cerro Tololo data in different pictorial formats. The area of the sky covered in the Cerro Tololo surveys covers a band lying between 37.5 and 42.5 degrees south celestial latitude that extends a fourth of the way around the celestial sphere [*see bottom illustration on this page*]. The band was recorded in a sequence of about 15 slightly overlapping plates made with the Curtis Schmidt telescope, each plate covering a field five degrees on a side. (The two stars at the end of the Big Dipper that point toward the North Star are separated by about the same angular distance.) The large four-meter reflector made deep exposures in seven small regions, each region about one degree square, within the broad band photographed with the Curtis Schmidt. In effect the four-meter telescope sampled long, thin tubes in space.

The Curtis Schmidt survey revealed 88 quasars with red shifts of 1.8 or more. In a much smaller area of sky the four-meter survey identified 53 quasars with comparable red shifts, including six also found in the Schmidt survey. (A handful of the brightest quasars identified in both searches had been discovered earlier by other astronomers.) The Curtis Schmidt survey disclosed one large-red-shift quasar for approximately every four square degrees of sky; the four-meter survey revealed roughly 40 times as many quasars per unit area searched, or about 10 large-red-shift quasars per square degree.

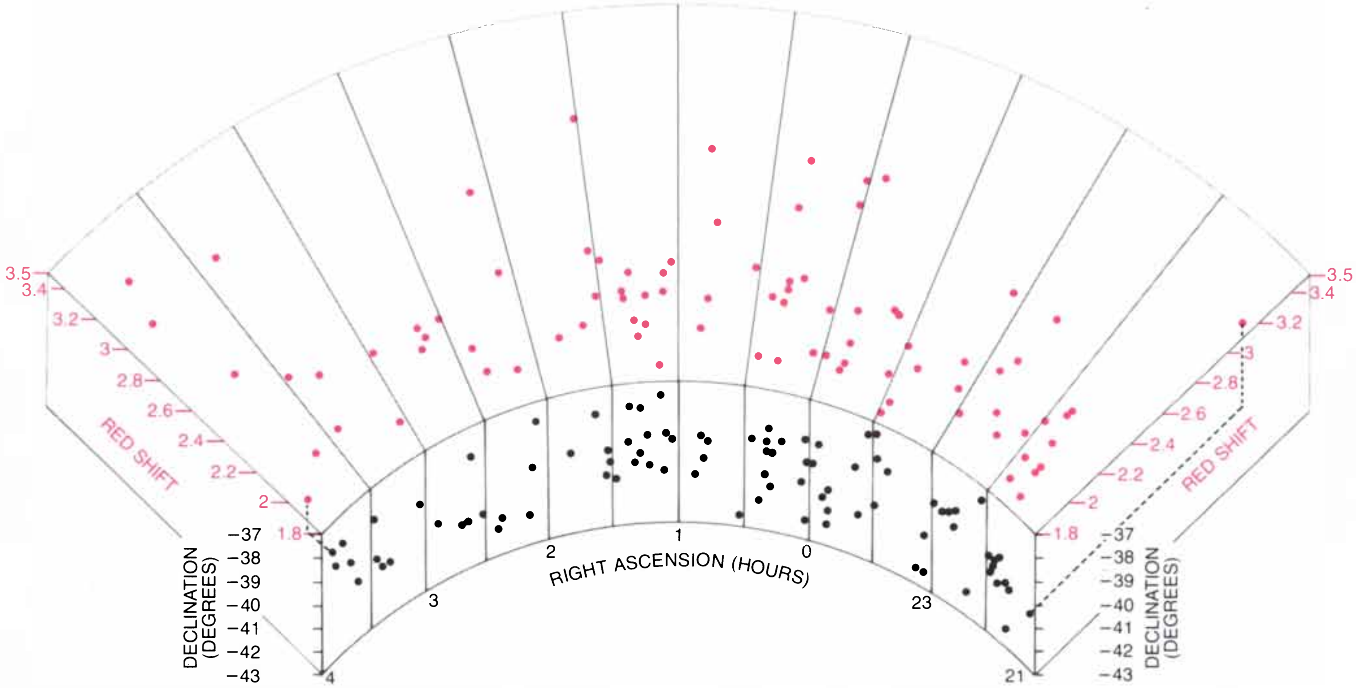
The reason for the higher discovery rate with the four-meter system, of course, is that it could detect considerably fainter objects. The four-meter survey probed to a limiting magnitude of 21, thereby revealing objects about 1.5 magnitudes, or four times, fainter than those that could be recorded by the small Schmidt telescope. Although the quasars found in the four-meter survey are generally fainter than those found in the Schmidt survey, the fainter objects do not in general exhibit larger red shifts. Most of the 135 quasars in the combined surveys exhibit red shifts near 2. This was expected both because the search method favors the discovery of such quasars and because quasars of that red shift are evidently the most plentiful. The 135 quasars examined in these Cerro Tololo surveys are about a fourth of all the quasars discovered so far with a red shift of more than 1.8.

To get a feeling for the actual distribution of quasars in space I have plotted the 88 Schmidt-survey quasars on a three-dimensional pie-shaped solid [*see top illustration on opposite page*]. The front face of the solid is an arc that locates the 88 quasars according to declination, or latitude, and right ascension (the astronomer's equivalent of longitude, which divides the celestial sphere into 24 slices, each slice 15 degrees wide). The red shift of each quasar is plotted radially on the top of the solid on a scale that runs from a red shift of 1.8 at the near edge to one of 3.5 at the far edge.

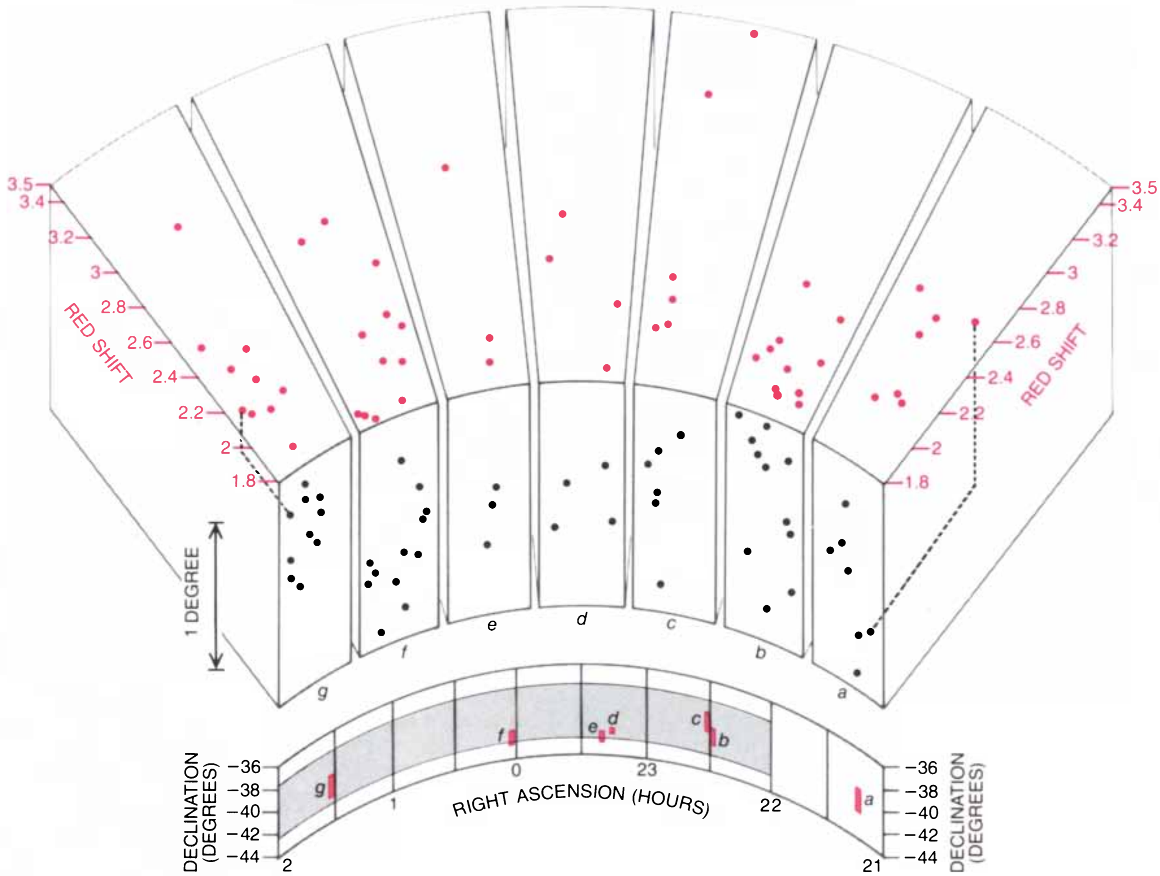
Such a plot discloses that in a region



QUASAR SEARCHES AT CERRO TOLOLO were conducted away from the plane of the Milky Way to avoid the obscuring interstellar dust that would dim the light from distant galaxies and quasars. The gray band shows the region of 340 square degrees searched by Smith working with the 60-centimeter Curtis Schmidt telescope equipped with a prism across the aperture. The seven small zones in color, covering a total of 5.1 square degrees, were sampled by Smith and Arthur A. Hoag with the four-meter telescope. The four-meter telescope was able to record objects some four times fainter than any detected by the Curtis Schmidt telescope.



CURTIS SCHMIDT SURVEY located 88 quasars with a red shift larger than 1.8. Since red shifts are correlated with distance, the 88 objects can be plotted as if they were distributed in a three-dimensional volume of space. The black dots on the front face of the diagram show the location of the quasars in the sky. The colored dots on the top surface show the red shift corresponding to each quasar. The plotting scheme is indicated by the broken lines connecting the quasar farthest to the right with its red shift at 3.16 and the quasar farthest to the left with its red shift at 1.96. Other connecting lines have been omitted for the purpose of clarity. The vertical dimension of the surveyed region has been expanded by a factor of two with respect to the horizontal dimension to enhance the separation between quasars.



FOUR-METER-TELESCOPE SURVEY located 53 quasars with a red shift larger than 1.8, including six quasars that also turned up in the Curtis Schmidt survey. Here the seven small zones searched with the large telescope are arbitrarily expanded into wedges of uniform volume. The black dots on the front face of the wedges, however, preserve the proper location of each quasar within its own sampling area. Quite a few pairs and small groups of quasars with similar red shifts are evident. In Zone *f*, for example, four quasars at a distance of about 12 billion light-years are grouped near red shift 1.84. All fall within a volume no more than 200 million light-years across.

Alas

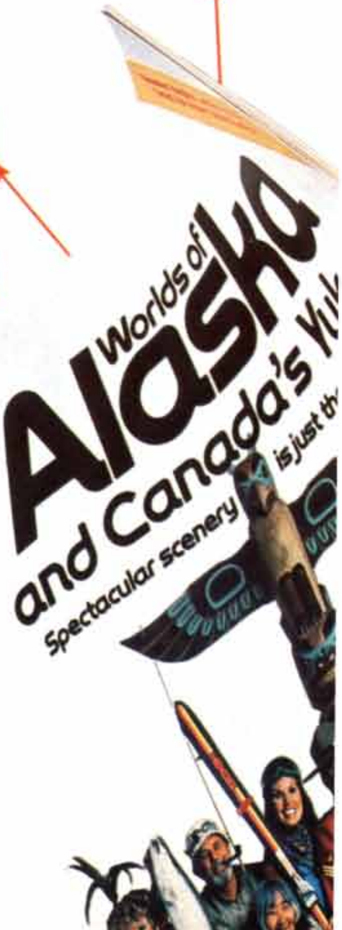
Elegant Hotels. Skylighted atrium lobbies, rooftop dining, nightly entertainment, and luxuriously appointed guest rooms are typical amenities offered by many hotels.

Luxury Cruising. An astonishing contrast of luxury and the primal forces of nature—in shirtsleeve weather.

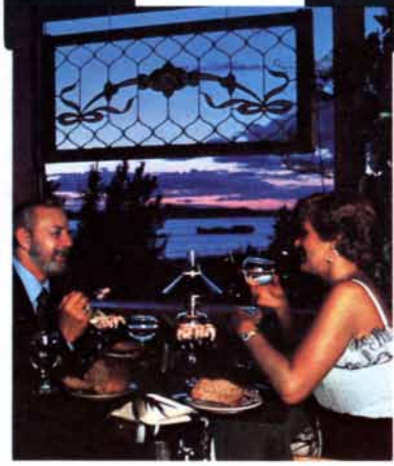
Anchorage. A major stopover for travelers en route to Europe or the Orient, Anchorage offers residents and visitors alike the excitement and conveniences of a modern city surrounded by the beauty and adventure of Alaska's wilderness.

Red Fox Cubs. Katmai National Park and Preserve, a medium-sized park by Alaska standards, is twice the size of Delaware. Red fox flourish here, along with populations of moose, wolf, lynx, bald eagles, whistling swans, sea lions, and sea otters.

Arctic Alaska. Eskimo dances, the blanket toss, demonstrations of skin sewing, ivory carving, and other traditional activities of a six-thousand-year-old culture welcome visitors to far north Alaska.



ska!

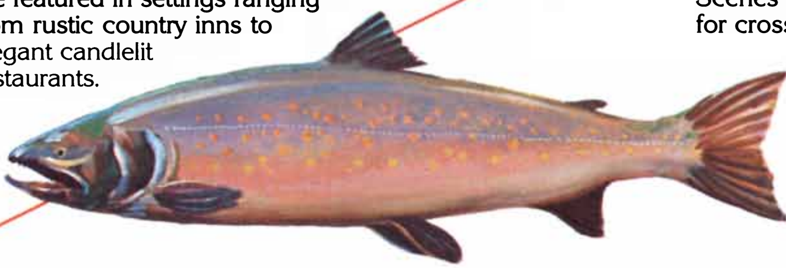


Totem Poles. In Sitka National Historical Park, scenic walking trails provide easy access to traditional totem poles set in a lofty forest of spruce and hemlock.

Dining. Fresh Alaska seafood, international cuisine, and a surprising variety of ethnic specialties are featured in settings ranging from rustic country inns to elegant candlelit restaurants.

Alaskan Artistry. Visitors discover value in a wide range of gifts produced by local artisans, including original art prints, scrimshaw, beadwork, furs, and carvings.

Matchless Winter Sports. Scenes like this are the backdrop for cross-country skiing, dog sledding, and winter hot-air ballooning over snowy peaks.



Arctic Char. One of many magnificent sportfishing species found in Alaskan waters. With more than a million freshwater lakes and thousands of streams and rivers, Alaska offers sportfishing experiences to suit the most discriminating angler.

Send for your FREE 132-page book that shows how to enjoy an Alaska vacation!

Alaska is the vacation experience of a lifetime and this book will start you on your way. It's packed with useful information on how to plan your Alaska vacation. How to get there, what to see, what to do, where to stay, what to look for when you are shopping, and how to make it all fit your budget. You'll discover Alaska is warmer, closer, more comfortable, and more affordable than you ever imagined. **Worlds of Alaska and Canada's Yukon.** It's free. Send for it today.

SEND THIS REPLY CARD FOR YOUR FREE BOOK TODAY!

If the pre-paid reply card has been removed, write:
Alaska State Division of Tourism, Pouch E–44, Juneau, Alaska 99811.

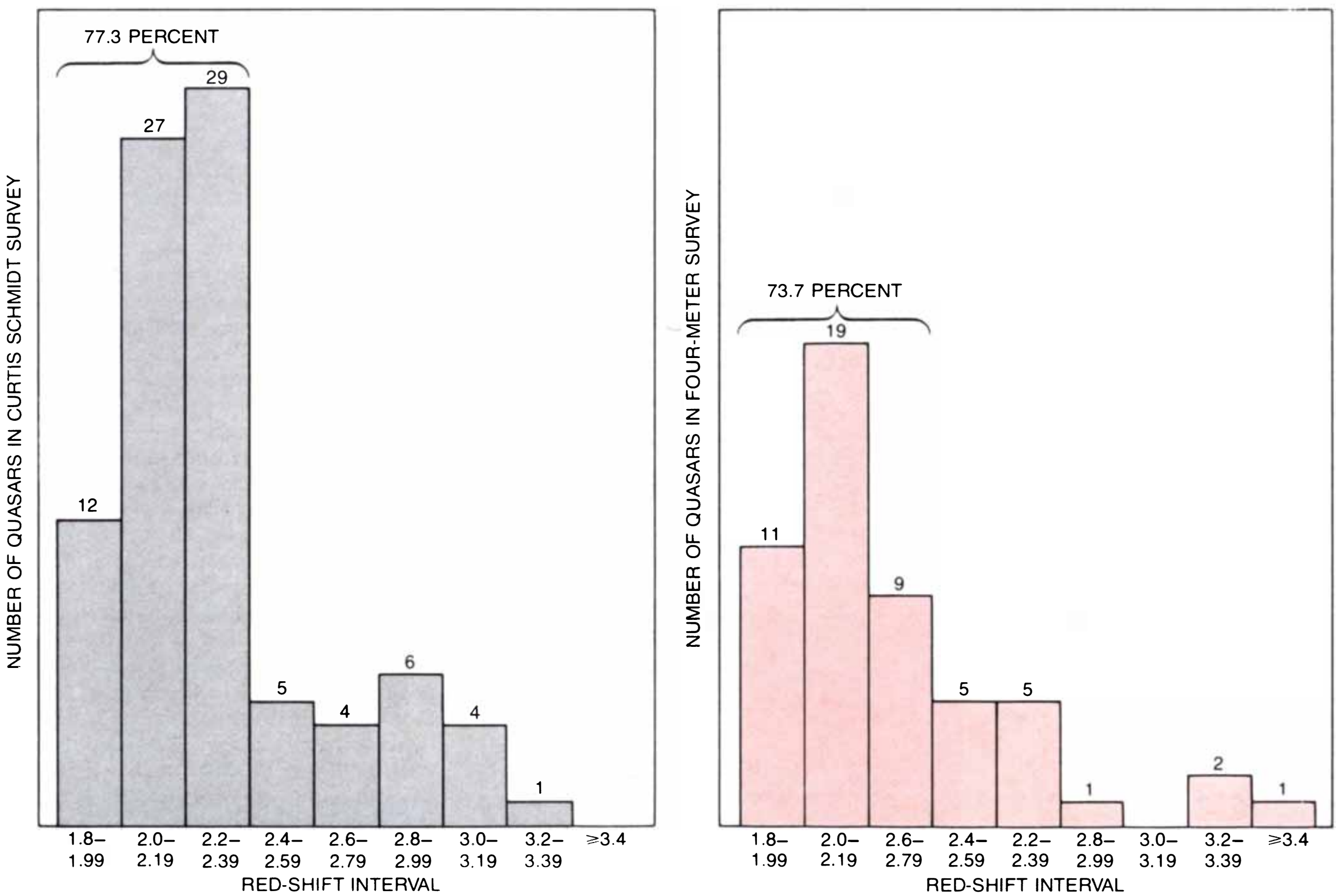
between two and four hours right ascension quasars seem to be sparser than they are elsewhere. If the scarcity reflects an actual nonuniformity of quasars in that part of space, it would be startling. Therefore my colleagues and I prefer to think provisionally that the apparent scarcity is due to a selection effect in the survey process. If one looks at the small-scale distribution of quasars, one can see instances where pairs and small groups of quasars lie fairly close together. The most prominent clumping shows up clearly when the quasars in the four-meter survey are plotted in a similar three-dimensional way [*see bottom illustration on page 133*]. A group of four quasars with red shifts between 1.83 and 1.86 are clustered in a region no more than 200 million light-years across near zero hours right ascension.

Such groups are fascinating because they have about the same size as the superclusters of galaxies spotted here and there in nearby space. Superclusters are regions some 300 million light-years in diameter that harbor several clusters of galaxies. They are the largest known structures in the universe. So far superclusters have not been observed at red shifts of as much as 2. Could the group of quasars with red shifts around 1.84 mask the location of a remote supercluster? Perhaps the Space Telescope will be able to tell us.

Important as visual examinations of the data are, the question of how quasars are distributed ultimately calls for statistical analysis. Is there an underlying pattern to the distribution? Are the observed groupings merely chance fluctuations? After all, if rice grains are scattered on the floor, some will fall closer together than the average distance between the grains. In recent years powerful statistical techniques have been brought to bear on galaxy surveys to investigate just such questions. With suitable modification for the large red shifts and the selection effects of the survey method such techniques can also be applied to the quasars.

The prime objective is to discover whether the quasars show any departure from a uniform, random distribution. If they do show such a departure, one would like to know whether quasars cluster and what form the clustering takes. Alternatively, one may find that quasars exhibit anticlustering and lie farther apart than one would expect. That could happen if the presence of a quasar at a given place were to inhibit the formation of other quasars nearby. The mathematical formulation of such possibilities and the tests to distinguish them need not concern us here, but the main possibilities lend themselves to graphical presentation in two dimensions. Strong clustering and strong anticlustering are recognizable at a glance. When the clustering or anticlustering is weak, however, the distribution is hard to distinguish from one that is actually uniform and random [*see illustration on next page*].

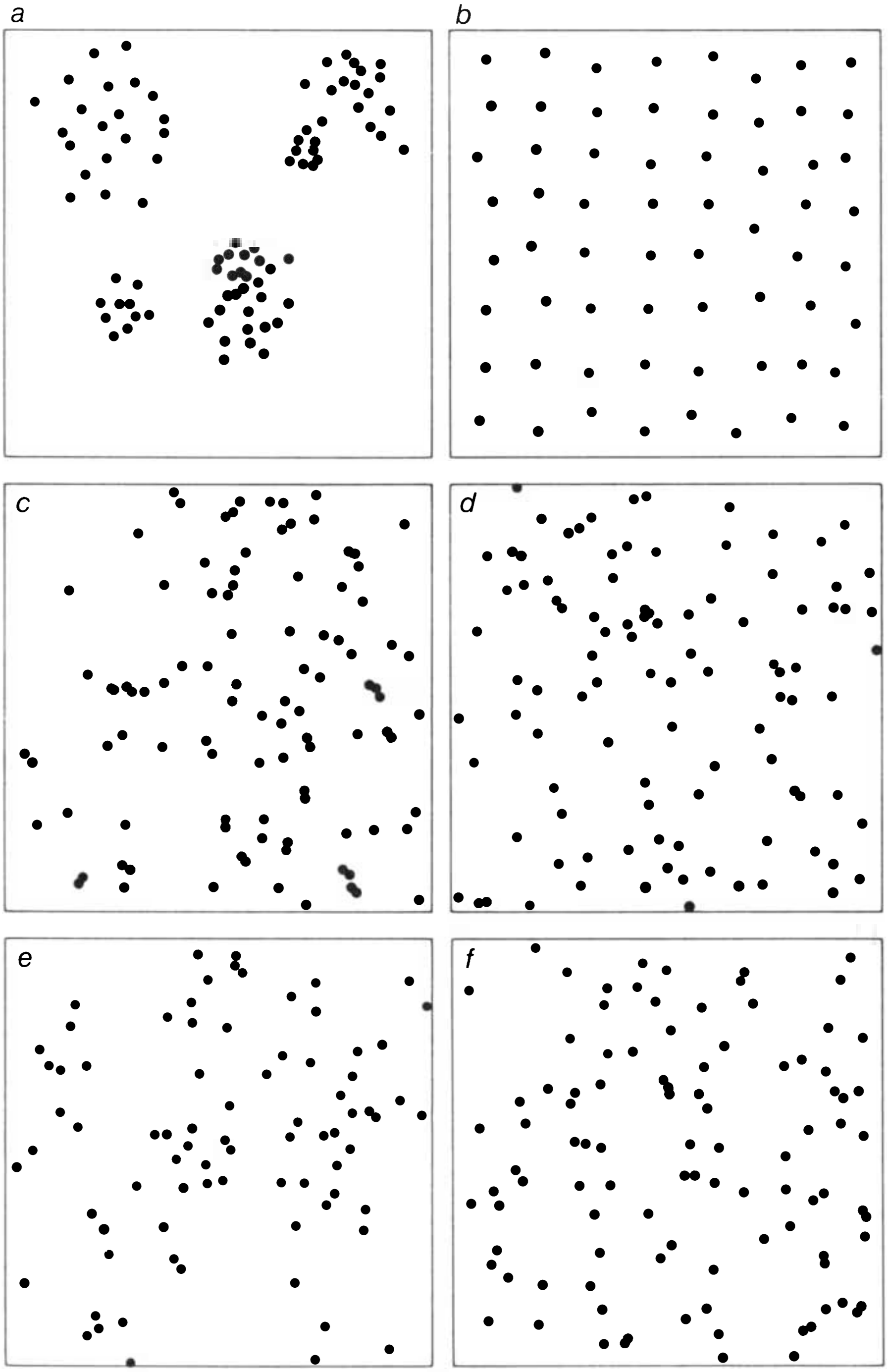
I can summarize the results obtained so far from the Cerro Tololo surveys by saying that the statistical analysis is consistent with the quasars' being distributed uniformly at random. The pairs and groups picked out by eye are evidently nothing more than chance fluctuations. For the 30-odd quasars with red shifts between 1.8 and 2.2 in the four-meter sample the mean nearest-neighbor dis-



DISTRIBUTION OF QUASARS BY RED SHIFT shows that about three-fourths of the quasars in the two Cerro Tololo searches have red shifts between 1.8 and 2.4. The 88 quasars in the Curtis Schmidt survey have the distribution at the left; the 53 quasars in the four-meter survey have the distribution at the right. The Curtis Schmidt survey disclosed one quasar for every four square degrees of sky. The four-meter-telescope survey, capable of detecting fainter objects, uncovered on the average slightly more than 10 quasars per square degree searched. Although four-meter quasars were generally less luminous than the Curtis Schmidt quasars, they were no more distant.

tance is about 400 million light-years, which is well within the expected range for a random distribution. For the quasars with larger red shifts, where mean separations are measured in billions of light-years, a similar conclusion can be drawn. These findings are evidence in favor of an assumption that is generally held but difficult to confirm, namely that on a large scale the universe is homogeneous. This assumption is a crucial requirement for current models of the universe.

Reassuring as such results are, they mark only the first step in a continuing inquiry that is likely to progress rapidly as more quasars are found. If the observed clustering of galaxies is extrapolated to the epoch and scale covered in the quasar survey, one can show that galaxy clustering should not be detectable in the existing samples. It will be interesting to see if the Space Telescope can improve the sensitivity of the distribution tests to the point where galaxy clustering can be detected. The expectation is that galaxy clusters exist at a red shift of 2, but that they stand out less prominently above the background of denser matter in that epoch than do clusters with smaller red shifts observed at a later epoch in a much expanded, and hence less dense, universe. At the same time the apparent pairs and groups of quasars in the four-meter survey cannot be ignored. After all, one theory of galaxy formation postulates that the galaxies originated with random density fluctuations in the early universe. The quasar groupings may represent similar regions of enhanced density.



PATTERN OF QUASAR DISTRIBUTION is being sought because it might lead to a fuller understanding of galaxy formation in the early universe. The six patterns appearing in this illustration are two-dimensional analogues of various distribution possibilities. If quasars are strongly clustered (*a*) or strongly anticlustered (*b*), the fact should be apparent. If quasars exhibit only weak clustering (*c*) or weak anticlustering (*d*), extensive statistical tests may be needed to demonstrate that the distributions are not uniform and random. Two computer simulations of uniform, random distributions (*e, f*) show that the eye is not a reliable judge of randomness.

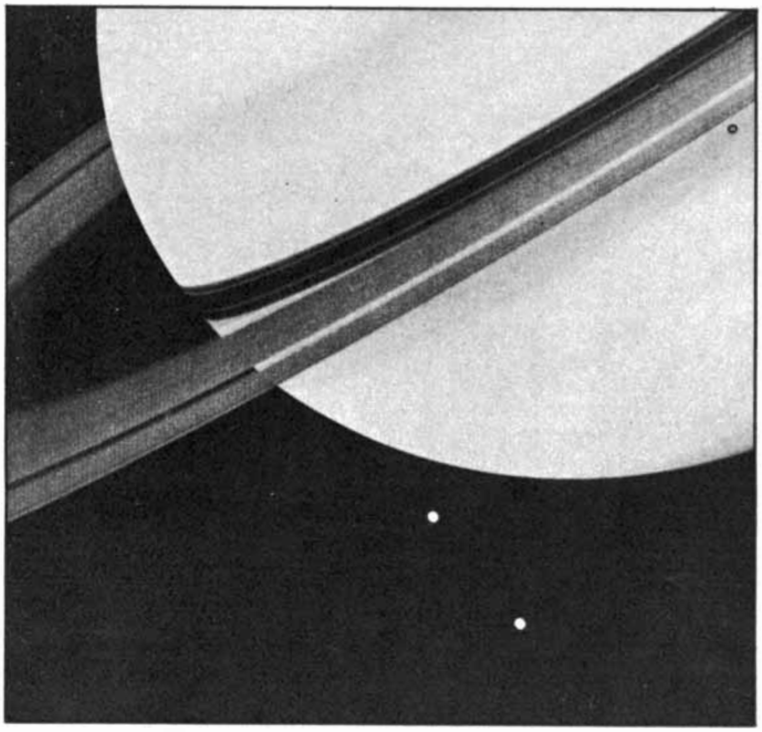
As a result of our Cerro Tololo investigations one can now try to describe what the universe looked like 12 to 15 billion years ago. Suppose our own galaxy, the Milky Way, were within the group of four quasars that to our present-day instruments exhibit a red shift of about 1.85. Assuming that human life could have evolved that early in the age of the universe, what would we see at night? First of all, the Milky Way itself would be much brighter than it is today because it would have a more prominent population of hot, young stars. The four quasars in the group would be clearly visible to the unaided eye as bright stars; in fact, quasars outside the group, and as far away as 300 million light-years, could also be seen without a telescope. Obviously quasars would have been discovered and recognized as strange phenomena early in the history of astronomy. Whether or not they are the luminous core of galaxies would have been ascertained about as soon as there were telescopes.

It is well to remember that, plausible as our current picture of quasars may be, there is some chance that it is entirely wrong and a good chance that it is wrong in important particulars. Some astronomers doubt that quasars are as distant as their red shifts indicate. Others question the reality of the high density of quasars at large red shifts and propose that quasars were either brighter then or that their intensity has been enhanced as a result of their radiation's passing close to intervening galaxies that act as gravitational lenses. Rarely, if ever, is a large body of data collected and suddenly explained at a stroke by an inspired theory. So it is likely to be with quasars. The subject would be the poorer without the current hypotheses, presumptuous as they may be. In any case it is hard to imagine the prevailing concept of quasars being overturned without their becoming even more amazing than they already are.

YOUR FIRST TRIP INTO SPACE IS FREE...

ABOARD ASTRONOMY MAGAZINE

From the delicate precision of the celestial sphere, to the hurly-burly of colliding galaxies, ASTRONOMY Magazine is your vehicle for exploring the universe.



Stunning graphics and breathtaking space photography add a fresh visual perspective to ASTRONOMY's articles. You'll feel you're there — in space — when you see the vivid photographs of the rings and moons of Saturn transmitted 930,000,000 miles back to Earth by the technology of Voyager, and examine the core of our galaxy through the "eye" of a giant infrared telescope.

ASTRONOMY's contributors are experts in their fields of research — the hands and eyes that aim the telescopes — but they don't clutter their writing with technical jargon. They give you the facts...and charge them with the rich sense of fascination and wonder that is the hallmark of ASTRONOMY.

ASTRONOMY takes a hands-on approach to exploring the heavens. You'll learn to build and use a telescope...pinpoint the stars and constellations...capture the beauty of the night sky on film...and chart the movement of the celestial sphere.

ASTRONOMY also delivers up-to-date information on the latest theories and controversies in the space sciences, and provides valuable tips you'll want to read before investing in related books and equipment.

And now you can launch the first leg of a year-long odyssey into space at no charge... with a free introductory issue of ASTRONOMY.

ASTRONOMY brings you face to face with the mysteries of the cosmos and explains the ongoing efforts to unravel them. You'll embark on speculative journeys through black holes...travel in time to the explosive birth of our universe...explore the formation of the solar system...and delve into the riddles of quasars, pulsars, supernovae and the enigmas of deep space.

How to Send for Your Free Trial Copy

Send no money now — just fill in the attached coupon and return it to us. We'll send your no-risk trial issue of ASTRONOMY, immediately. To subscribe, just send payment upon receipt of our bill.

If you decide you don't want to subscribe, just mark the invoice "cancel" and return it but keep the free trial issue.

It's as easy as that to take your first trip into space!

FREE

As a new subscriber you will receive absolutely FREE this huge NASA Space Shuttle poster. It's a 30x42" full-color cutaway drawing detailing our first true spaceship. We've sold hundreds at $3.25, but yours is FREE upon payment!

ASTRONOMY, 625 E. St. Paul Ave., P.O. Box 92788, Milwaukee, WI 53202

FREE TRIAL ISSUE ☐ Send my first FREE issue of ASTRONOMY If I like it, I'll return my payment when you bill me and receive 11 more issues(one year) or 23 more issues (two years). If not, I'll return the invoice marked "cancel," keep the first issue and owe you nothing. My free poster will be sent upon payment.

☐ 1 Year (12 issues) $12.95 ☐ 2 Years (24 issues) $21.95

NAME____________________

ADDRESS ____________________

CITY ____________________

STATE __________ ZIP __________

All Foreign add $5 per year 5010

Color Vision in Fishes

The visual environments of fishes are blue, green or near infrared. The retinal pigments acquired by diverse fish species in adapting to these environments are a valuable clue to the evolution of the eye

by Joseph S. Levine and Edward F. MacNichol, Jr.

The photosensitive cells of the retina are the outer extension of an animal's visual system. The cells catch the photons of visible light and in effect count them, thereby triggering a complex series of chemical and neural events that can result in visual sensation. In many animals the data-collecting function of the cells has developed to take advantage of discriminations possible on the basis of color as well as brightness, because in many environments color vision is an efficient means of gaining accurate information essential to survival. In deepwater environments, however, the number of photons that bathe an organism is much reduced and the color of the ambient light is confined to a narrow spectral band. The system of photoreceptor cells in the retina of a deep-water organism has correspondingly developed to maximize sensitivity to the available light rather than to provide color discrimination. Between the darkness at depth and the full sunlight at the surface there is a continuous range of luminous environments that vary in both color and intensity. In such a natural laboratory one encounters a wide variety of visual systems.

Several years ago we began an intensive study of the photoreceptors of teleost fishes (bony fishes, as opposed to cartilaginous ones such as sharks) in the hope of understanding how complex visual systems might have evolved from their precursors. The evolutionary reconstruction of the visual system cannot rely on fossil evidence because the soft tissues of the eye and the nerve pathways are not preserved. We chose therefore to investigate representatives from a variety of teleost families so that we could survey some of the relations among visual systems, visually dependent behavior and the environment. If the evolutionary relations among living teleost species are known on independent grounds, an examination of their visual systems and their visual ecology could lead to the reconstruction of the likely stages in the evolution of vision.

It might be supposed such a reconstruction would be relatively straightforward. One might assume, for example, that fishes known through fossil evidence to have undergone little recent evolutionary change would retain relatively primitive visual systems. Accordingly one might also assume that species evolving in relatively recent times would incorporate visual systems more advanced than the ones found in the older species. Both assumptions, however, take no account of the possibility that eyes may have evolved independently of other organs, because eyes themselves may confer selective advantages for such vital activities as finding food, avoiding predators and selecting a mate. The selective advantages of one visual system over another, in a particular environment for a particular behavioral repertory, may bring about changes in the eyes of a fish that is otherwise unaltered by evolutionary forces. The teleosts illustrate such evolutionary possibilities with particular clarity.

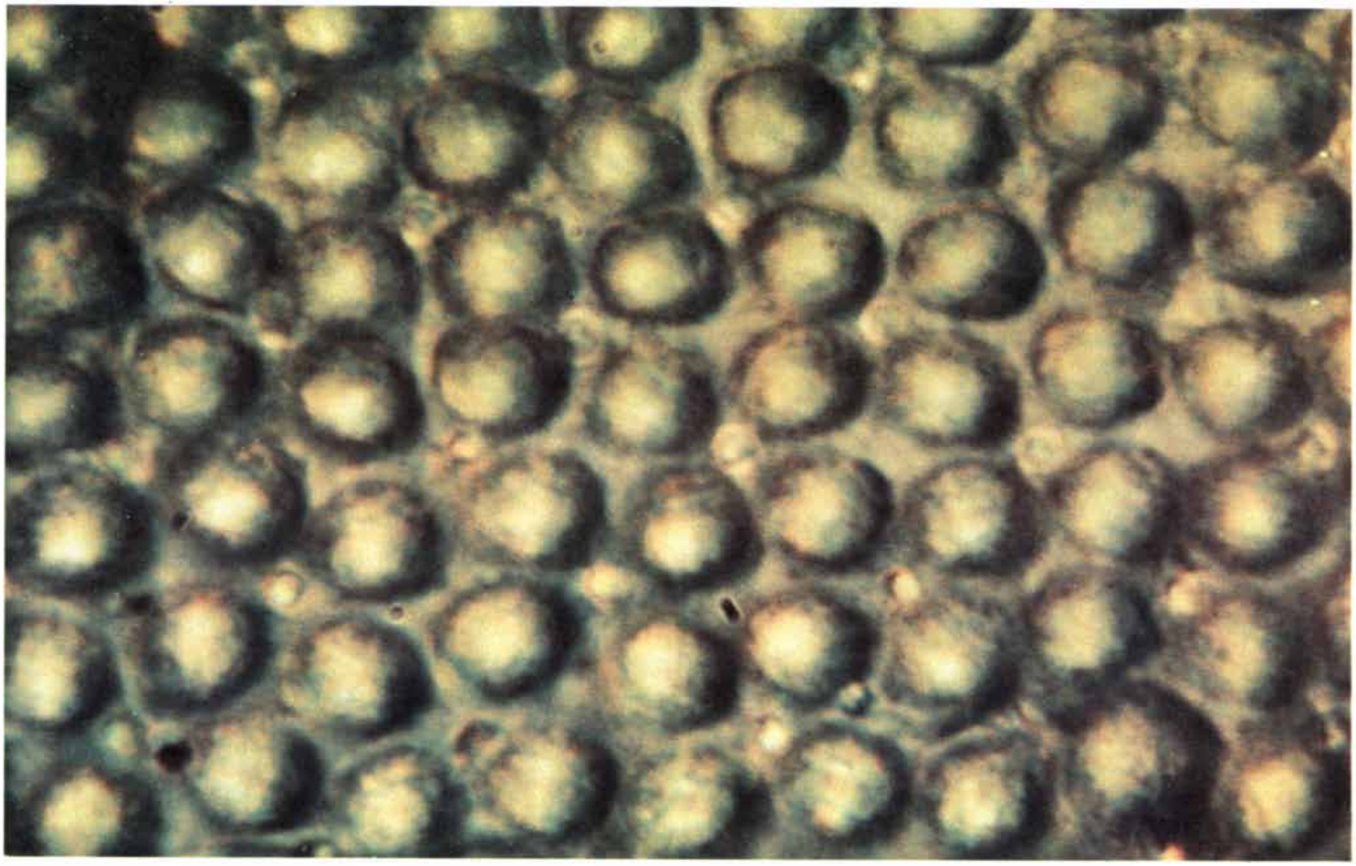
Explosive evolutionary radiation has emplaced between 17,000 and 30,000 species of teleosts in a broad range of aquatic habitats. Even single families of teleosts, with a wide variety of visually dependent types of behavior, occupy habitats that span nearly the full range of light zones in the sea and in freshwater lakes. Their eyes can differ so dramatically in size, shape, retinal organization and color sensitivity that one would never suspect, from examining the eyes alone, that the species are closely related. On the other hand, unrelated species inhabiting the same light zone can show remarkable similarities in their visual systems. For this reason primitive visual systems that may be preserved in living teleosts cannot be unequivocally distinguished from degenerate or specialized systems.

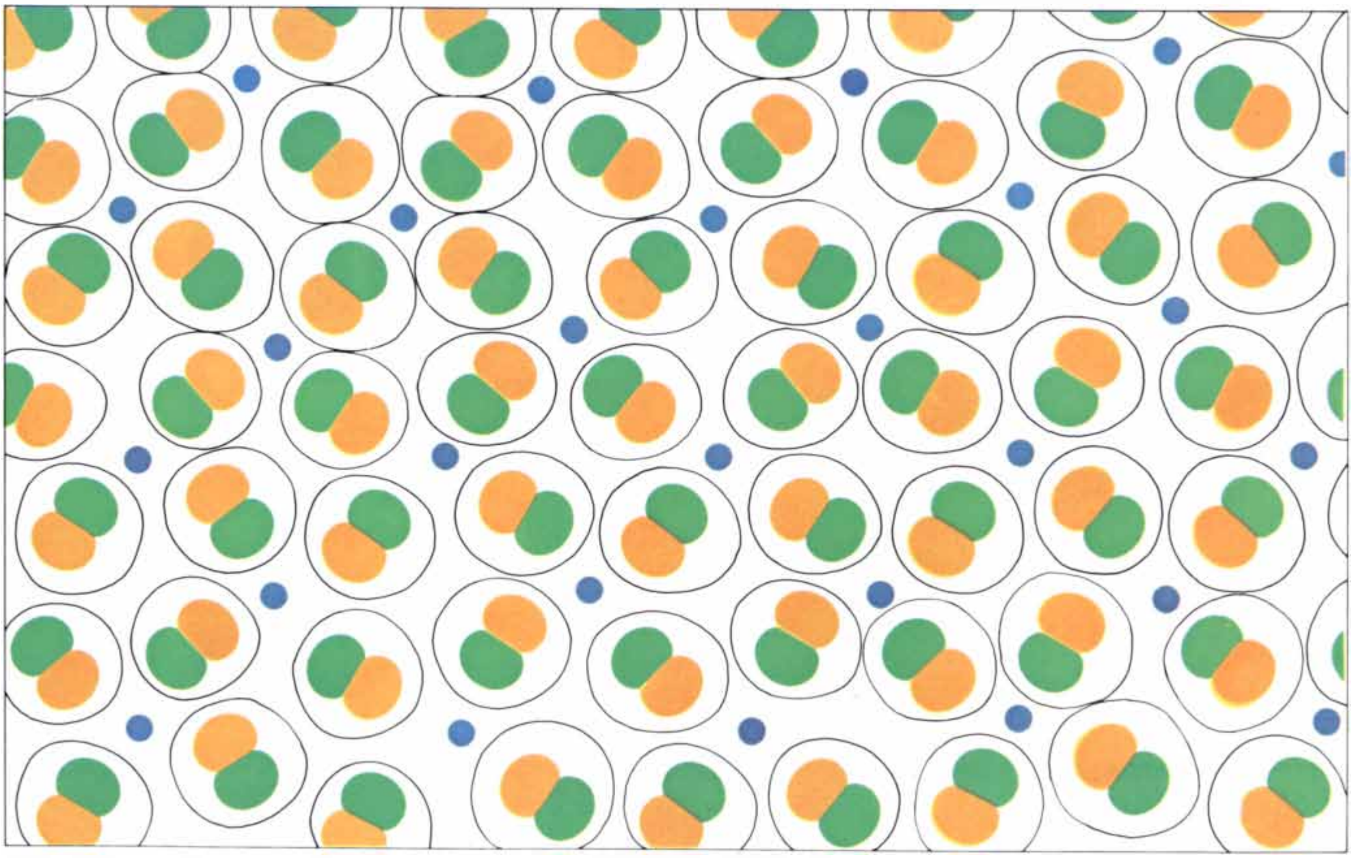
In spite of this difficulty, comparative ecological study can answer certain important evolutionary questions. What combination of selective advantages and operating constraints favors the possession of visual systems capable of discrimination over a broad range of colors and light intensities? Under what circumstances is the possession of such a complex system an encumbrance rather than an advantage? Although relatively little is known about the constraints on vision that are specifically related to feeding, fleeing or mating behavior, much has been learned about the lighting conditions that prevail underwater.

The solar radiation reaching the surface of the water is made up of photons of every visible wavelength, together with photons of the infrared and ultraviolet regions of the electromagnetic spectrum. The energies of visible-light photons correspond to colors or wavelengths, from violet with a wavelength of 400 nanometers to deep red with a wavelength of 700 nanometers. When light passes through the water, its intensity decreases, and the loss of intensity varies with color. Bodies of fresh water and salt water that do not contain much organic matter absorb violet and red light much more than they absorb light of intermediate wavelengths. Hence tropical oceans and clean lakes such as Crater Lake in Oregon and Lake Malawi in eastern Africa get bluer with depth as colors other than blue are absorbed. In such waters blue light can penetrate to a depth of more than 75 meters, whereas red and violet light are essentially absent below a depth of about 25 meters.

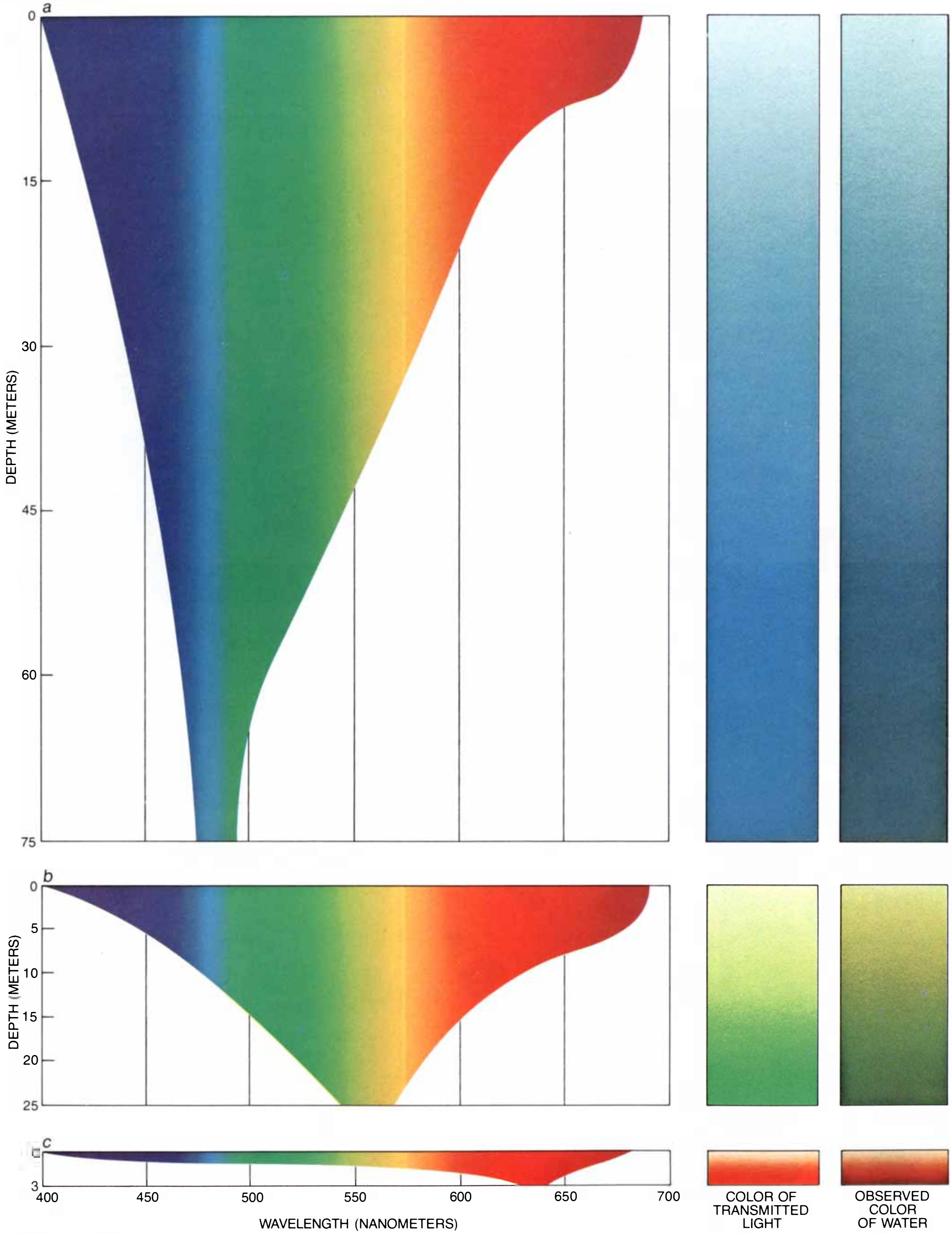
Unlike these "clear water" habitats, most Temperate Zone lakes, streams and coastal waters contain both yellow-green phytoplankton and dissolved organic matter from decaying plants and animals. In such waters light of all wavelengths is absorbed more strongly than it is in clear water, but the relative color intensities at each depth are quite different from what they are in clear water. The short-wavelength violet and blue light is the most strongly absorbed, and green or yellow-green light at wavelengths between 540 and 560 nanometers becomes increasingly dominant with depth. To the human eye the water looks green.

Marshes, swamps and "black water" rivers contain the same short-wave-





PHOTORECEPTOR CELLS in the retina of the black tetra fish (*Gymnocorymbus ternetzii*) form a regular mosaic pattern. In the photomicrograph at the top the cone cells normally essential for color vision are shown as they appear from the back of the fish's retina, magnified 1,000 diameters. (Most of the rod cells, which are more sensitive to light than the cone cells, have been removed from the specimen for clarity.) Three kinds of cones are present in the retina of the black tetra: single cones holding a blue-sensitive visual pigment and double cones, made up of separate cells, holding both a green- and an orange-sensitive pigment. In the drawing at the bottom the color of the light that is maximally absorbed by each pigment is associated with the position of the pigments in the retina. The photomicrograph was made by the authors. The color of the cells in the photomicrograph is an artifact introduced by the technique of microscopy.

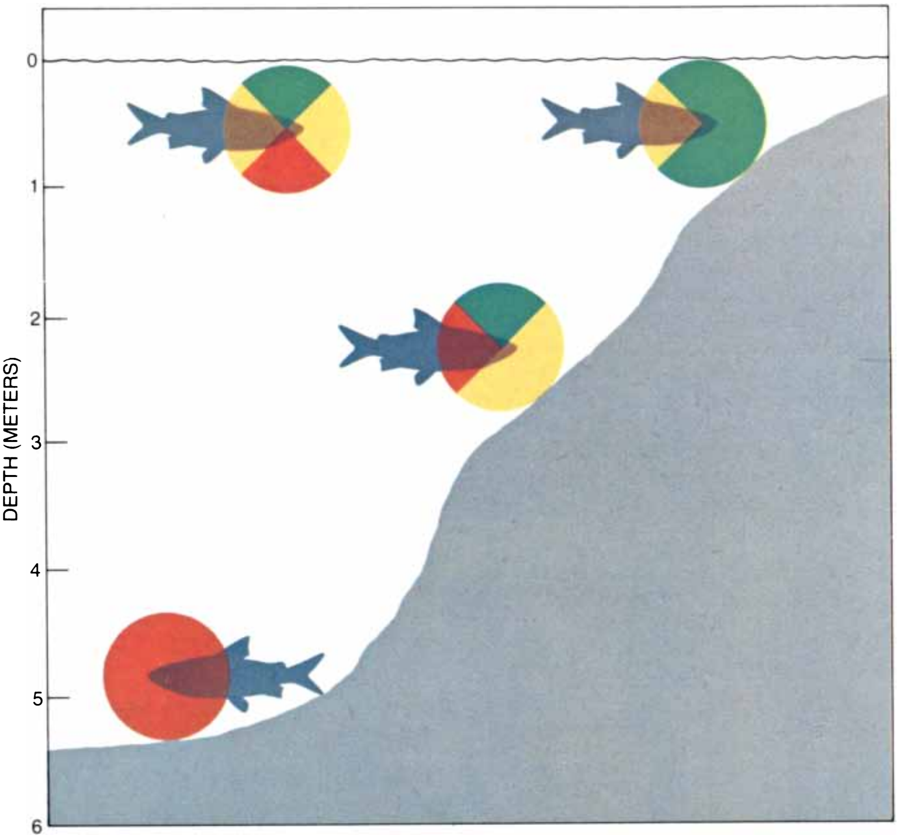


TRANSMISSION OF LIGHT by water is dependent on the color or wavelength of the light. In clear oceans and lakes (*a*) the light becomes increasingly monochromatic and blue as its path length increases. In fresh water that carries green organic matter (*b*) light at all wavelengths is absorbed more quickly than it is in clear water, but the light becomes greener with path length. In rivers, swamps and marshes that carry large amounts of the products of plant and animal decay (*c*) absorption is rapid and the spectral distribution of the light shifts to the red. Such waters are called black because the human eye is relatively insensitive to light at long wavelengths; a less anthropomorphic name would be infrared water. The depths given for the maximum penetration of light are typical, but they vary widely.

absorbing compounds as lakes and streams, along with tannins, lignins and other products of more complete plant decomposition. The combined effects of the suspended and the dissolved matter can be the absorption of all light within about three meters of the surface; the maximally transmitted wavelength is 600 nanometers or longer, well into the red. To the eye such water usually appears dark or ruddy brown.

Light passing through water is also scattered in all directions by suspended particles and the water molecules themselves. When light is scattered, it must travel farther through the water to reach a given point; moreover, because of scattering every line of sight in the water is a source of photons, and so the water itself appears to be colored in all directions. This coloration is called background space light, and the combined effects of scattering and absorption cause it to vary with the direction of the line of sight.

Background space light has long been familiar to scuba divers, but it was not until the mid-1970's that spectroradiometric measurements of it were made in tropical marine habitats. William N. McFarland of Cornell University and Frederick W. Munz of the University of Oregon quantitatively confirmed earlier observations that light along horizontal lines of sight is much more monochromatic than light along a vertical line of sight near the surface. McFarland and Munz recognized their finding as a major factor in the evolution of aquatic visual systems. Unlike terrestrial animals, which must distinguish food, predators and mates from whatever happens to be in the background at the time of sighting, fishes usually must detect objects against background space light that has a relatively constant color along any particular line of sight.



SCATTERING OF LIGHT by water molecules and particles suspended in the water causes the dominant color of the transmitted light to vary with the direction of the line of sight. Because every point on every line of sight is a source of photons the water itself appears to be differently colored in all directions. The coloration imposes a background of space light against which any object must be viewed, and it also interposes a veiling brightness between the observer and the object. The illustration shows how the coloration varies with direction in four microhabitats of a "black water" river: at the surface and on the bottom in shallow and in deep water.

In the eye of a fish, as in the eye of a mammal, the photoreceptor cells onto which the incoming light is focused are filled with the photosensitive substances called visual pigments. Visual pigments were independently discovered in 1877 by the German physiologists Franz Boll and Wilhelm Friedrich Kühne. Boll noted that when a frog's retina is removed from the eye, it is initially bright red but bleaches in light, becoming first yellow and finally colorless. Kühne found soon afterward that the pigment in the rod cells (the rod-shaped photoreceptor cells) of many animals is also red and bleaches when it is exposed to light. In a living animal the color of the pigment is restored if the animal is placed in the dark, but in cells that have been removed from the animal the bleaching of the pigment is permanent.

The visual pigments catch photons of light by absorbing them, just as water does. To appreciate the importance of the pigments, however, it is necessary to consider the wavelengths of light that are maximally absorbed rather than those that are maximally transmitted. The pigment in human rod cells, for example, is pink, hence the name rhodopsin, from the Greek words meaning "rose" and "vision." Rhodopsin appears pink because it transmits red and blue light, whereas its absorption maximum is at a wavelength of 500 nanometers, in the blue-green region of the spectrum.

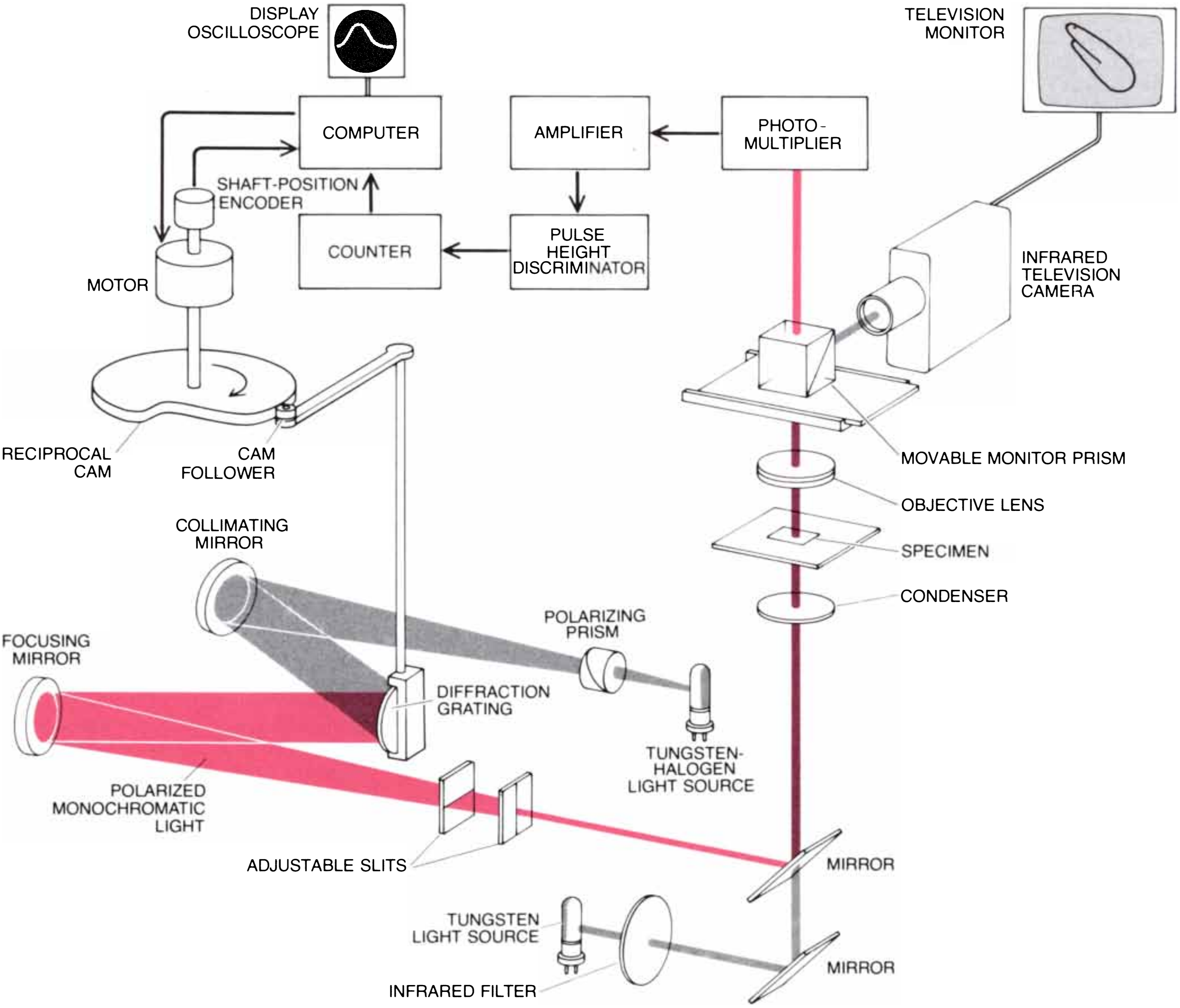
When a visual pigment absorbs photons, the molecules in the pigment undergo structural changes and trigger electrochemical responses in the photoreceptor cells that hold the pigment. The response of each cell is determined in part by the number of photons absorbed by the pigment but not by the wavelengths of the photons; once a photon is absorbed its wavelength no longer affects the output of the photoreceptor cell. The only means the visual system has of extracting information about the color of incident light is to take into account the statistical probability that one of its pigments will absorb a photon with a given wavelength.

How can one determine these absorption probabilities? Early investigators extracted pigments chemically from a whole eye, and so their investigations were limited almost entirely to the pigments in the rod cells. Rod pigments are more abundant and easier to extract than the pigments in cone cells (the cone-shaped photoreceptor cells found in many retinas). After a rod pigment was extracted its spectrum was analyzed by a spectrophotometer. In the instrument a beam of monochromatic light of selected visible wavelengths was passed through the pigment and the intensity of the transmitted light was measured. The measured intensity was compared with the intensity of a reference beam from the same source that did not pass through the pigment. It is possible to transform the data mathematically into a graph that shows the optical density of the pigment as a function of wavelength; the graph is called an absorbance spectrum. One can also calculate the probability that a visual pigment will absorb a photon of a given wavelength from the absorbance spectrum of the pigment.

The method we employ to determine the absorbance spectra of visual pigments is conceptually identical with the method described above, except that we are able to record the spectra of pigments in individual photoreceptor cells. The instrument we use is the microspectrophotometer, a spectrophotometer that incorporates a high-powered microscope so that a magnified area of the image of the specimen can be spectrally analyzed. Such instruments are available in fewer than a dozen laboratories throughout the world. Earlier versions of the microspectrophotometer in our laboratory were developed by William B. Marks of Johns Hopkins University and one of us (MacNichol) in the early 1960's and by Ferenc I. Hárosi of the National Institute of Neurological Diseases and Stroke in the early 1970's. The instrument enables us to examine not only the different pigments in different kinds of photoreceptor cells but also the distribution of the cells in the retina.

To prepare a retinal specimen for microspectrophotometry the thin layer of photoreceptor cells must be separated from the pigment epithelium, a layer of retinal tissue behind the cells that absorbs stray light in the eye. This is accomplished first by soaking the retina in a saline solution and then teasing the layers apart surgically. The layer of photoreceptor cells is mounted on a thin microscope cover glass, and a second cover glass is placed over the cells in such a way as to flatten them on the first glass. All the work must be done in the dark or under infrared illumination, since otherwise the pigments in the specimen would bleach and so be of no use for getting spectral data. Hence in order to position the slide and focus the image of the specimen for scanning by the microspectrophotometer we employ infrared illumination at a wavelength that is not appreciably absorbed by any visual pigment. We observe the specimen on a television monitor connected to an infrared-sensitive television camera, which receives an image of the specimen magnified 1,000 diameters.

Once a photoreceptor cell has been aligned with the path of the scanning beam of the microspectrophotometer the recording of the absorbance spectrum is virtually automatic. A low-intensity, monochromatic beam of light is passed through the apparatus, first without the specimen cell and then with the specimen cell in place. The intensities of the two transmitted beams for each wavelength are recorded by a small



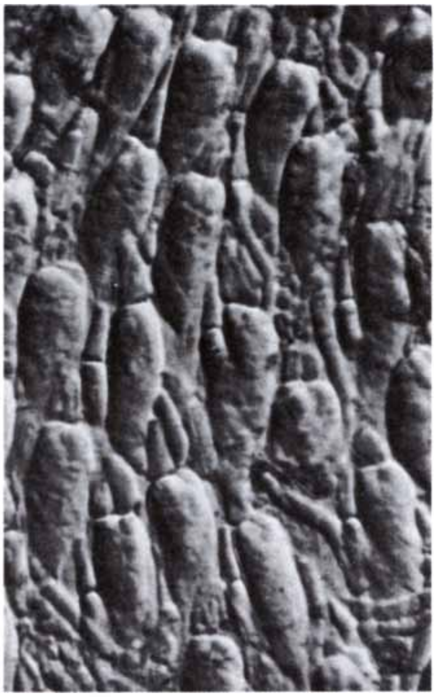
MICROSPECTROPHOTOMETER determines the absorbance spectrum for the visual pigment in a single photoreceptor cell. The apparatus measures the ratio of the intensity of a beam of monochromatic light that passes through the pigment to the intensity of an identical beam that does not pass through the pigment. In the schematic diagram each beam is polarized and then split into its spectral components by a diffraction grating. The wavelength of the monochromatic beam is calibrated to the position of a camshaft that controls the grating. Hence the precise wavelength of the input beam and the intensity of the output can be registered simultaneously in a computer. The pigments will bleach and lose their ability to absorb light if they are exposed to bright light. Therefore the beams passing through the specimens are held to low intensity, and the individual photons in the beam are converted into electric pulses by a photomultiplier. The pulses are then amplified, sorted from spurious random signals and tallied in a counter, whose output is periodically read into a computer. To position and focus the specimens infrared illumination is used in combination with an infrared television camera and monitor.

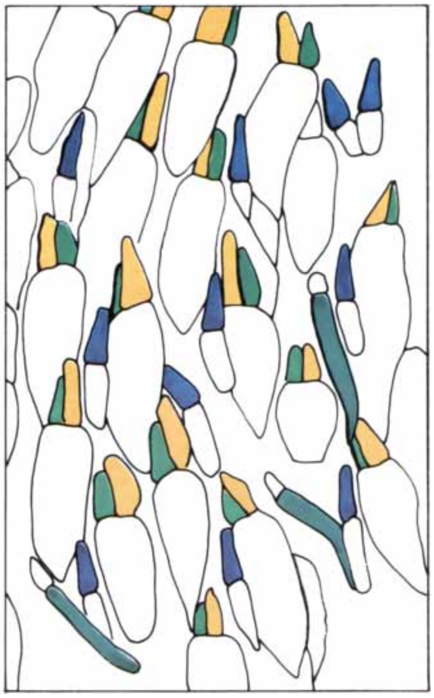
computer and compared. By working with low-intensity beams and by scanning the spectrum rapidly and repetitively we avoid having to compensate for the bleaching of the pigment every time the light passes through the specimen cell.

Almost all visual pigments absorb light of all visible wavelengths to some degree. Each pigment, however, has a well-defined wavelength for which the absorption probability is maximized, and the probabilities fall off rather sharply for both longer and shorter wavelengths. A deepwater fish with only one visual pigment would have the highest possible visual sensitivity if the maximum absorption probability of the pigment were matched to the color of the background space light. One of the earliest successes of visual ecology was establishing that in nonmigratory, deepwater marine fishes such matches are common. The light available to fishes more than 100 meters down is both dim and blue, at wavelengths between 470 and 490 nanometers. The maximum absorption wavelengths of the rod pigments closely fit the spectral distribution of the ambient light.

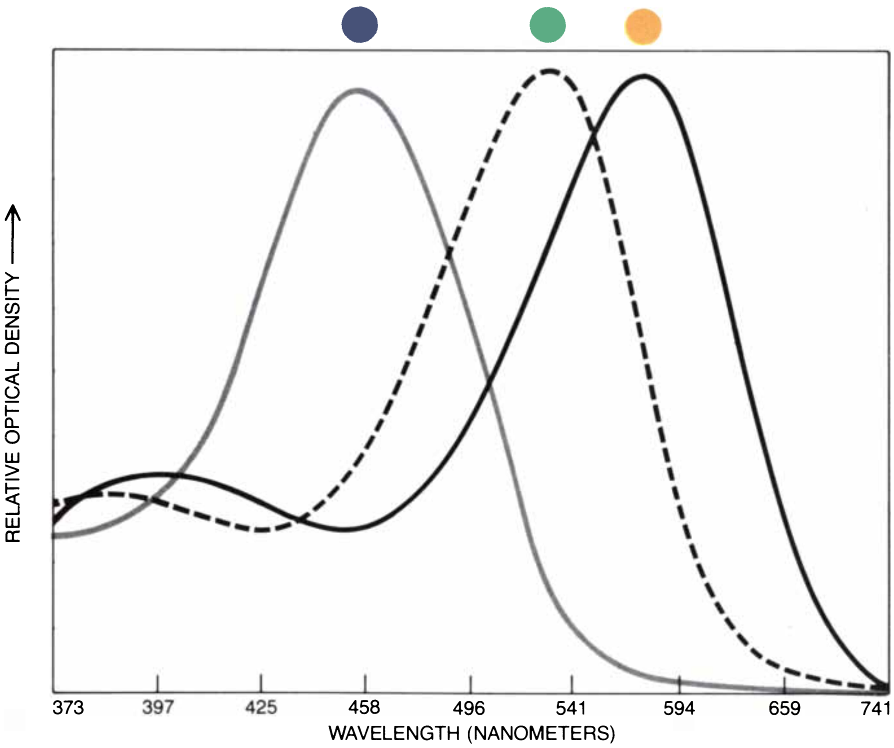
The rod pigments of coastal marine and freshwater species do not, however, follow the spectral distribution of available light as closely as those of deepwater species. The rods of Temperate Zone fishes whose habitat is shallow salt water are most sensitive to blue-green light with wavelengths between 500 and 510 nanometers, even though the local space light is yellow-green at wavelengths between 525 and 550 nanometers. Rod pigments in freshwater fishes from both deep and medium-depth habitats appear to be limited to a maximum absorption probability at wavelengths no longer than 540 nanometers, even though reddish orange light with a wavelength of 600 nanometers dominates the space light in their environment.

In 1966 J. N. Lythgoe of the University of Sussex suggested that the apparently mismatched pigments were set off from the color of the space light in order to maximize visual contrast. Lythgoe pointed out that when a pigment is matched to the space light, there is a relatively high contrast between a dark object, which deposits relatively few photons on the visual pigment, and the space light, which deposits many photons. A bright object, however, will deposit many photons on the pigment no matter what its color, and so it will not be strongly distinguished from the background space light. On the other hand, if the light illuminating the object is spectrally broad and the pigment is not matched to the space light, the situation will be reversed. A dark object and the space light will both appear dark, so that they will be relatively indistinguishable. A bright object will look bright, how-

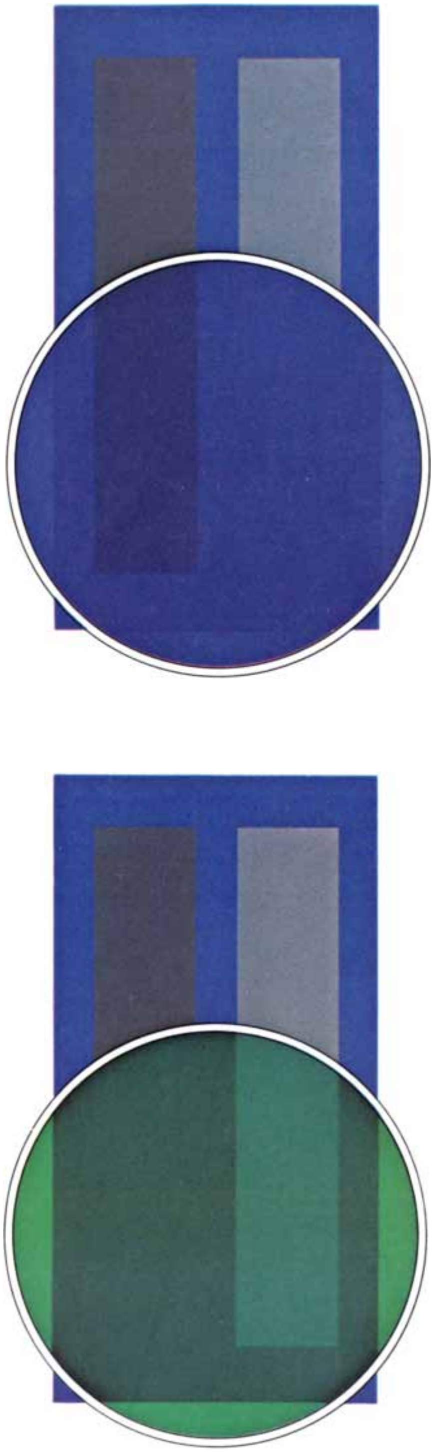




CONE CELLS of the black tetra, here magnified 500 diameters, differ according to shape as well as pigment content. Three kinds of cones are visible: large double cones, composed of separate cells, which carry the orange- and the green-sensitive pigments, and the smaller single cones, which carry the blue-sensitive pigment. A few long rod cells can also be seen; in the black tetra rods are maximally sensitive to green light. The photomicrograph was made by flattening the cells shown end on in the top illustration on page 141. The cells must be held flat between two cover glasses before the absorbance spectra can be made. In the diagram the position of the pigments in the cells is indicated by the colors to which they are maximally sensitive.



ABSORBANCE SPECTRA show, for each wavelength of light, the relative optical density of a visual pigment; they can be related mathematically to the probability that a photon of that wavelength will be absorbed by a visual pigment. Three absorbance spectra, one spectrum for each pigment of the cichlid fish (*Cichlasoma longimanus*), are graphed. The colors at the top of each curve indicate the appearance to the human visual system of the wavelengths of light to which each pigment is maximally sensitive, at 455, 532 and 579 nanometers respectively. For convenience in analysis the machine employed by the authors plots the absorbance of the pigments as a function of frequency. Hence when the same curves are employed to represent absorbance as a function of wavelength, the wavelength increments vary along the horizontal axis.



MULTIPLE-PIGMENT visual systems in fishes may initially have evolved because they enhanced the contrast of light and dark objects with the space light rather than because they were helpful in discriminating hue. For a monochromat whose blue-sensitive visual pigment matches the blue space light found in open-water marine habitats (*upper illustration*) a dark object is clearly visible against the background, but a light object is nearly invisible. A green-sensitive visual pigment that does not match the background (*lower illustration*) provides contrast between a light object and the background, although a dark object is nearly invisible. Fishes that incorporate both pigments can distinguish light and dark objects in their environment, even though they may lack the neural interconnections necessary to discriminate solely on the basis of hue.

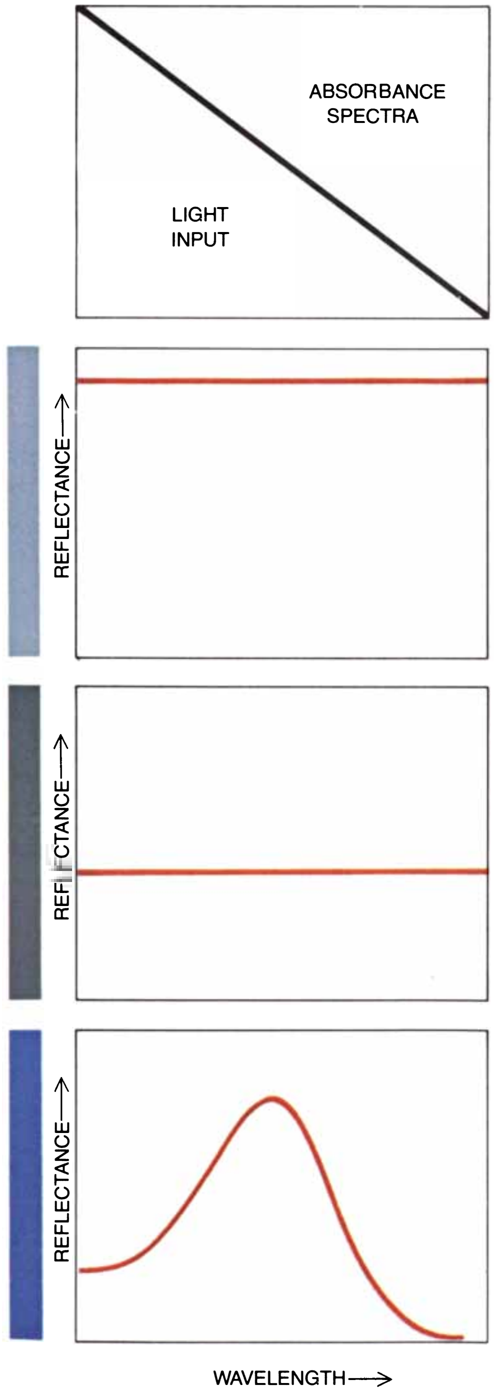
ever, and so now it will be strongly distinguished from the space light [*see illustration at left*].

The rod cells are specially adapted for low-intensity illumination. Since in a given species they usually contain only one kind of pigment, vision that must depend on rods is responsive only to variations of light intensity. To a fish with rod cells only (a rod monochromat) the world looks as though it were photographed on black and white film through a color filter. Dark blue objects are indistinguishable from dark red or dark green ones, and bright colors are easily confused with white.

Nevertheless, a monochromat can conform its behavior to take maximum advantage of the contrast between the space light and relatively dark objects. For example, the skipjack tuna (*Katsuwonus pelamis*) appears to carry only one visual pigment, even though it possesses both rods and cones. Munz and McFarland noted that the skipjack maneuvers so that it can rush up to attack its prey from below. Viewed from below the prey is silhouetted against vertical space light, and the skipjack, whose single-pigment sensitivity is matched to the space light, can exploit the contrast between its relatively dark prey and the relatively bright space light.

Many fishes, however, have two or three classes of cone cells in addition to their rod cells. Each class of cones is usually associated with a particular visual pigment, and microspectrophotometric measurements have shown that rod and cone pigments have approximately equal probabilities of absorbing photons. Nevertheless, because of differences among rods and cones and the nature of their interactions with other cells in the retina, activating cone vision requires light much more intense than that required to activate rod vision. Cones are generally associated with color vision, and the discrimination of colors bestows a clear evolutionary advantage in habitats where light is bright and made up of many colors. Individual cone cells, however, can distinguish colors no better than rod cells: their output too depends on the number of photons absorbed by their pigments but not on the wavelengths of the photons. Even when several kinds of pigment are present in the photoreceptor cells in the retina, additional neural "wiring" is required before discriminations can be made solely on the basis of color. Hence it is likely that the evolutionary advantages initially conferred by the cones were not directly related to color vision. The advantages of the cones probably derived instead from their better adaptation to bright light and from the fact that a multiple-pigment visual system makes visual contrast possible across a broader spectral band than a single-pigment system does.

Some saltwater fishes that live near the surface, such as the mahimahi (*Coryphaena hippurus*), have two different cone pigments, one that absorbs mostly in the green and one that absorbs mostly in the blue. Such a dichromatic visual system is sensitive to light across a broader spectral band than a monochromatic system is, simply because more pigments give rise to efficient absorption

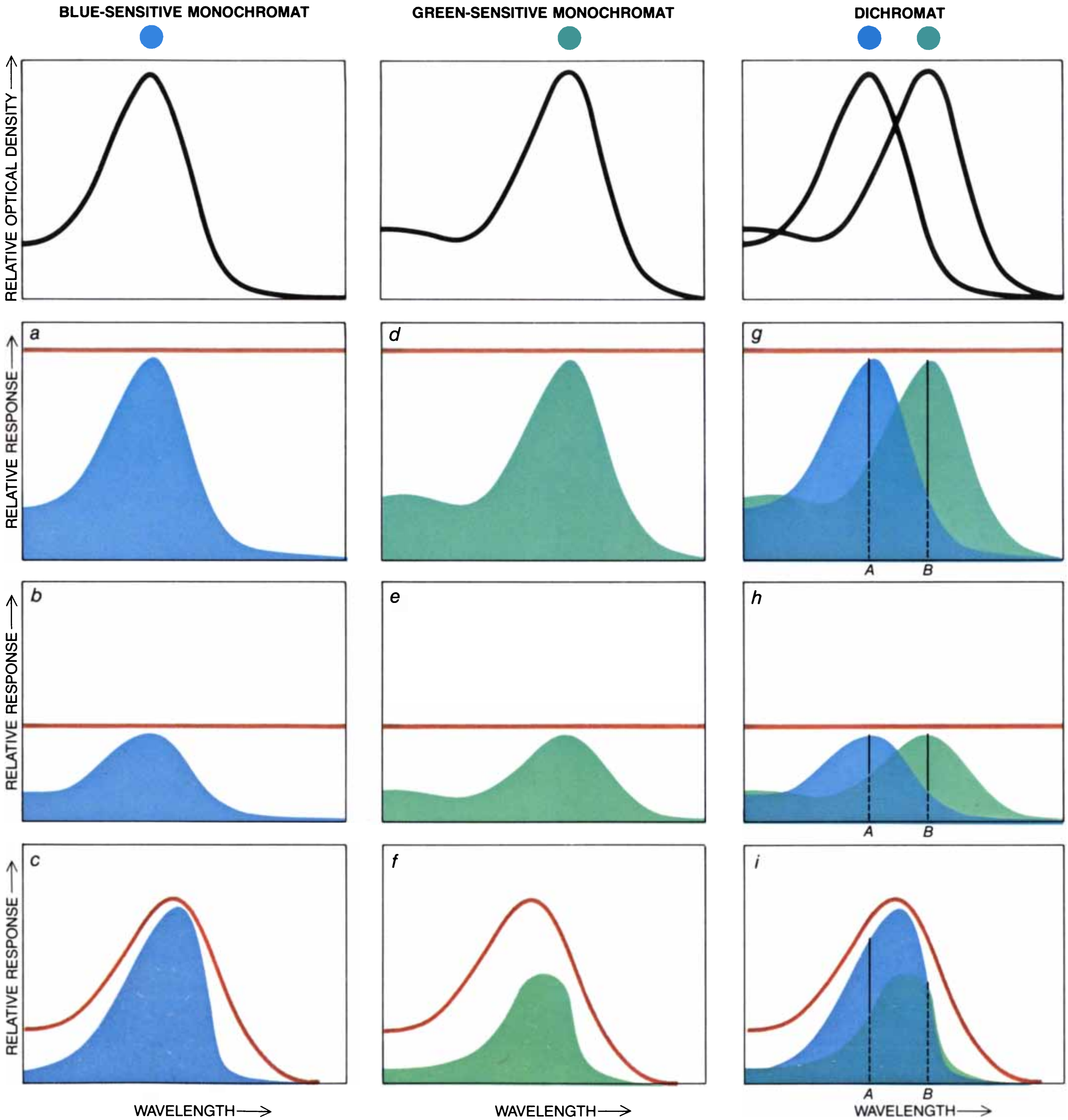


RESPONSE of a photoreceptor cell to a visual stimulus depends on the intensity of the stimulus and on the filter characteristics of the absorbing pigment. The wavelength of an incoming photon is irrelevant to the output of the cell once the photon has been absorbed.

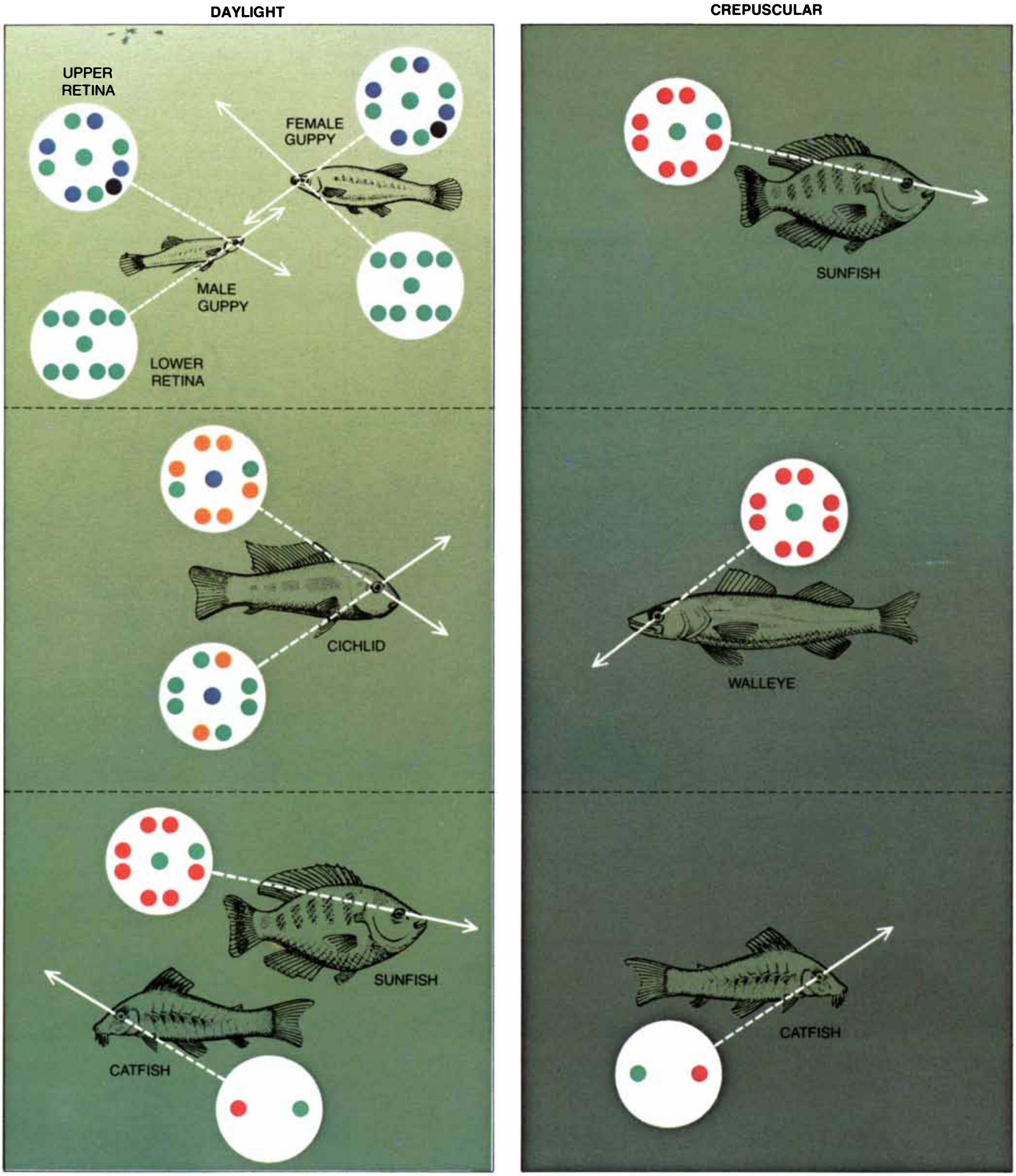
in more regions of the spectrum. When a saltwater dichromat looks upward, the green pigment is well matched to the space light, whereas when it looks forward or downward the blue pigment is the better match.

McFarland and Munz have contended, however, that from an evolutionary standpoint the major advantage of two-pigment visual systems over one-pigment ones is that two-pigment systems allow both dark and bright objects to be discriminated against the background space light. When a blue-and-green-sensitive dichromat looks upward, the green-absorbing pigment enables the fish to distinguish dark objects against the predominantly green space light; the blue-absorbing pigment allows bright objects to be differentiated. When the fish looks forward or downward, the roles of the two pigments are reversed, but the net effect remains that both bright and dark objects can be distinguished against the background.

Our data show that the cone pigments of teleost fishes correspond rather well to the illumination available in the fishes' local environment. The broadest range of visual pigments is found among



The table shows how the output of the cells (*areas shaded in blue and green*) varies with the input of the three colors shown in the illustration at the left. The output of the blue-sensitive cones is similar for the light gray and the blue inputs (*a, c*). Dark gray areas give reduced output (*b*), so that they would contrast with the background for a blue-sensitive monochromat. The output of the green-sensitive cones, on the other hand, is similar for the dark gray and the blue inputs (*e, f*), whereas the output for the light gray input (*d*) contrasts with the background. A dichromatic pigment system can also be the basis of color vision (*g, h, i*). A cell that monitors differences between the outputs of the two kinds of cones would distinguish light at wavelength *A* from light at *B*, no matter what the intensity of the light.



COLOR SENSITIVITY of fishes is adapted to environmental conditions and to the requirements of visually based behavior. Six visual habitats are illustrated for a freshwater lake; the fishes are portrayed in the type of visual habitat in which they are active. (Not all the species shown would normally be found in the same body of water.) The colored circles schematically represent the mosaic patterns of the photoreceptor cells in each species; the colors correspond to the maximum color sensitivity of each visual pigment present in the cells. The sensitivity to color shifts to the red with depth, corresponding to the increasingly predominant long-wavelength light in the water. Two species of fish show marked intraretinal differences in cell distribution. The upward-looking lower retina of the guppy (*Poecilia reticulata*) is dominated by green-sensitive cones. The guppy can therefore detect dark objects against the downwelling light, but the lower retina is not suitable for discriminating colors. The downward-looking upper retina is made up of all three types of cone and is therefore sensitive to a broad spectrum of colors. In mating displays the male guppy takes a position below or in front of the female so that his colors will be registered on the most color-sensitive area of the female's retina. The upward-looking lower retina of the cichlid is also primarily sensitive to middle-wavelength radiation, whereas its downward-looking upper retina is more sensitive to long-wavelength light. The differences correspond to the spectral distribution of the background space light along different lines of sight and to the variations among the visual tasks each species must perform. The retinal mosaic patterns of the sunfish (*Lepomis macrochirus*), the catfish (*Corydoras meyers*) and the walleye (*Stizostedion vitreum*) are all relatively uniform. The sunfish feeds on phytoplankton near the surface at dusk or dawn and follows the plankton downward during the day, so that the fish is active in visually similar environments throughout most of the day.

species that are active near the surface, where the light is bright and spectrally broad. In salt water at intermediate depths the red, long-wavelength light from the surface is almost entirely absorbed, and at such depths most saltwater species lack pigments that absorb red light. Their cones are maximally sensitive to blue and green light.

One exception to this trend is the sea raven (*Hemitripterus americanus*), which we studied with Hárosi and Barbara-Ann Collins of the Marine Biological Laboratory at Woods Hole, Mass. Sea ravens possess three pigments, maximally sensitive to blue, green and yellow-green light, which cover a much broader spectral range than one would expect in a fish that spends most of its life at depths of more than 100 meters. To understand the adaptive utility of a fish's visual system, however, one must consider the fish's total behavioral repertory. In the sea raven's breeding season, which is in late fall, the fish moves to shallow water and thus to brighter and spectrally broader light. The bodies of reproductive individuals turn bright yellow, orange and scarlet, and after fertilization the eggs are deposited on a species of sponge that is colored bright orange or yellow. Such behavior requires visual discriminations in the middle- to long-wavelength regions of the spectrum and so calls for a wide range of visual pigments.

In most freshwater habitats the utility of blue-sensitive cones diminishes with depth because the transmitted light from the surface is red-shifted instead of blue-shifted. The visual systems of the freshwater fishes we examined can be divided into four groups, based on the maximum absorption sensitivity of their visual pigments and on the physical appearance of their photoreceptor cells. In the first group are fishes such as the guppy (*Poecilia reticulata*) and the giant danio (*Danio malabaricus*), which confine themselves to the region directly below the surface. Their visual systems are similar to those of shallow-depth marine species; they are characterized by three visual pigments, maximally sensitive in the violet, the blue-green and the yellow-green.

The retinal sensitivity of the guppy is closely linked to its behavior patterns. The upper retina, onto which is focused light from below and in front of the guppy's visual field, is a mosaic of cone cells sensitive to three widely spaced spectral regions; it is well designed to make color discriminations. The lower retina, however, onto which is focused light from above the visual field, is lined with cells that all contain the same green-sensitive pigment. John A. Endler of the University of Utah noted recently that when male guppies display their conspicuous coloration in an effort to attract females, they do so from a position ahead of and slightly below the object of their desire. Such a position in the visual field of the female ensures that the image of the male falls on the most color-sensitive region of the female's retina. The green-sensitive lower retina is best suited to finding dark bits of food silhouetted against the predominantly green vertical space light.

The second group of freshwater fishes also have three kinds of visual pigment, but their maximum sensitivities lie in longer-wavelength regions of the spectrum than those of the pigments in the first group. Such fishes include the tiger barb (*Barbus tetrazona*) and the cichlid fish (*Cichlasoma longimanus*); their sensitivity to long-wavelength light is an apparent evolutionary response to the rapid loss of short-wavelength light in the heavily stained waters they inhabit.

The bluegill (*Lepomis macrochirus*), the walleye (*Stizostedion vitreum*), the piranha (*Serrasalmus* sp.) and other fishes in the third group have few if any blue-sensitive cells in their retinas, although they do have green- and red-sensitive cells. They are active primarily at dawn or dusk near the surface, although many species feed in deeper parts of the water column during the day. Hence their spectral environment remains roughly the same throughout their active period. The red-sensitive pigments of these fishes make them far more sensitive to long-wavelength radiation than human beings are. The so-called black water inhabited by many such fishes is actually illuminated with near-infrared space light that is visible to them.

The retinas of the species in the fourth group are made up entirely of rods and green- and red-sensitive cones. No blue-sensitive pigments have been found in these cones, and in some species green-sensitive pigments seem to be missing as well. The fishes probably retain red-sensitive cones only because the dim background space light is highly red-shifted; otherwise they would probably be rod monochromats. The species in the group include numerous catfishes (such as *Corydoras meyers*) and the ecologically similar red-tailed black shark (*Labeo bicolor*).

It is likely that color vision, the ability to distinguish light on the basis of wavelength as well as intensity, evolved some time after two-pigment visual systems were in place. To be capable of color vision a fish not only must have two or more visual pigments whose absorbance spectra overlap; it also must have neural interconnections that can compare the signals of the photoreceptors in an appropriate way. To understand how this is accomplished consider the effect of a monochromatic red stimulus at a wavelength of 630 nanometers on the eye of a cichlid. The cichlid's eye has three visual pigments, including a yellow-absorbing pigment at a wavelength of 579 nanometers and a green-absorbing one at 532 nanometers. No matter what the intensity of the red stimulus, the yellow-absorbing pigment will absorb more of the red photons than the green-absorbing pigment. On the other hand, a yellow-green stimulus at a wavelength of 540 nanometers will have approximately the same effect on the yellow-absorbing pigment that it has on the green-absorbing one, no matter what the light intensity of the stimulus is.

The output of a nerve cell that monitors both kinds of photoreceptor cell can be dependent on the difference between the two incoming signals. Such a cell would enable the cichlid to distinguish red from yellow-green, even though the brightness of the two colors as measured by one class of the cichlid's cone cells might be identical. Cells that process signals from photoreceptors of different classes were first discovered in fish retinas by the Swedish physiologist Gunnar Svaetichin and one of us (MacNichol) in the 1950's. They are called color-opponent cells, and they have been found at several levels of the visual system of many species.

Since color-opponent cells depend on the difference between photoreceptor signals, the discrimination of hue is most accurate in regions of the spectrum where the absorbance spectra of at least two visual pigments overlap and exhibit markedly different slopes. If the absorbance spectra have the same slope or if they do not overlap, the difference in output of the two associated classes of photoreceptors is a constant and hue discrimination is poor. The addition of a third pigment or even a fourth to a visual system, together with the appropriate color-opponent cells, can make hue discrimination possible in such "dead" spectral regions. Of course, a third or fourth pigment can also support the absolute sensitivity of a visual system to an even wider band of wavelengths than two pigments can.

Given the evolutionary advantages of hue discrimination over the ability merely to distinguish the intensity of light, why is color vision not universal among teleosts? For a hue-sensitive retina to function properly it must be bathed in enough light over a broad band of wavelengths to stimulate substantial signals from all types of cones. Any photoreceptor cells that do not emit signals leave "holes" in the visual image. Holes degrade the image, and so any cone pigments in the retina that do not contribute frequently to the visual image are actually dysfunctional. A fish that is active only at night or at twilight, or one that lives in the deep sea, seldom encounters enough light to stimulate cone cells at all. For such a fish, and apparently for many of man's primate relatives, the price of survival in the dark is the loss of sensitivity to color.

Henry A. Rowland

This 19th-century American physicist has been undervalued by history. Among his accomplishments was the building of the first engine for ruling the diffraction gratings on which modern spectroscopy depends

by A. D. Moore

"You know that from a child I have been extremely fond of experiment; this liking, instead of decreasing, has grown upon me until it has become a part of my nature and it would be folly for me to attempt to give it up...."

"I intend to devote myself hereafter entirely to science. If she give me wealth, I will receive it as coming from a friend, but if not I will not murmur."

With these words, in letters written to allay his mother's misgivings, Henry Augustus Rowland in 1868 embraced his calling. It had been intended that he would follow his father, grandfather and great-grandfather into the ministry of the Presbyterian church, but he was now a student at the Rensselaer Polytechnic Institute in Troy, N.Y. Within a decade his original experiments on questions with which he was already then engaged would win recognition in the international community of physics. In 1875, at age 27, Rowland was invited to join the founding faculty of Johns Hopkins University.

With the confident authority of one who could say "there is nothing more delightful than discovering truths for one's self," he set himself the task of pinning down the constants of physics: the mechanical equivalent of heat, the unit of electrical resistance and of its magnetic analogue permeability, the ratio of the electrostatic force to the electromagnetic force, the speed of light. Uncertainty about these values was at the center of theoretical controversies in physics in those days. Rowland's answer was to make better instruments.

In 1882 Rowland made and demonstrated the perfect instrument, his concave diffraction grating. A grating is a series of fine parallel grooves evenly spaced on a flat or curved surface at a tolerance much less than the wavelength of light. Light transmitted through or reflected from the grating is spread out in the order in which its component wavelengths excite the sensations of color vision, as the painter Thomas Eakins showed in his portrait of Rowland [*see illustration on opposite page*].

In order to scratch 14,400 parallel, evenly spaced three-inch lines per inch across five or six inches of the mirrored, gently spherical surface of his gratings Rowland invented and built the perfect machine, his ruling engine. To make that machine he had to make the perfect machine part. This is the screw that moves the grating, line by precisely spaced line, under the engine's diamond scribing tool.

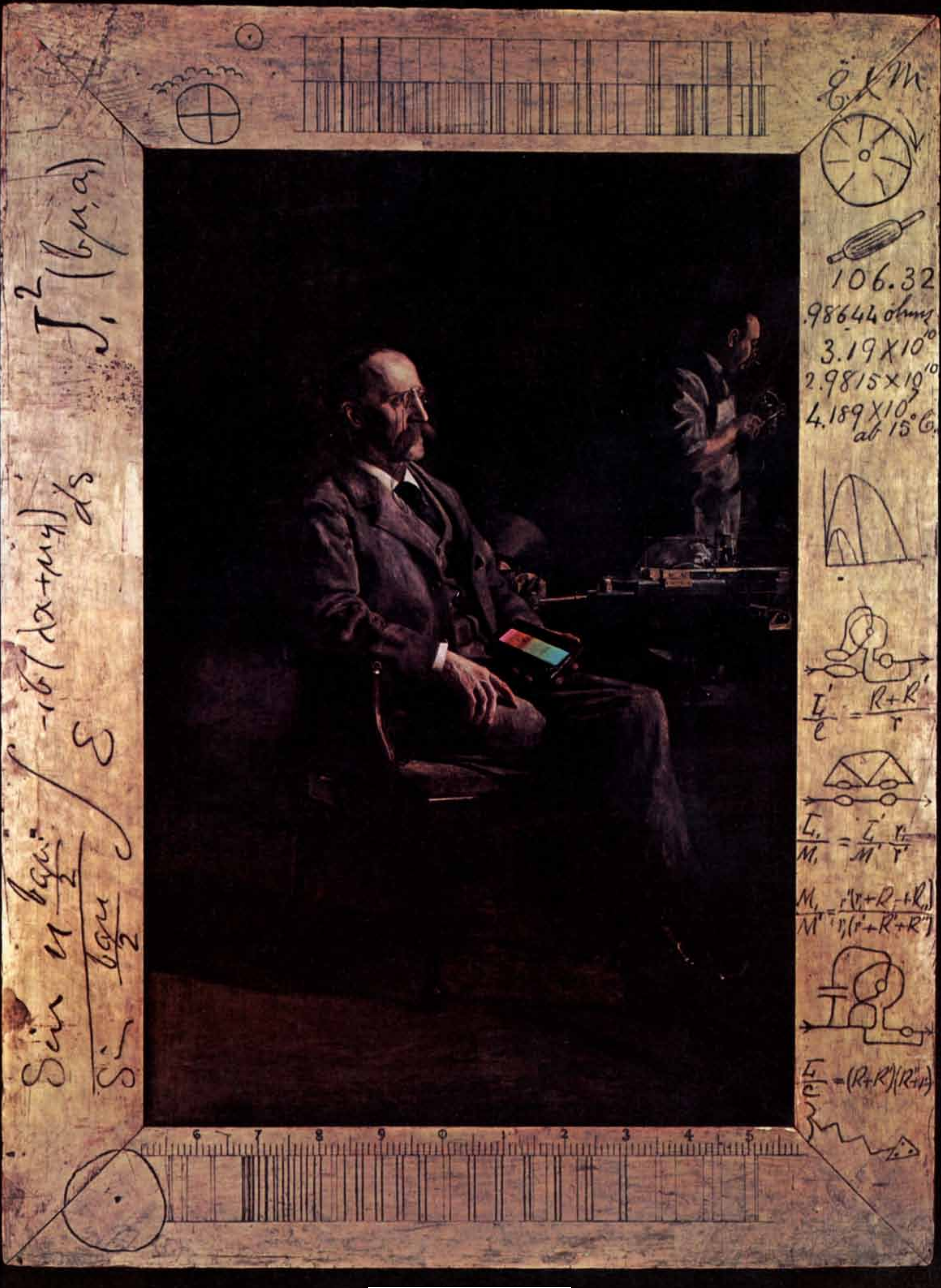
Rowland's ruling engines went on working long after his untimely death in 1901. By the middle of this century they had made more than 1,000 gratings. Over the past 100 years light diffracted by those gratings in laboratories and observatories around the world has taught us most of what we know for sure about the physical world.

The biography of Henry A. Rowland presents us with questions about the reliability of fame as a measure of a scientist's accomplishments. Even in a gathering of scientifically literate people many fewer hands will go up signifying recognition of Rowland's name than will respond to the mention of Josiah Willard Gibbs or Albert A. Michelson. Yet for the part Rowland played in his time and for the persisting and pervading influence of his accomplishments he deserves more than this biographical note to remind his beneficiaries of his useful and creative life.

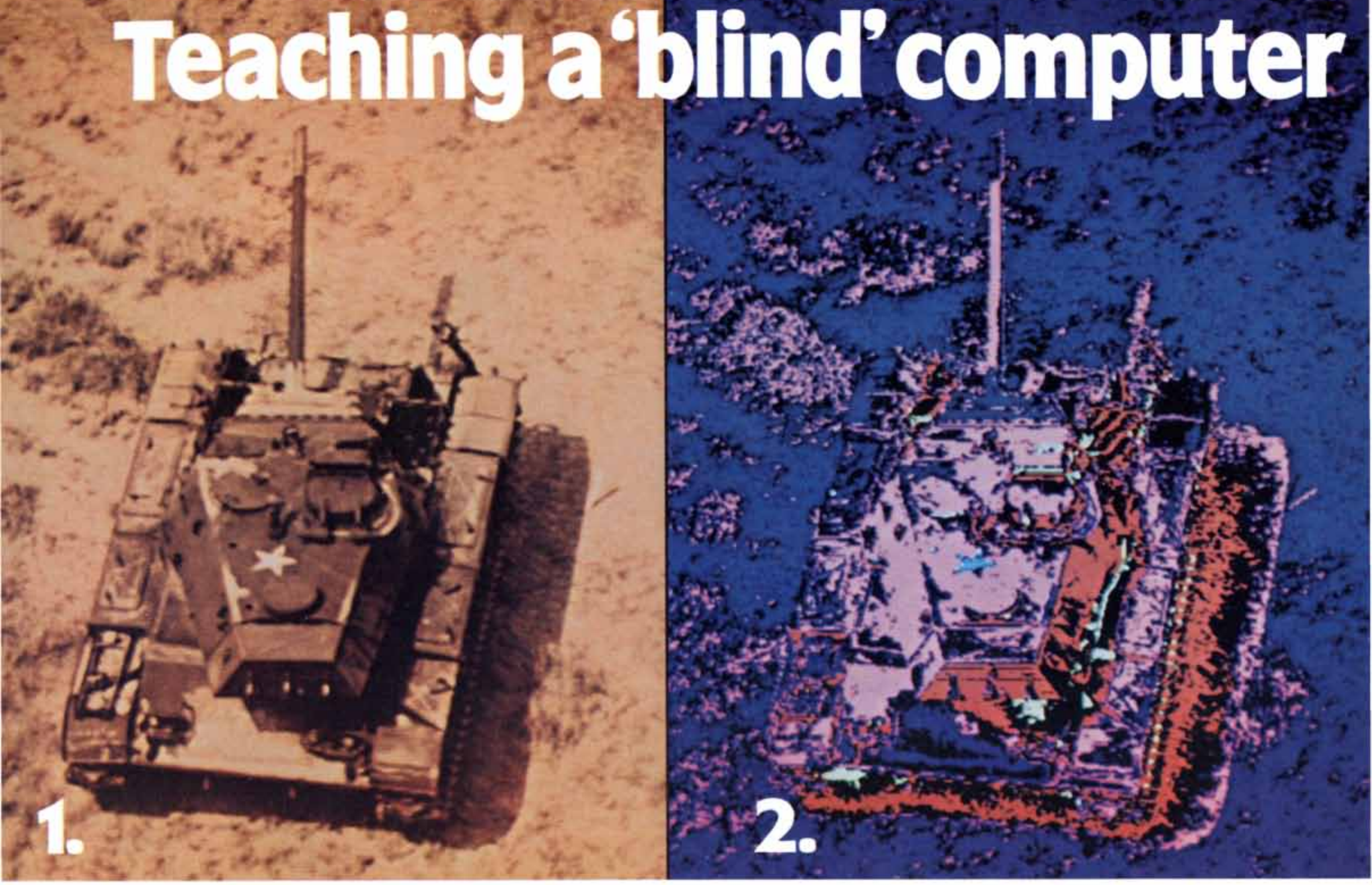
Rowland, born on November 27, 1848, at Honesdale, Pa., must have got his bent for science at least in part from his father. Strictly educated for the ministry at Yale University and at Andover Theological Seminary in Massachusetts, Henry Augustus Rowland, Sr., was nonetheless "interested in chemistry and natural philosophy, a lover of nature, and a successful trout fisherman." His death, when Henry, Jr., was 11 years old, left the boy to contend alone with his mother's strong will.

Starting at age three, when he made a model of a clock from a cardboard carton, Henry conducted experiments and fashioned apparatus that he insisted on demonstrating to his assembled family, his parents and two sisters, sometimes at their peril. At age 14 he began experimenting with the then still novel phenomena of electricity and magnetism. He made his own Leyden jars to supply current for a "shock" machine and for electromagnets of his own making. To insulate his wires, he wound them with layers of paper, a method later patented by someone else. By the time he was 17 he had made a furnace to melt iron to make his own castings, from which he

PORTRAIT OF ROWLAND BY THOMAS EAKINS hangs in the Addison Gallery of American Art at Phillips Academy in Andover, Mass., which Rowland attended. In his left hand he holds, ablaze with spectral color, one of the large diffraction gratings made with his ruling engine. In the background at the right is the engine itself, tended by Rowland's instrument maker Theodore Schneider. Eakins also made the frame of the painting, which he decorated with symbols of Rowland's work, photographed, drawn or written by Rowland himself. The two aligned spectra at the top of the frame show the bright *D* lines of sodium in the solar spectrum (*bottom*) and in a laboratory spectrum (*top*). The circle with the radial lines at the upper right is the disk in an experiment Rowland designed, at age 20, to detect the "convection current" postulated by Michael Faraday (*see illustrations on pages 156 and 157*). The numbers below the disk recall Rowland's measurements of the ohm, the speed of light and the mechanical equivalent of heat. The drawings and equations below the numbers pertain to Rowland's electrical work; the symbols are not conventional ones. The spectrum at the bottom, like one of those at the top, is part of the sun's. The round object at the left end of the spectrum is the symbol for the sun. The differential equation at the left side has to do with the sharpness of lines in a spectrum made by a grating. Above the equation is a Bessel function for one of the terms in it.


106.32
.98644 ohms
3.19X10^10
2.9815X10^10
4.189X10^7
ab 15°C.

Teaching a 'blind' computer



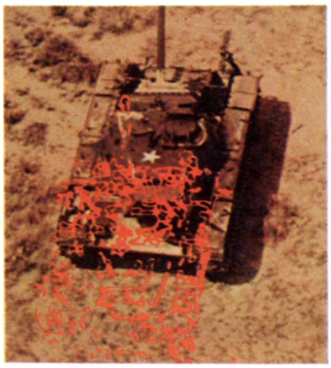
Lockheed knows how.

One of the growth technologies of the future is computerized image processing.

It's a vital technological thrust. More and more, industry, medicine, science and defense are making use of huge numbers of images gained from many sources—radar, infrared, acoustic, television cameras.

Until recently, most images have been analyzed or compared only by specialists. But now, computers are being taught to "read/understand" and make decisions about torrents of images that would otherwise swamp the human eye and brain.

"The tanks just moved."



The orange lines, which are offset from the original image of the tank, are part of the recognition technique that shows the tank has moved.

This can be critical input to a threatened commander, and he needs it almost instantly.

A reconnaissance aircraft or satellite carrying a sensor such as a television camera or radar would make the discovery. The sensor's images then are converted into electrical signals and sent to a computer, where they are stored as arrays of numbers. Computer programs manipulate the numbers until the tanks are "recognized." An enhanced image showing the tanks can be reconstructed and displayed for the commander.

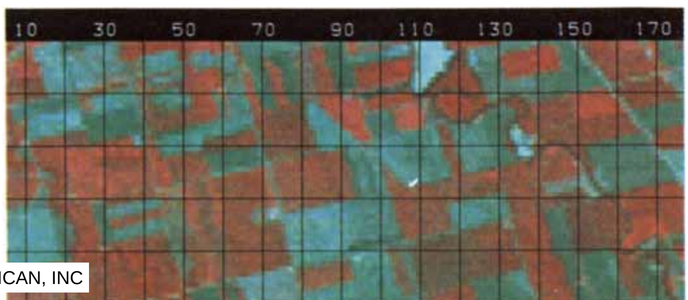
But a computer, unlike a human being, has to be given a mathematical description in advance to know what a tank looks like. The sensor, covering everything it sees, sends a staggering volume of signals. So the computer must be "trained" to eliminate unimportant images and concentrate on significant ones... for example, to look for the tanks that were spotted yesterday to see if they moved.

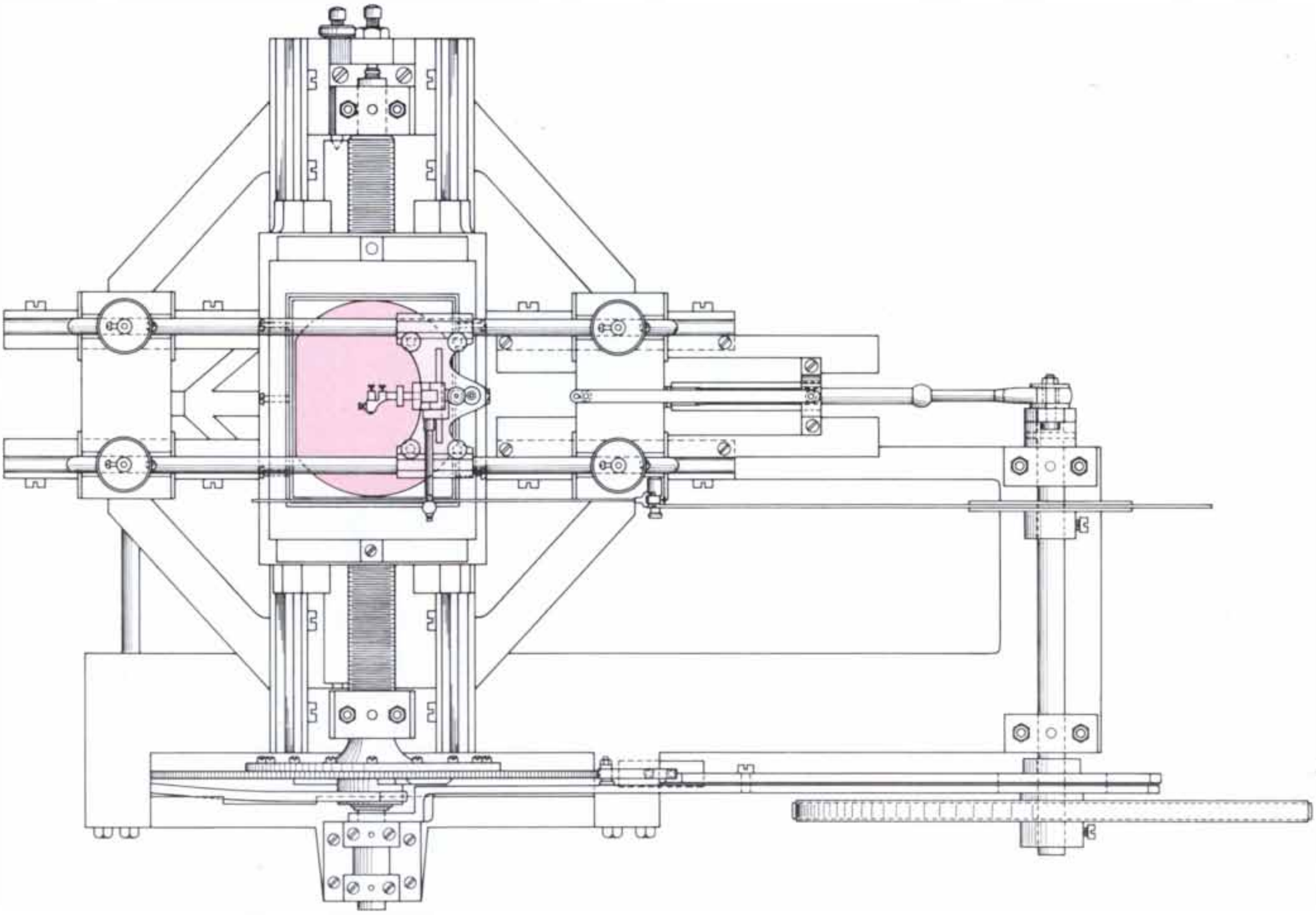
This is not just laboratory theory. Lockheed teams right now are training computers in this automated image understanding, or "computer vision."

How big is the wheat crop?

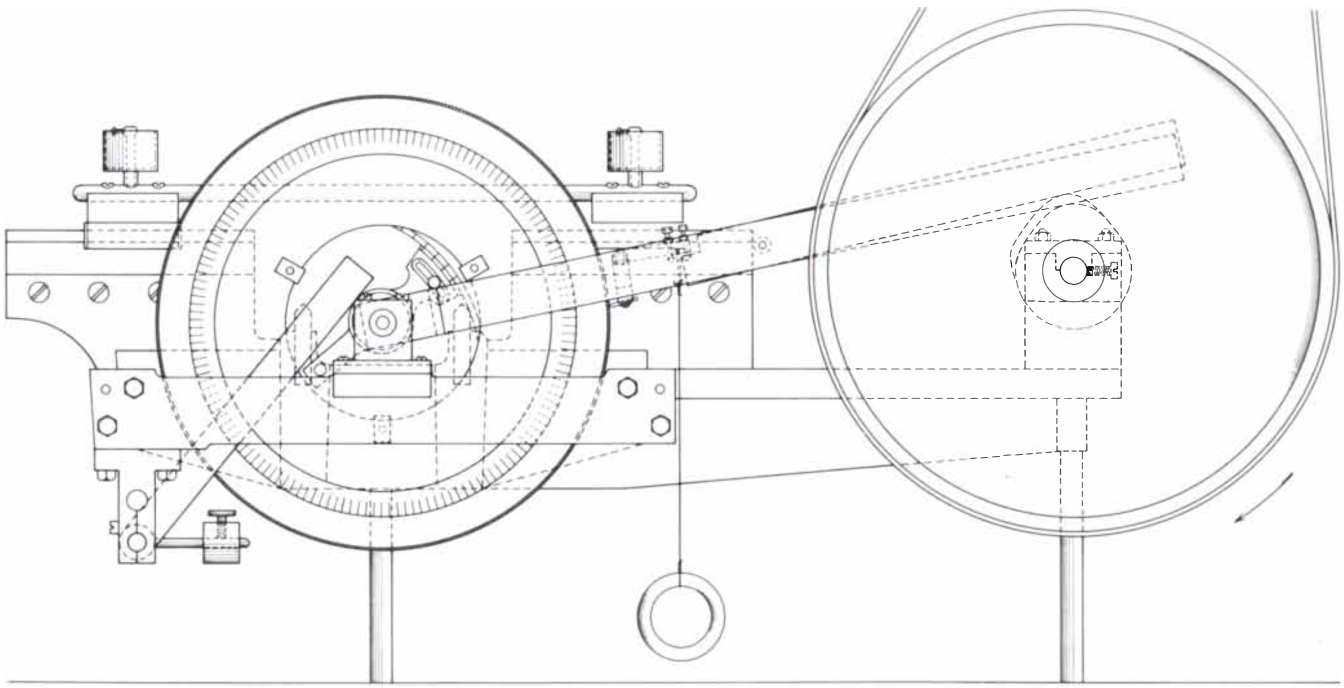
Identifying wheat from corn from sagebrush or other growth is another triumph of computerized image processing.

The photo here came from a Landsat satellite. The red color, assigned by a computer, indicates wheat. Bare ground, plowed ground, and stubble are in shades of green. Working with NASA, Lockheed scientists, engineers and mathematicians have developed techniques to process huge numbers of images through computers, which they have "taught" to recognize various crops and their conditions.





ROWLAND'S RULING ENGINE is shown in plan view. The two parallel horizontal rods at the upper left center hold the mounting of the diamond tool that is moved back and forth (to the left and right) to cut the grooves of the diffraction grating. The grating blank is the truncated circle (*color*) under the rods. Here the diamond tool is positioned at the center of the circle. Above and below the rectangular carriage holding the grating blank are the ends of the large lead, or driving, screw that moves the blank the tiny distance from one groove to the next. The carriage holding the diamond tool is moved back and forth by a crank mounted at the end of the shaft at the right.

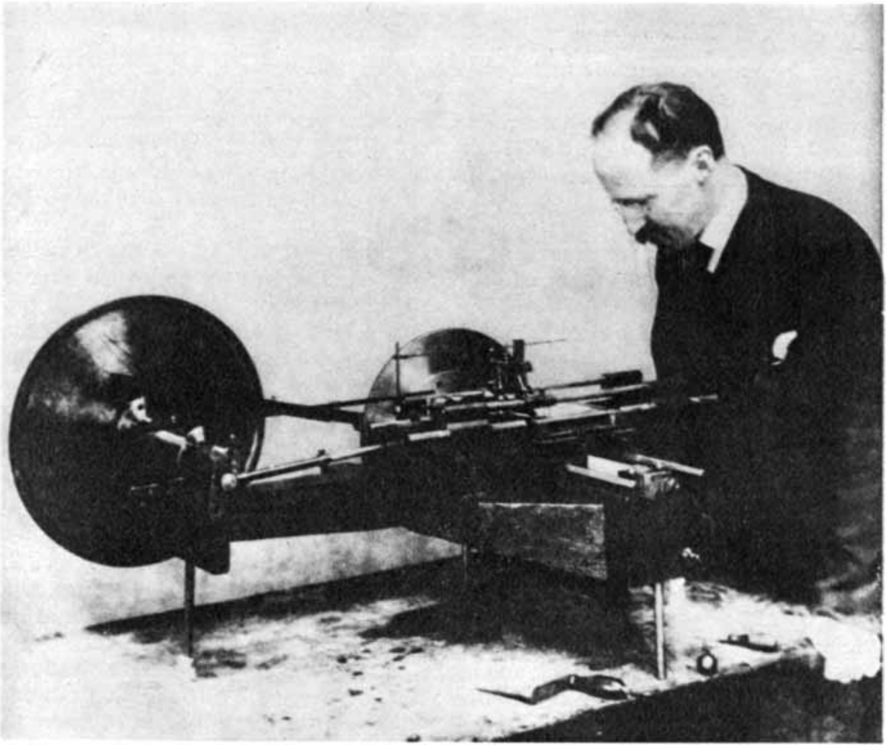


ELEVATION VIEW OF THE ENGINE mainly shows the mechanism that advances the lead screw. At the right is the wheel that drives the entire engine; in the original engine it was turned by a belt pulled by a small water motor. On the shaft of the drive wheel is a cam that lifts the large bar running diagonally across the drawing. At the bottom of the bar where it intersects the graduated wheel at the left is a ratchet that engages the geared edge of the wheel. The gear has 720 teeth. The graduated wheel is fixed to the end of the lead screw of the engine. Therefore when the cam lifts the diagonal bar, the ratchet turns the wheel and the lead screw an angular equivalent of one gear tooth, or half a degree. The ring hanging from a cord at the bottom is a weight to keep the ratchet engaged with the gear.

built a large electric motor, and he had made delicate instruments: an electrometer, a polariscope and an astatic galvanometer with its needle suspended by a new method to compensate for the earth's magnetic field.

One day, venturing outside of his first interest, he pasted up the pages of *The New York Observer* to make a hot-air balloon, which he lofted before the astonished neighborhood. When it descended in flames on a nearby roof, the Honesdale fire department had to be summoned to save the house.

None of this enterprise and the learning that went into it was encouraged by his schooling. Harriette Rowland held firm to her determination that her son be educated for the ministry. When Henry was 14, she sent him to Newark Academy in New Jersey. The principal, Samuel A. Farrand, agreed that yielding to Henry's penchant for science would be "like throwing her boy away." He himself undertook to tutor the boy in Latin and Greek. After three discouraging years Farrand gave up and recommended that if Henry was to be continued in classical studies, he should be sent to another school. Still unyielding, Harriette Rowland sent her son in 1865 to Phillips Academy in Andover, where, she hoped, the wintry principal Samuel H. Taylor would bring him to book.



ROWLAND INSPECTS THE ENGINE at Johns Hopkins University, where it was built. The crank mechanism that moves the diamond tool is visible in front of the wheel at the left. One of the engines is still in the department of physics at Johns Hopkins. There are two others, one at the Maryland Academy of Sciences and the other at the Smithsonian Institution.

Many years later, in an essay on this distinguished alumnus, an Andover instructor had to concede that Henry A. Rowland had "gained nothing from the school." The curriculum that confronted him for the first term was Latin Lessons, Latin Grammar and Latin Reading; for the second term Latin Grammar, Latin Reading and Latin Prose Composition, and for the third term Latin Grammar, Nepos (Roman history in Latin), Ancient History and Arithmetic. "Oh, take me home," Henry pleaded in a large scrawl at the end of his first letter from Andover to his mother. In a later letter he invoked Latin: "Non feram, non patiam, non sinam! [I shall not bear it, suffer it, tolerate it.]" Harriette relented, and by September he was installed at the Rensselaer Polytechnic Institute.

Named for its benefactor Stephen Van Rensselaer, one of the last patroon landholders of the Hudson Valley, this school was the first of its kind in the country and was then in its 40th year. It had developed, in response to the stirrings of industrial enterprise in the country and the building of its public-works infrastructure, a strongly technical curriculum leading to the Civil Engineer degree. Henry's letters from Rensselaer carried a happier message to his family: "I am getting along finely and like it first rate."

Although Rowland's studies at Rensselaer gave him command of Newtonian mechanics and of mathematics through the calculus of variations, he had to pursue his interest in electricity and magnetism on his own time and in his own boardinghouse quarters. His letters declare his satisfaction with his success in his schoolwork; he was leading his class in competition with students mostly older than himself, some of them veterans of the Civil War just ended. They tell mostly, however, of the instruments he was making ("a quite delicate ballance [sic], galvanometer and electrometer") and of the experiments he was conducting ("the prettiest is when electricity is sent through a tube in which a vacuum has been made by boiling mercury") and ask for funds to support his extracurricular enterprise ("it has not interfered with my standing at the Institute").

In 1868 Rowland wrote: "Made dynamo armature. The first continuous current armature ever made." His publication on his later development of this work put the fundamentals of dynamo technology in the public domain and thoroughly clouded the patent situation in that field. He was moved on another occasion, however, to ask his mother for the large sum of $200 ("perhaps from my patrimony") to get a patent on a "machine to send a number of messages by itself all night long." He did not pursue the possibility and did not think again of profiting from his work in science until the last decade of his life, when he attempted the commercial resurrection of this early thought invention.

Looking back on his experience at Rensselaer some years later, Rowland observed: "I do not remember ever learning a single thing about magnetism and electricity by word of mouth and... the first electric machine, galvanic battery, electromagnet, etc., etc., were those made by my own hands."

For a few months in his third year, apparently hoping to find a more congenial environment for his scientific enterprise, Rowland transferred to the Sheffield Scientific School at Yale. Back at Rensselaer he settled down in earnest to his lifework. He arranged his class schedule so that he could spend uninterrupted mornings on his reading and experiments. After the example of Michael Faraday, whose work and life now became his ideal, he began to keep bound notebooks. In these he logged his experiments, sketched designs for future experiments and made notes of ideas and conjectures that occurred to him from his reading, particularly from his reading and rereading of Faraday's *Experimental Researches in Electricity.* Under "Thoughts suggested by the reading of Faraday" he entered his earliest notions about how to attack two questions—how to measure magnetic permeability and how to detect the magnetic effect of moving a body carrying an electrostatic charge—in experiments that were to win his first fame.

Articled with his Rensselaer C.E. in 1870, Rowland found employment as a surveyor in the building of the Western

New York Railroad. From his discontent with this work he was rescued after a year by his appointment as assistant professor of natural science at Wooster University in Ohio. He was happy after another year to be called back to Rensselaer to teach, among other courses, the first at that institution in his own subject, electrodynamics. On the promise of $1,500 to equip his laboratory, he accepted the low salary of $1,200 per year.

Classroom teaching proved to be neither to Rowland's liking nor much to his credit. He was pleased that his course in his own subject, for seniors, was rated "the hardest in the Institute." He would not "put up with any nonsense" from freshmen and deplored a class made up of "all the bummers who spit tobacco juice." Rowland's students, in their yearbook for 1874, rated him "the smartest man in the Faculty." They went on to complain that he could not "impart knowledge to others.... No one can excel him in blind questions, and his explanations are useful only to himself."

Rowland made happier and more productive connections with students as the faculty sponsor of the Pi Eta Scientific Society, which he had helped to organize when he was an undergraduate. Here they saw him as "a pure devotee of science; he labors unceasingly in her shrine, tinkering and dickering from morn till night with his old apparatus." The students who were active in the society were the first to know him as a teacher by example rather than by precept. They were privileged to hear from his own lips his first reports on his researches.

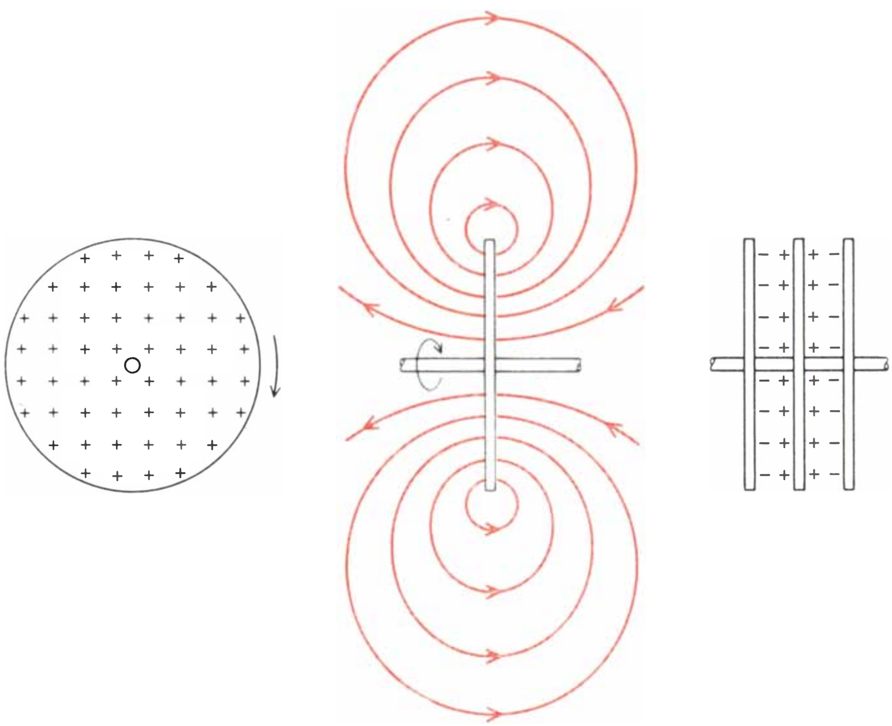
In his service as teacher Rowland managed to find not only what Charles William Eliot was calling at Harvard University at that time "the leisure to pursue research" but also the funding needed. Rensselaer, he complained, had provided only $500 of the money promised for apparatus, and his own outlay for books alone had mounted to $1,070. Not daunted, he got on with his investigation of magnetic permeability. "I have been working for the last week," he wrote in the winter of 1872–73, "in a little shed attached to the Institute with the thermometer down so far that my breath froze on my instruments."

Faraday had captured Rowland's attention five years earlier with his discussion of the resemblance between the magnetic field and the electric field. The distribution and concentration of the lines of force around a horseshoe magnet ("the sphondyloid of power"), Faraday observed, resemble those around the electric eel when it bends itself into a semicircle to surround and electrocute a small fish. (Faraday had used his bare hand to trace and estimate the intensity of electric force around the eel.) Other authors, surveyed in Rowland's voracious reading, had produced no full account of the relation between the magnetic field and the shape and material composition of the magnet and the strength of the energizing electric current in the case of the electromagnet. Working with a galvanometer "specially constructed for the purpose," Rowland meticulously plotted the magnetic field around cores of various materials excited by electric current of various strengths. He made ingenious correction for the shape of the magnet by measuring the field around toroids of iron and other materials in place of the magnets of infinite length implied by his equations. His experiments showed that the strength of the magnetic field varies in proportion to the exciting force and in inverse proportion to the resistance or permeability of the material, a quality characteristic of each material that varies with the exciting force, reaching a "maximum of magnetism" in each. Thus Rowland demonstrated for magnetic permeability a principle that nicely complemented Ohm's law for electrical resistance.

At this point in Rowland's career his personal bibliography showed nothing more than two letters to the editor published in *Scientific American*. The first letter written in his freshman year at Rensselaer, he had written "as a kind of joke" in learned explanation of the vortex at a bathtub drain. The second, written in the summer of 1871 when he found time off from laying out railroad tracks, exposed the "quackish nonsense" of a promoter who "pronounced the conservation of force an old-fashioned idea" and had imported perpetual motion into the new technology of electrical power. At the threat of exposure by Rowland, the promoter "intimated that... a force might proceed from his body which would act in conjunction with gravitation in causing me to be projected through the window and strike with violence on the ground below."

Beginning in 1871, Rowland had been sending reports of his work on magnetic permeability to *American Journal of Science*, published at Yale. That journal had just as often returned them, rejected. Its pages were given mostly to papers by natural historians (about "some crab with an extra claw which he had found on his morning ramble," as Rowland said of such authors). No one at Yale was up to his physics or mathematics.

In June, 1873, acting from the bold self-assurance that was the most striking facet of his personality, Rowland sent a comprehensive account of his work

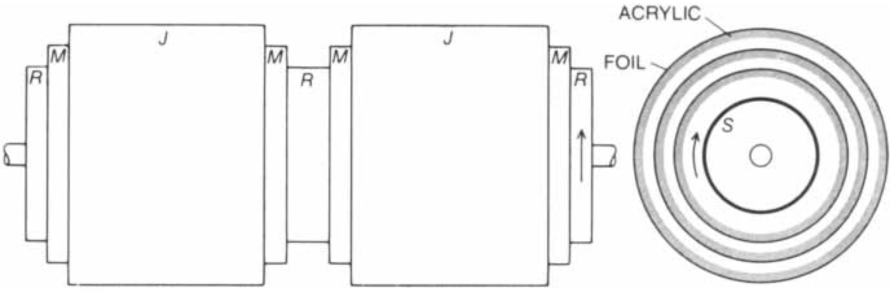


EXPERIMENT ON THE CONVECTION CURRENT was done by Rowland in the laboratory of Hermann von Helmholtz in Berlin. When a charged disk was rotated and the charges (*plus signs on the disk at the left*) were thereby moved, they constituted what was called a convection current. The question was: Would such a current set up a magnetic field (*field lines around the disk, seen from the edge in the middle*)? Faraday said that it would. James Clerk Maxwell said that there should be a field but that there was no proof, either experimental or theoretical. The weakness of the field had defeated all attempts to show its existence. Rowland was the first to demonstrate it. The three disks seen on edge at the right were a feature of his design; the two outer disks were oppositely charged to intensify the charge on the middle disk.

off to James Clerk Maxwell at the University of Cambridge, who had just published his definitive *Treatise on Electricity and Magnetism.* Maxwell must have been impressed by the brash young man's conversance with the literature, because the paper cited and dismissed the work of the foremost authorities of Europe, including James Prescott Joule and Heinrich Lenz, and he must have noted that Rowland's equations could be derived from his own. He arranged at once to get Rowland's paper published in *The Philosophical Magazine,* volunteering to read the proofs himself.

In letters to his family Rowland declared that he was now recognized as an "authority on magnetism" in Europe. "It will not be very long before my reputation reaches this country." By 1875 that hope was fulfilled.

Daniel Coit Gilman had resigned as president of the University of California to undertake the organization of Johns Hopkins University after the German graduate-school model. On a visit to the U.S. Military Academy at West Point, not far from Troy, he heard about Rowland and asked the young man over for an interview. They spent several hours walking in Kosciusko's Garden, Gilman recalled, "he telling me his dreams for science and I telling him my dreams for higher education." Maxwell's letters to Rowland were "worth more than a whole stack of recommendations," and Gilman signed Rowland on to organize the physics department of the new university. With classes not scheduled to begin until 1876, Gilman was able to offer him in addition three months of travel in Europe, to establish contacts with physicists there and assemble apparatus for the laboratory of his department. The three months stretched to nearly a year.



CONVECTION CURRENT COULD BE VERIFIED, perhaps with greater accuracy, by a new experiment proposed by the author. Here the design is cylindrical, consisting of concentric acrylic tubes with their outer surface covered with conducting foil. The jackets *J* and the middle tubes *M* are fixed. (Their supports are not shown.) The one-piece rotor *R* goes all the way through. When a power pack supplies to the jackets a 4,000-volt sine-wave alternating current at 50,000 hertz (cycles per second), the system becomes four capacitors in series. The charges on the rotor, when it is turned, become a.c. convection currents. The two sheets of current on the rotor amount in effect to two short opposed solenoids. A search coil (not shown) at the center of the unit at the left gets nearly all its flux from the left rotor current. It also gets a small opposed flux from the unit at the right. These fluxes can be calculated quite closely. With a coil of perhaps 200 turns and a rotor speed of 10 revolutions per second the coil electromotive force will be adequate for amplification (e.m.f. being proportional to both speed and frequency). The amplifier, with a tuned band-pass filter, will screen out variations in the earth's magnetic field and any other electromagnetic disturbances. The apparatus would have tube diameters of six, 6.5 and seven inches, and each of the two units would be about six inches long.

Rowland was welcomed everywhere as a peer and nowhere more warmly than by Maxwell at Cambridge. In Berlin he secured space for a winter's work in the laboratory of Hermann von Helmholtz by challenging the great man with his design for an experiment to test a significant but still unproved conjecture of Faraday's.

"If a ball be electrified in the middle of a room," Faraday had written, "and then be moved in any direction, effects will be produced as if a 'current' in the same direction had existed." The "effects" to be looked for were, of course, magnetic: a current in a wire sets up a magnetic field at right angles to its direction of flow. In his notebook, back in 1868 at age 20, Rowland had proposed to see whether a magnetic field could be detected in the vicinity of a spinning disk carrying an electrostatic charge. Helmholtz had tried a different experiment with null results, and he was glad to see Rowland try his scheme.

The strength of the expected magnetic field, however, was tiny. In accordance with Maxwell's equations, it was calculated to be the product of the charge and its velocity, but the velocity of the charge on the spinning disk would fall far short of the velocity at which a current travels in a wire, a velocity that is a substantial fraction of the velocity of light. Rowland had to detect a magnetic field of .00002 times the strength of the earth's magnetic field.

With his apparatus set up in a basement room in the Helmholtz laboratory, Rowland succeeded. Beyond this proof of the existence of Faraday's "convection current," he made measurements he thought were reliable enough to test the ratio of the electrostatic force to the electromagnetic force, the constant Maxwell had postulated to be equal to the velocity of light. His numbers came closer to the now accepted standard than most measurements made at that time.

Maxwell hailed the publication of this work of his protégé with the kind of mock-heroic verse he composed for such occasions, beginning:

The mounted disc of ebonite
Has whirled before, nor whirled
  in vain;
Rowland of Troy, that doughty knight,
Convection currents did obtain,
In such a disc, of power to wheedle
From its loved North and subtle
  needle.

In these last years before the discovery of the electron Rowland's convection-current experiment had significant bearing on the nature of electricity. The result was in favor of the classical "fluid" theory and comports well with the present understanding of electric current as the flow of electrons. The theoretical debate at the time, however, was too muddled to be settled by his experiment. Curiously, Rowland himself rejected the classical fluid in favor of a role for the ether, still to be defined, in "all ordinary electrical and magnetic actions." Moreover, Rowland's finding came into controversy when others failed to secure it independently. By the time he tried to repeat the experiment trolley cars had come to Baltimore and the electrical age was generating too much stray electricity to allow reliable readings.

(Here I must intervene to say that the Rowland convection-current experiment held special significance for me when I first learned about it not so long ago. Electrostatics has been my own work. It would surely shake the foundations of electromagnetism and much else if the experiment did not come out as Rowland thought it did, that is, if an electrostatic charge in motion did not excite a magnetic field. I have designed an experiment [*see illustration above*] that would take advantage of developments in materials and instrumentation since Rowland's day that I hope will verify his findings.)

"Give me time and apparatus," Rowland said to Gilman on his arrival in Baltimore in the spring of 1876, "and if our University is not known, it will not be my fault." Although it was set up at first in two former boardinghouses, the Johns Hopkins physics department was acknowledged to be the best equipped in the U.S., if not the world. The apparatus, furthermore, was for research and "*not* [Rowland's italics] for the illustration of lectures." When Rowland was asked what he planned to do with his students, he replied, "Do with them? Do with them? Why, I shall neglect them, of course!" He lost no time getting on with his work.

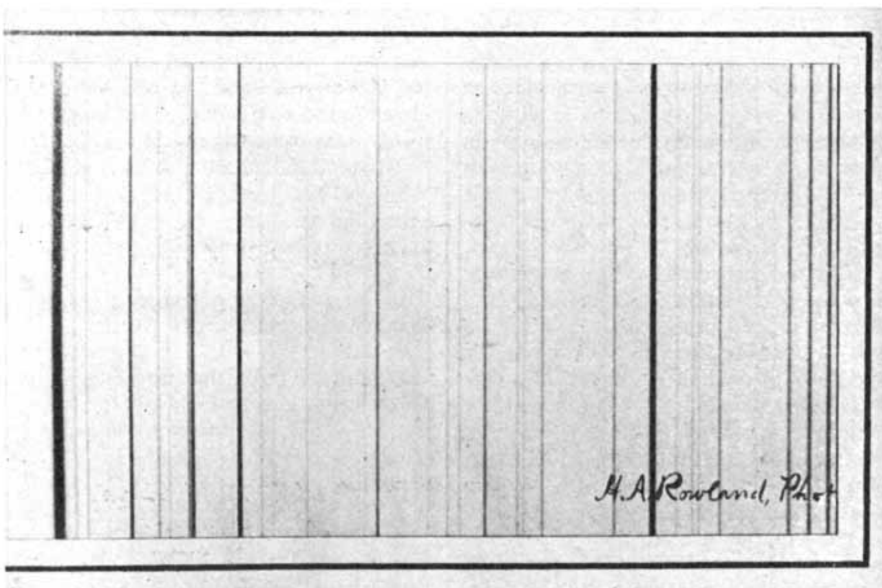
He turned first to the fixing of the mechanical equivalent of heat. This unit had been carefully measured three decades earlier by Joule (after whom it is called the joule). Yet the uncertainty as to its value was sufficient to prolong the life of the "materialistic" theory of heat (as a fluid akin to the electrical fluid) and to compromise the full acceptance of the law of the conservation of energy. "The number of units of work necessary to raise one pound of water one degree in temperature" has been called one of the "golden numbers" of the 19th century.

Rowland's strategy was to secure as large and as rapid a rise in the temperature of the water as possible in order to reduce the uncertainty due to the radiation of heat from the apparatus. Compared with the .64 degree Celsius increase per hour by which Joule had measured the constant, Rowland's apparatus achieved increases of up to 35 degrees C. per hour and a temperature excursion of 15 to 25 degrees within half an hour to an hour. He took care to measure each .1-degree change and so accumulated more than 100 observations with each run. "From any two [readings] the equivalent of heat could be determined," Rowland said, "and any single series has more weight than all Joule's experiments in 1849 put together." The strategy allowed not only correction for change in the specific heat of water with temperature but also precise measurement of the specific heat at each point. To drive at high speed the paddles that stirred the water Rowland built a special steam engine, and the great labor also required fundamental improvements in calorimetry and thermometry. For this work he was rewarded with tenure and an honorary Ph.D. by his university.

The principal yield from the Rowland laboratory's work on the constants of electricity and magnetism came with the demonstration of the Hall effect. This was the work of a junior colleague, E. H. Hall, in refinement of an experiment done earlier by Rowland on which he had not published. It established the existence of the electromotive force at right angles to both the current flow and its magnetic field.

Rowland had meanwhile addressed his energies to the perfection of spectroscopy. For this task he was prepared because, as he was heard to say more than once, he had never seen a machine, no matter how complicated, whose working he could not at once comprehend. He saw that the crucial part of a machine that would rule a grating was the lead, or driving, screw. The thread of this screw must be so nearly perfect that the rotation of the screw through a given angle—more precisely, through a given fraction of an angle—would result in the exactly corresponding forward motion of the grating carriage. A screw turned on a lathe necessarily inherits all the imperfections of the lead screw of the lathe. William A. Rogers of the Harvard College Observatory had invented a procedure for adjusting the position of the cutting tool to compensate for some of the periodic error in the lead screw. This left local errors for correction by grinding.

Rowland's invention started from the



SOLAR SPECTRUM made by "H. A. Rowland, Phot[ographer]" hangs in the Johns Hopkins physics department. Only a very small part of the entire visible solar spectrum is shown. The two prominent dark lines at the left and the right are the same two sodium *D* lines that appear in the spectra at the top of the frame of Eakins' portrait of Rowland on page 151.

recognition that a long, solid nut will quickly bind on a screw owing to error in the threading of both. To achieve a nine-inch length of perfect threading he made a grinding or lapping nut 11 inches long that he split lengthwise in four equal segments. With the segments reassembled in a collar the tightness of the nut could be adjusted through the course of the grinding, with emery powder and oil and then optical rouge as the grinding material. "Now grind the screw in the nut," the Rowland prescription reads, "making the nut pass backwards and forwards over the whole screw.... Turn the nut end for end every ten minutes and continue for two weeks." In this time the high spots on the threading of nut and screw had found one another and worn themselves away. Meanwhile endless other fussy measures had been taken to hold temperature constant, to counterbalance the weight of the grinding nut and so on.

"No workmanship is perfect," Rowland said; "the design must make up for its imperfections." He found the mounting of the nearly perfect screw "more difficult to make without error than the screw itself." For the nut that advanced the grating for the scribing of the next line he adapted his stratagem of splitting the nut. The two halves were then collared to ride along together but independently on the screw, an elegant mechanical feedback linkage correcting for periodic error in the screw. With 20 threads to the inch, each full rotation of the screw advanced the grating a twentieth of an inch; each of the 720 settings in each rotation advanced it an exact 14,400th of an inch.

The most sensitive instrument to detect periodic error in a ruling engine, Rowland wryly observed, is the grating it makes. Such error, he said, "must be perfectly eliminated since a periodic displacement of the lines by one one-millionth of an inch from their mean position will produce ghosts in the spectrum." Ghosts confuse the record not only by their presence but also by stealing light from the "real" lines in a spectrum, making faint ones indistinguishable from ghosts.

Gratings had been made for a century. Over the prism they had or promised to have such advantages as wider dispersion and higher definition, absorption of less light or virtually no light and a reach into the infrared and the ultraviolet. More or less egregious errors flawed them all. No grating was more than an inch or two across, a serious defect because definition is a function of width.

Once Rowland got his ruling engines running they could consistently turn out ghost-free gratings five inches and more across. He sold them for the cost of the wages of the skilled machinist who set up the engines and tended them through the six days of nonstop ruling it took to

plow the three miles of grooves on a five-inch grating.

Rowland's supreme invention is the concave grating. Spherical in curvature on a radius of eight to 21 feet, these gratings look nearly flat. Their ruling, however, calls for the engine to space the lines evenly on the chord of the arc, not on the arc itself. With these gratings no telescope is required to focus the light, and all the light reflected by the grating at each angle comes into focus directly on the photographic plate.

With his own concave grating Rowland proceeded to map the lines of the solar spectrum, using a micrometer with a screw perfected by his own method. His *Preliminary Table of Solar Spectrum Wave-lengths*, published in 1898, served until 1928 without revision as the world standard; its later revision at the Mount Wilson Observatory, on readings from a Rowland grating, introduced certain corrections, the largest correction amounting to one part in 30,000. Rowland also undertook to catalogue the known elements, getting the spectral signature of each. Comparing these with the solar spectrum, he was challenged by the number of "important lines not accounted for." He could see that some lines in the solar spectrum corresponded to unknown lines in the spectra of the earth's minerals. His work presented many questions to his successors, and he equipped them with the tool to find the answers.

Before the Physical Society in London in 1881, to demonstrate his gratings and report on his solar-spectrum work, Rowland declared, with no false modesty, that he could do as much in an hour as had hitherto been done in three years. Sir James Dewar, Faraday's successor at the Royal Institution, rose to say: "I struggle with a very mixed feeling of elation and depression: elation for the wonderful gain to science, and depression for myself; for I have been at work three years in mapping the ultra-violet."

Emerging as a public figure, wearing all the honors the scientific community could then bestow, Rowland used his status to urge, with passion, the cause of science to his countrymen. Speaking in Minneapolis at the 1883 meeting of the American Association for the Advancement of Science, he said: "It is not an uncommon thing, especially in American newspapers, to have the applications of science confounded with pure science. And some obscure American who steals the ideas of some great mind in the past and enriches himself by the application of the same to domestic uses is often lauded above the great originator of the idea who might have worked out hundreds of such applications, had his mind possessed the necessary element of vulgarity." In a more explicit reference to the press celebration of Thomas A. Edison, then at full tide, Rowland declared: "The spark of Faraday blazes at every street corner."

A bachelor until he was 42, Rowland was as strenuous in his leisure activities as he was at work. A successful trout fisherman like his father, he was also an accomplished horseman. He loved to ride off the bridle paths and used to say that every tree in the countryside was decorated with his eyeglasses. He rode to hounds with the local hunt clubs around Baltimore and raced his own hunters in their steeplechases. On winning one race and simultaneously learning he had won the substantial Venetian Prize for his work on the joule, he could not say which made him prouder, and he spent the Venetian Prize money to buy a group of rearing, plunging horses done in bronze in the style of the day.

In 1890, in a medical examination for an insurance policy soon after his marriage, Rowland learned he had diabetes. Medicine could then, many years before the isolation of insulin, do nothing about the disease. He gave up fishing and horsemanship for sailing during summers in Maine with his family.

The recognition that his days were numbered brought a more painful change in Rowland's career. Concerned to provide for his family, he sought consulting assignments, and at this time he secured a patent on the multichannel telegraph system conceived in his youth and undertook to launch it as a business. He was not entirely successful in these enterprises, which he found otherwise unrewarding. He bitterly regretted the time they took from his real work. Toward the end of his last decade he returned to the convection-current experiment that had so frustrated him. On his deathbed he heard the last encouraging findings from it.

Although Rowland found a crusty satisfaction in his reputation as a poor teacher, he was revered by his students, one of whom said, "To be neglected by Rowland... was more stimulating and inspiring than the closest personal supervision of lesser men." He was a teacher of the kind for whom Johns Hopkins University was created, one who taught apprentices by example rather than classrooms by broadcast precept. For his success there is one compelling measure. Of the 90 American physicists "starred" by vote of their colleagues in an edition of *American Men of Science* of the mid-1920's, 10 came from Cornell, eight from Columbia, six each from Clark and Princeton. No other university was named by more than five, except Johns Hopkins—which was claimed by 29! As a teacher, therefore, as well as a scientist, Henry A. Rowland must be reckoned among the founders of physics in America. The students of his students and their students as well must be at work everywhere, along with his gratings.

THE AMATEUR SCIENTIST

The "speckle" on a surface lit by laser light can be seen with other kinds of illumination

by Jearl Walker

The first gas lasers (some 20 years ago) not only made laser light but also called attention to a curious phenomenon: when the light fell on a surface, the surface often had a granular or speckled appearance. The speckled pattern, now simply called speckle, was surprising even though similar patterns had long been observed on surfaces illuminated by sunlight. Speckle appears whenever laser light scatters from a surface whose irregularities are about equal to the wavelength of the light.

Speckle from sunlight normally passes unnoticed because of its faintness and lack of contrast and motion. Recently, however, Robert N. Sollod, a professor of psychology at Cleveland State University, described to me an optical effect that I believe is a kinetic example of sunlight speckle. When he puts a spoon with a shallow puddle of coffee with cream in it in direct sunlight, he observes a dynamic, colorful array of flashing points moving randomly through the fluid. The array seems to me to be speckle created by the scattering of sunlight from the colloidal particles in Brownian motion in the liquid. Previous studies have suggested that such an effect would be unlikely or even impossible. If my interpretation is correct, Sollod's demonstration may be the only one in which Brownian motion can be detected by the unaided eye.

A speckle pattern generated by light from a helium-neon laser

A sheet of ground but otherwise transparent glass will help to explain speckle. A laser beam travels through the glass and illuminates a screen or a piece of photographic paper some distance away. Think of the initial beam as consisting of many parallel rays, with a wave propagating along each one. A unique characteristic of a laser is that the process whereby the atoms in the laser medium are made to emit light guarantees that the waves are all in phase, that is, they are all in step as they cross an imaginary line perpendicular to the beam. Such light is said to be coherent. Light from other common sources, such as an incandescent bulb, is incoherent. The waves crossing the imaginary line lack a fixed relation in phase and constantly shift in the difference between their phases.

Laser light passing through the glass scatters from the rough surface of the ground side of the glass. The scattered rays are no longer in phase and are unlikely to be traveling in the same direction. The change in phase is caused by the slight variation in the thickness of the glass because of the irregularity of the ground surface. The speed of light is effectively slower in glass than it is in air because of the time required for the light to interact with the atoms in the glass. The thicker the glass is, the more time is lost to the interactions and the longer the light takes to pass through the glass. Since the surface of the ground glass is rough, the rays emerging from the surface of it have passed through the glass with a variety of transit times. Hence they emerge with a variety of phases.

An example of what might happen is shown in the top illustration on page 165. Two waves traverse a sheet of glass that has a "step" on one side to represent roughness. The waves enter the glass and travel through it exactly in phase. Within the glass they are reduced equally in wavelength (because of the effective slowing of the light), and once they emerge they again have their former wavelength. For the sake of simplicity they are shown emerging with no change in their direction of travel. Because of the roughness of the glass one ray emerges exactly out of phase with the other. (This is one of many phase differences the waves could have.) Such a difference in phase determines how the waves interfere with each other at the screen.

Consider light reaching a particular point on the screen after emerging from points A and B on the ground glass, as

is shown in the bottom illustration on the opposite page. If the waves arrive exactly in phase, they interfere constructively, giving rise to the maximum possible brightness. Because of the rough surface of the glass and also because of the different distances from A and B to the point on the screen the rays are unlikely to arrive at the point exactly in phase. Their interference probably results in a brightness that is less than the maximum. If they arrive exactly out of phase, they interfere destructively, giving rise to darkness at the point.

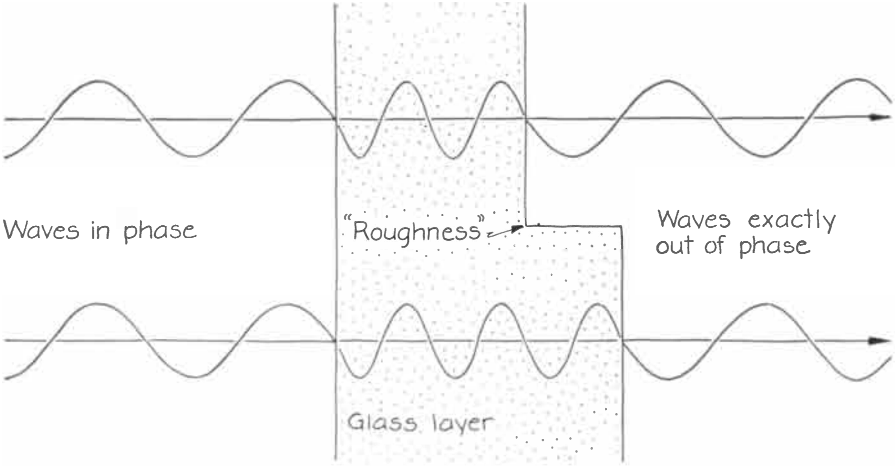
All the points on the ground glass contribute light waves to the point on the screen. The waves interfere with one another to determine the brightness perceived at the point. Other points on the screen also receive rays of light, which also interfere. The result is that the screen is dappled with a complex array of bright, dim and dark regions. This is the speckle pattern. It lacks any regular design partly because the surface of the ground glass lacks any.

The brightness at a given point on the screen remains constant only if the rays maintain a constant phase relation. If a wave crest emerges from A whenever a wave trough emerges from B, the rays from the two points always interfere in exactly the same way at the screen. If the phase relation shifts, the interference varies. A rapidly shifting relation causes a rapidly varying degree of brightness at the point. Since the eye and the brain average the brightness over many fluctuations, the pattern disappears, and the screen appears to be uniformly illuminated.

When an ordinary light bulb is substituted for the laser, the speckle pattern vanishes. Although at any given instant the light rays emerging from the ground glass have some particular phase relation with one another, the relation changes randomly in the next instant. The speckle pattern therefore varies randomly. The perceptual apparatus averages the patterns over time, and so no trace of an interference pattern remains (which is fortunate, because otherwise every common light source would coat its surroundings with speckle patterns).

Speckle patterns can be seen in light from a source other than a laser if the source occupies a sufficiently small angle in the field of view. For example, speckle can be seen in sunlight because the sun is a light source only .5 degree in diameter. A distant pinhole can serve the same purpose. (Although the pinhole might be illuminated by a light bulb, speckle can still appear if the pinhole subtends only a small angle as it is seen from the screen.)

The speckle pattern from sunlight or a pinhole illuminated by white light is complicated by the fact that the light consists of the full visible spectrum. It is simpler to consider light filtered to a sin-



A change in the phase of two waves of laser light caused by surface roughness in glass

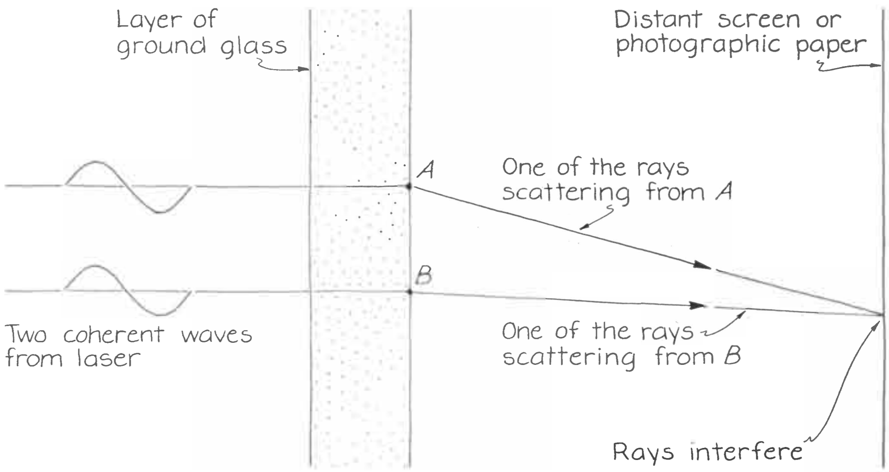
gle wavelength. The creation of speckle by a nonlaser source depends on the spatial coherence of the light rather than on the temporal coherence. Suppose you could sample the light passing a particular point on its way to the ground glass. Temporally coherent light would show at the sampling point a continuous variation between crest and trough, meaning that the frequency of the wave is constant. Such light must originate in a laser because there are no such naturally occurring sources on the earth.

Suppose you also sample light simultaneously at another point equally distant from the source. If you find that the waves passing the two sampling points maintain some particular phase relation, the light is said to be spatially coherent. (Sometimes it is said to be transversely coherent.) For example, if a wave crest passes through one point whenever a wave trough passes through the other, the two waves are spatially coherent. The measure of the extent of spatial coherence is the distance by which the two sampling points can be separated without destroying the fixed phase relation. The contrast of speckle depends on the extent of the spatial coherence of the light. Since laser light has great spatial coherence, it yields high-contrast speckle.

Light from a pinhole illuminated by a bulb has some spatial coherence if the pinhole is relatively distant from the sampling points. The smaller the angle the light occupies in the field of view from the sampling points, the greater the spatial coherence of the light. When light traveling from a distant pinhole passes through ground glass on its way to a screen, it creates a speckle pattern of interference.

At any given instant the rays from the glass all have some phase relation. An instant later the light has no particular phase relation to the earlier light, but still all the rays emerging from the glass have the same phase relation to one another. The interference pattern is maintained and perceived.

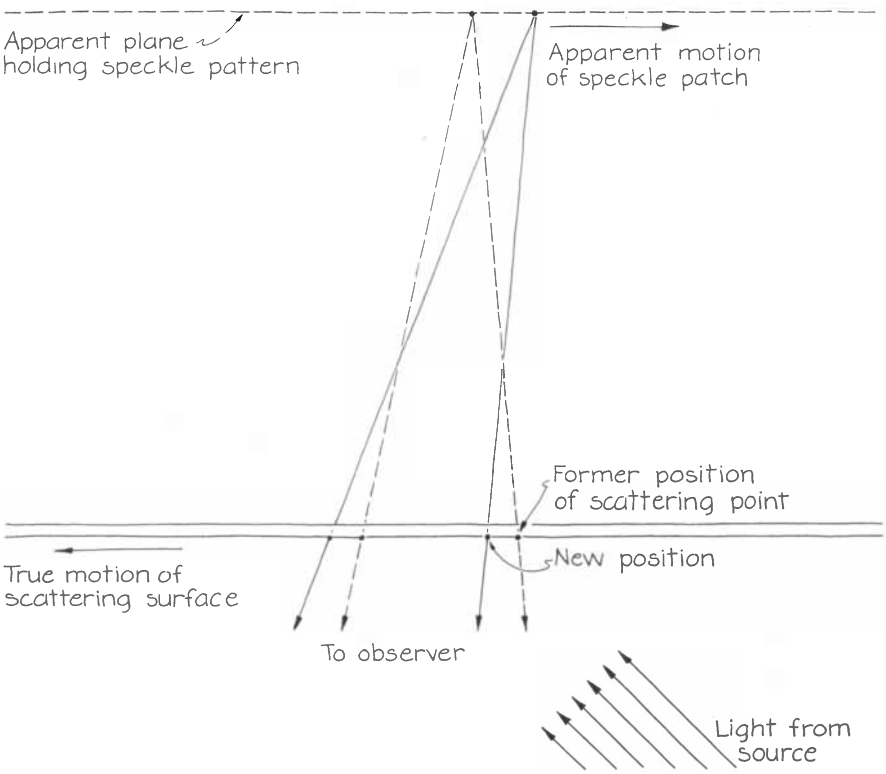
As the pinhole is brought closer to the ground glass the light illuminating the glass becomes less spatially coherent and the contrast of the speckle wanes. Eventually the spatial coherence is essentially lost. Then the rays emerging from the glass have no fixed phase relation and the speckle pattern disappears.



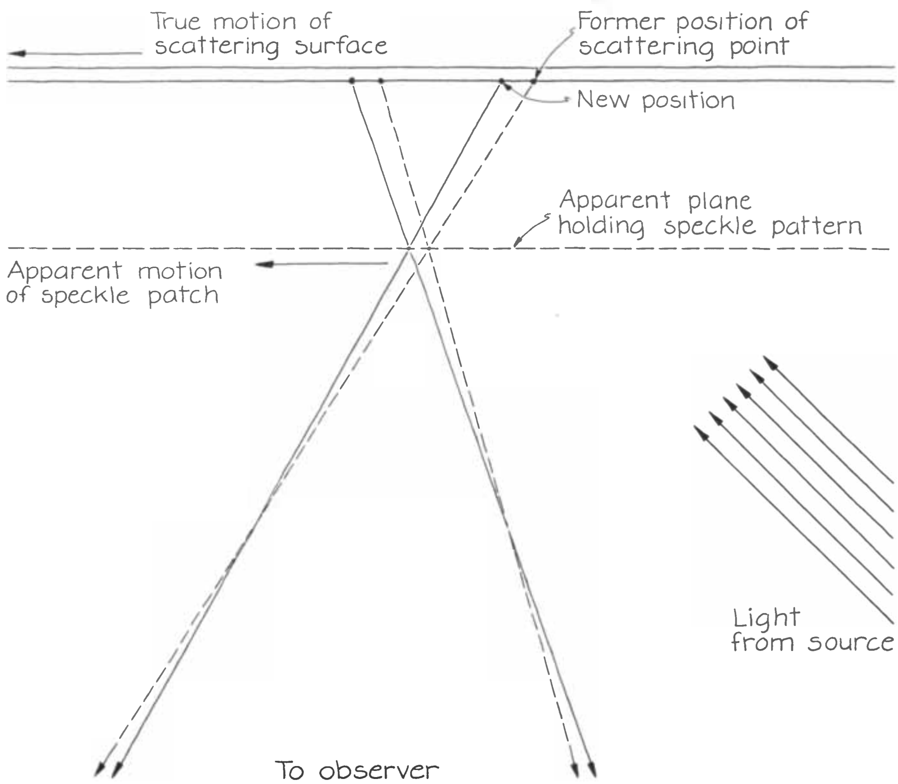
How two laser waves passing through ground glass can interfere

Speckle was first studied in the reflection of laser light from a rough surface. When spatially coherent light reflects from the surface, the rays of light reaching an observer have traveled over paths of differing length. Although they may have approached the surface while they were exactly in phase, they are no longer all in phase when they reach the observer. Hence they interfere with one another at the observer's retina.

White light from a pinhole or the sun provides a multicolored speckle pattern because the light has a broad spectrum of colors. The ultimate phase difference of the interfering rays depends partly on their wavelength. Therefore a slightly different speckle pattern appears for each color in the spectrum. The result is colorful but lacks contrast.



The apparent movement of a speckle pattern when the observer's visual focus is overlong



The apparent movement when the observer's focus is short of the lighted surface

Eugene Hecht of Adelphi University employs a black viewing surface to see speckle patterns in sunlight. Spray flat-black paint on smooth paper. Examine the illuminated paper at an angle of about 45 degrees. With a little practice you can see a multicolored, fine-grained pattern on the black surface. Hecht suggests that once you have seen the pattern you will also see it on many common objects illuminated by direct sunlight. His examples include "a tarnished coin, the weathered hood of an old car and even a fingernail."

Vincent P. Mallette of the Georgia Institute of Technology has described a speckle pattern in the light from a carbon arc. Speckle appears if the scattering surface is relatively distant from the arc. As the surface is brought closer to the arc the light falling on the surface becomes progressively less coherent spatially and the pattern disappears.

The apparent size of the elements in a speckle pattern derived from a reflection depends in part on the aperture of the eye. When the pupil is wide, rays of light within a large range of angles can enter the eye. They interfere with one another at the retina to create a pattern with relatively small elements. With a narrower pupil the range of angles of the rays decreases. Fewer rays enter the eye, and they generate at the retina a less complicated interference pattern with larger elements.

You can demonstrate how the speckle size depends on the size of the aperture through which the pattern is observed. Examine the size of a speckle pattern arising from reflection as you sight through an aperture that can be varied in size. As you decrease the aperture the speckle elements should appear to grow larger.

When either the observer or the reflecting surface that exhibits speckle moves quickly perpendicular to the observer's line of sight, the speckle pattern disappears. Within the averaging time of human perception the moving speckle pattern smears into what seems to be a uniform reflection from the scattering surface. If the motion is slow, however, the speckle pattern usually seems to migrate across the surface. The movement can be to the left or to the right.

The early reports of speckle attributed the direction of movement to the condition of the observer's vision. If the observer moves his head to the right (or the reflecting surface moves to his left), a nearsighted observer sees the speckle migrate to the left. For a farsighted person it moves to the right.

In the early explanation the apparent motion was attributed to parallax, which displays its effect when you can see two objects, one near and one far. Move your head to one side. The objects do not retain their alignment. You might interpret the realignment as an apparent motion of one or both of the objects. If you do assign motion to them, the near-by object will be given a motion opposite to the movement of your head. The more distant object will be assigned a motion that is in the same direction as your head. In truth neither object moved. Given enough clues you know this, but without clues you might believe the objects really did move.

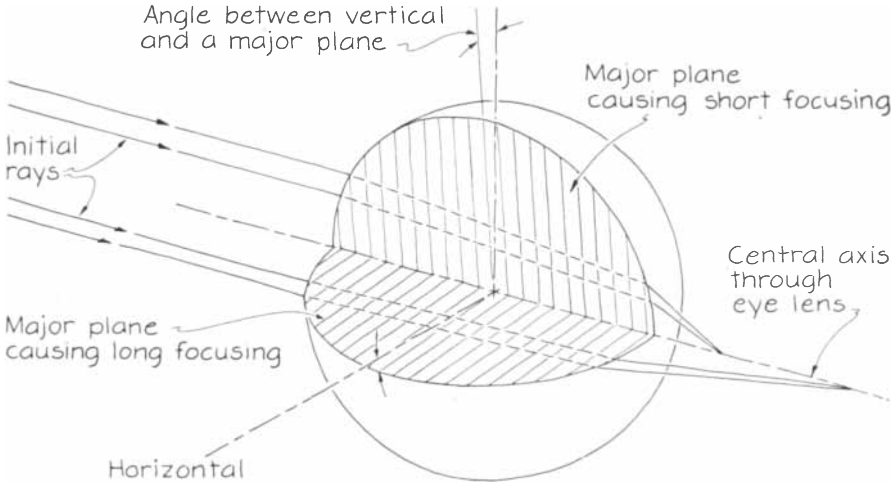
In the demonstration of speckle motion the apparent motion derives from the accommodation (focusing) of the eye. A person with far sight or normal sight can accommodate to focus a distant object on the retina. When the perception lacks clues to the distance of the object, the brain allows the eye to focus as if the rays were coming from a considerable distance. This is what happens when a farsighted person or one with normal vision studies a speckle pattern. The speckle is truly a pattern created directly on the retina, but the insufficient accommodation of the eye gives the illusion that the pattern is distant.

In this situation the screen is closer to the observer than the apparent position of the pattern. When the observer moves his head to one side, the more distant object (the speckle pattern) appears to move in the same direction. If the reflecting surface is moved perpendicular to the line of sight of a stationary observer, the speckle appears to move across the screen in the opposite direction.

A nearsighted person cannot focus on an object at a large distance. Suppose the screen is beyond the maximum focusing distance for such an observer. The speckle pattern will be interpreted as lying at the maximum distance of focusing. When the observer moves his head, the closer object (the speckle pattern) seems to move in the opposite direction.

In 1965 Douglas C. Sinclair of the University of Rochester pointed out that this explanation is incomplete for a person with normal sight. Under proper conditions the apparent motion of the speckle pattern for such a person depends on the color of the light. In the early days of gas lasers all the laser emissions were in the red. Reports on the apparent motion of speckle assumed that the same results would be obtained with other colors. Discrepancies appeared, however, when lasers emitted in the blue and green.

Suppose a person with normal vision examines the apparent motion of two speckle patterns, first one from a blue laser beam and then one from a red laser beam. The person is to focus on the scattering surface, which is also illuminated by ambient white light from the room. The red speckle pattern appears to lie behind the scattering surface and the blue pattern in front of it.



The geometry of light passing through an astigmatic eye

The difference arises from the chromatic aberration of the eye. Light of a long wavelength (such as red light) is refracted less by the eye and requires a longer focal distance inside the eye than light of a shorter wavelength (such as blue). This difference means the apparent position of a red speckle pattern is beyond that of a blue speckle pattern.

When the observer moves his head across the line of sight, the apparently distant red pattern moves in the same direction as his head and the apparently closer blue pattern moves in the opposite direction. If the color of the laser light is closer to the center of the range of colors to which the eye responds, the speckle pattern appears to be closer to the scattering surface. The apparent motion of the pattern when the observer's head moves becomes progressively harder to discern. When the pattern seems to lie at approximately the same distance as the scattering surface, the pattern does not migrate as a whole across the surface when the observer's head moves. It does display a random motion that is often called boiling, since it resembles bubbles reaching the surface of boiling water.

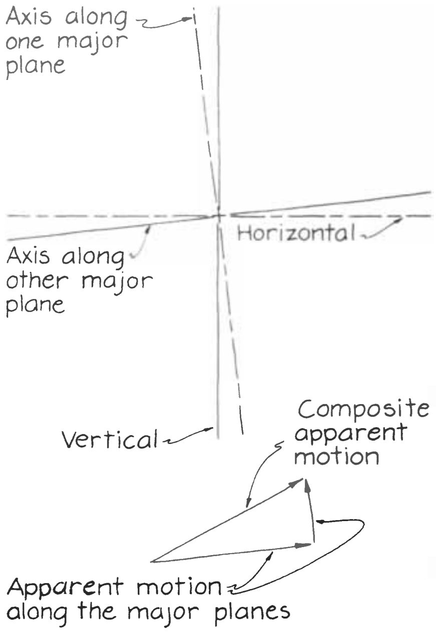
A more detailed examination of the migration of speckle involves the motion of the points scattering the rays responsible for a bright patch in the pattern. The upper illustration on the opposite page shows a rough surface scattering two rays to an observer. Suppose they result in a bright patch in the speckle pattern. Accommodation by the observer's eye makes the rays seem to emerge from a single point on an imaginary plane behind the actual scattering surface.

Suppose the surface moves to the left across the observer's line of sight, but not so much that the bright patch loses its identity in the changing pattern. The apparent location of the single origin of the two rays also moves, but in a direction opposite to the true motion of the scattering surface. When the observer's eye is underaccommodated for the scattering surface (that is, it is focused for something behind the surface), the apparent motion of the speckle is opposite to the true motion of the scattering surface and in the same direction as the true motion of his head. The reverse is true when the eye is overaccommodated. The apparent speed of the speckle patch depends on its apparent distance from the scattering surface.

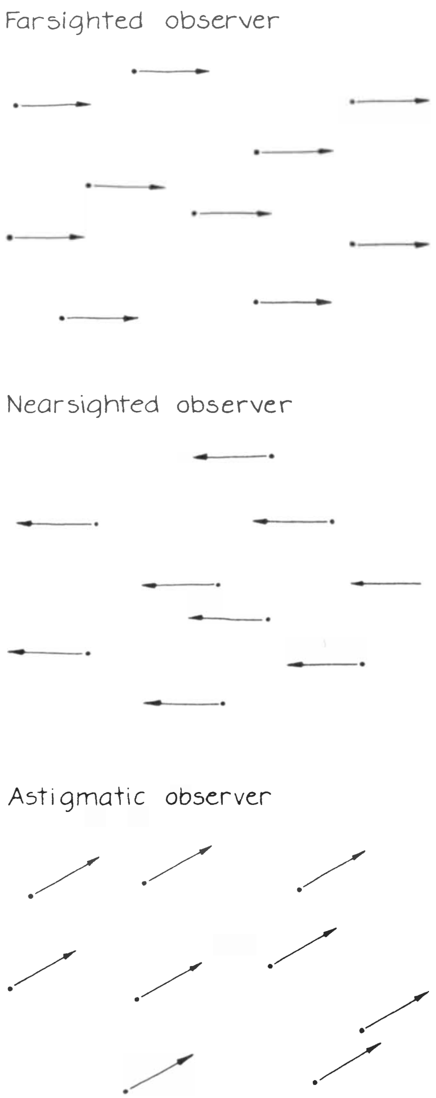
In neither instance does a patch of the pattern actually travel completely across the scattering surface. Soon after a patch seems to have traveled a short distance it disappears because the motion of either the scattering surface or the observer's head changes the geometry of the rays entering the eye. Fresh patches pop up, however, and move as the former patch did. The observer therefore has the illusion that all the patches survive for a full trip across the scattering surface.

When the eye is accommodated for the scattering surface (the surface is in focus on the retina), motion of the surface or the head results in no organized migration of the pattern. Instead the individual patches in the pattern move across the line of sight until the angles of the scattered rays change so much that the patches lose their identity. With continued motion of the surface or the head new patches appear, move in some random direction and then disappear. The surface appears to boil.

Some investigators have suggested that the observation of speckle patterns might replace the conventional tests of eyesight, particularly for people who cannot read the letters employed in the



Speckle motion affected by an astigmatic eye



Possible apparent motions of speckle

normal test. A surface is moved across the field of view of an observer. Laser light of low intensity is scattered from the surface to provide a speckle pattern. The observer notes the direction of migration of the pattern. Lenses of increasing strength are then inserted into the visual field until the direction is reversed. Reducing the lens strength, the examiner finds the strength at which the speckle appears to boil on the scattering surface. This is the strength of lens needed to correct the observer's vision.

Some observers encounter another peculiarity in the perceived motion of the speckle patterns. If the eye of the observer is astigmatic, the pattern might move at an angle to the horizontal axis along which his head moves. This two-dimensional motion of the speckle arises because of different focal lengths for rays crossing the eye through different cross-sectional planes.

The astigmatic eye is normally described as having two major planes through it. They are depicted schematically in the illustration on the preceding page. Initially parallel rays crossing the eye along one of these imaginary planes are focused at a certain distance inside the eye. Initially parallel rays crossing the eye in the other plane require more distance for focusing. Although the two planes are perpendicular to each other, they are probably not horizontal and vertical.

Rays contributing the speckle pattern to an observer enter the eye at a large range of angles. To simplify the situation I consider the rays as crossing the eye in one or the other of the imaginary planes through the eye. When the observer moves his head horizontally, the speckle patches appear to move partly parallel to the short-focus plane and partly parallel to the long-focus one. The speed of the patch appears to be greater along the long-focus plane. The brain interprets the composite motion as being along a line lying between the two planes. Since that line is likely also to be tilted to the horizontal, the patches seem to move along a tilted line when the astigmatic eye moves horizontally.

I was fascinated by Sollod's description of kinetic colors in a spoon of coffee with cream in it. Previously published accounts of speckle had dismissed the possibility of seeing speckle in solutions. For example, one study concluded that speckle cannot be seen when spatially coherent light is reflected and scattered from the surface of milk unless the milk is frozen. In principle speckle should be visible because each of the colloidal particles in milk scatters light to an observer. Since the path lengths of the various scattered rays differ, the observer should see an interference pattern.

Speckle is missing, however, because the colloidal particles are constantly moving in random directions. The motion, called Brownian motion, is due to the constant jostling of the particles by the surrounding molecules, which have thermal motion. At any given instant the array of particles does send a speckle pattern to an observer, but the continuously shifting position of the particles smears out the pattern over the averaging time of the perception process. The result is that the surface of milk seems to reflect light uniformly. The speckle patterns would emerge if the colloidal particles were slowed in their random motion. Freezing the milk essentially halts the motion. The speckle pattern is then stationary, as it is in all my preceding examples.

I believe Sollod's demonstration is one in which speckle can be seen in a fluid still free to flow. The trick is to arrange for a very thin or diluted layer of the fluid to be illuminated with spatially coherent light. Light scatters from a few particles on its way to the reflecting surface of the spoon, scatters from a few more on its return path to the observer and then creates the speckle pattern on the observer's retina. A thicker or more concentrated layer has too many particles in the way. The light reflecting from the spoon is feeble and the pattern is smeared.

A very thin layer of milk may have another function. The colloidal particles in it have slower Brownian motion not because the molecules are any cooler but because the viscosity of the fluid is greater when the fluid is in a thin layer. The slower motion shifts the fluctuating speckle patterns into the time range in which individual patterns can be perceived.

The specks of light from Sollod's arrangement are colorful because the speckle patterns depend on the wavelength of the light. At one instant the observer might be at the correct orientation to receive blue light from a particular particle. In the next instant the particle has moved and green light might be sent to the observer from a particle somewhere near the former position of the first particle.

When I hold a spoon partly filled with milk in the sunlight, I sometimes have trouble spotting the kinetic speckle pattern. Over the dry areas of the spoon lies a faint, stationary speckle pattern as on most objects illuminated by direct sunlight. Only at the edge of the pool of milk is the milk thin enough to create the kinetic speckle pattern. To study it better I gently swirl the milk in the spoon so that most of the surface is coated with milk. As the milk drains back into the pool, the coating on the rest of the spoon thins. Gradually these thin layers break up into kinetic speckle arrays. After a while the layers dry completely and the speckle pattern becomes stationary and less colorful.

I spot kinetic speckle patterns in thin

layers of milk when the layers lie on reflecting surfaces such as a mirror or the spoon in Sollod's demonstration. Any surface painted flat black does not work. I believe the reason lies in how light is scattered from the colloidal particles. The scattering appears to be more intense in the forward direction than it is back toward the observer. Hence the light scattered by the particles should be reflected by a surface in order to reach the observer if the speckle pattern is to be bright enough to be perceived. If the surface fails to reflect the light, the kinetic speckle is probably too dim.

To verify this hypothesis I caused sunlight to pass through a thin layer of milk in a transparent plastic dish. When I looked at the transmitted sunlight from below the dish, I found the familiar kinetic speckle pattern. From above the dish no speckle was apparent. The demonstration showed that the particles in the milk were scattering the light primarily in the forward direction.

Suspended particles much larger than small molecules are essential for kinetic speckle patterns. Thin layers of water give no kinetic displays, but a thin layer of water doped with a little milk does. The kinetic speckle apparently requires particles in the size range of roughly a micrometer.

Scattering from molecules of water, which are considerably smaller, does not result in speckle. One reason is the almost continuous distribution of these molecules. Another is that they move faster than the larger particles. Both the distribution and the speed result in a uniform reflection. Particles larger than about 100 micrometers do not generate kinetic speckle patterns because they are too large to be set moving by Brownian motion. They do, however, yield stationary speckle patterns.

Last August I described interference patterns that one can see in a misty or dusty mirror. The patterns require light that has at least partial spatial coherence. Laser light, sunlight or light from a pinhole will serve. As a final note I mentioned that the central sections of the pattern called the Fraunhofer rings (or corona) should be uniformly textured but instead had noticeable streaks. The streaks are part of a speckle pattern created by the scattering of the light from the dust motes or water droplets. Coherent waves can scatter from adjacent motes or droplets and then travel to the observer to interfere at the retina. The result is that the observer sees either bright or dim light or darkness at the position of the motes or droplets. The intensity depends partly on the distance separating the two motes or droplets. Since the separations between other pairs of scattering particles are sure to be different, the interference varies between bright, dim and dark over the entire scattering surface.

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

METAMAGICAL THEMAS

FADS AND FALLACIES IN THE NAME OF SCIENCE. Martin Gardner. Dover Publications, Inc., 1957.

CULTS OF UNREASON. Christopher Evans. Delta Books, 1973.

ESP AND PARAPSYCHOLOGY: A CRITICAL RE-EVALUATION. C. E. M. Hansel. Prometheus Books, 1980.

FLIM-FLAM. James Randi. Harper & Row, Inc., 1980.

SCIENCE: GOOD, BAD AND BOGUS. Martin Gardner. Prometheus Books, 1981.

THE BIOLOGICAL EFFECTS OF LOW-LEVEL IONIZING RADIATION

LEUKEMIA AND IONIZING RADIATION. E. B. Lewis in *Science*, Vol. 125, No. 3255; May 17, 1957.

HUMAN HEALTH AND THE ENVIRONMENT—SOME RESEARCH NEEDS. Department of Health, Education, and Welfare. U.S. Government Printing Office, 1977.

SOURCES AND EFFECTS OF IONIZING RADIATION: REPORT TO THE GENERAL ASSEMBLY. The Scientific Committee on the Effects of Atomic Radiation. United Nations, 1977.

THE EFFECTS ON POPULATIONS OF EXPOSURE TO LOW LEVELS OF IONIZING RADIATION. Advisory Committee on the Biological Effects of Ionizing Radiation. National Research Council, National Academy of Sciences. National Academy Press, 1980.

THE CONTRIBUTION OF IONIZING RADIATION TO CANCER MORTALITY IN THE UNITED STATES. Seymour Jablon and John C. Bailar III in *Preventive Medicine*, Vol. 9, No. 2, pages 219–226; March, 1980.

EFFECTS OF LOW-LEVEL RADIATION AND COMPARATIVE RISK. W. K. Sinclair in *Radiology*, Vol. 138, No. 1, pages 1–9; October, 1981.

THE FOSSIL FOOTPRINTS OF LAETOLI

FOSSIL HOMINIDS FROM THE LAETOLIL BEDS. M. D. Leakey, R. L. Hay, G. H. Curtis, R. E. Drake, M. K. Jackes and T. D. White in *Nature*, Vol. 262, No. 5568; August 5, 1976.

MELILITITE-CARBONATITE TUFFS IN THE LAETOLIL BEDS OF TANZANIA. Richard L. Hay in *Contributions to Mineralogy and Petrology, Vol. 67, No. 4, pages 357–367; December 15, 1978.*

PLIOCENE FOOTPRINTS IN THE LAETOLIL BEDS AT LAETOLI, NORTHERN TANZANIA. M. D. Leakey and R. L. Hay in *Nature*, Vol. 278, No. 5702, pages 317–323; March 22, 1979.

PALEOENVIRONMENT OF THE LAETOLIL BEDS, NORTHERN TANZANIA. R. L. Hay in *Hominid Sites: Their Geologic Settings*, edited by G. Rapp, Jr., and C. F. Vondra. Western Press, 1981.

HOW AN ANIMAL VIRUS GETS INTO AND OUT OF ITS HOST CELL

ASSEMBLY OF THE SEMLIKI FOREST VIRUS MEMBRANE GLYCOPROTEINS IN THE MEMBRANE OF THE ENDOPLASMIC RETICULUM *IN VITRO*. Henrik Garoff, Kai Simons and Bernhard Dobberstein in *Journal of Molecular Biology*, Vol. 124, No. 4, pages 587–600; October 5, 1978.

COATED PITS, COATED VESICLES, AND RECEPTOR-MEDIATED ENDOCYTOSIS. Joseph L. Goldstein, Richard G. W. Anderson and Michael S. Brown in *Nature*, Vol. 279, No. 5715, pages 679–685; June 21, 1979.

ON THE ENTRY OF SEMLIKI FOREST VIRUS INTO BHK-21. Ari Helenius, Jürgen Kartenbeck, Kai Simons and Erik Fries in *The Journal of Cell Biology*, Vol. 84, No. 2, pages 404–420; February, 1980.

THE BUDDING MECHANISMS OF ENVELOPED ANIMAL VIRUSES. Kai Simons and Henrik Garoff in *The Journal of General Virology*, Vol. 50, Part 1, pages 1–21; September, 1980.

TITAN

THE PLANET SATURN: A HISTORY OF OBSERVATION, THEORY AND DISCOVERY. A. F. O'D. Alexander. Faber and Faber, 1962.

TITAN. James B. Pollack in *The New Solar System*, edited by J. Kelly Beatty, Brian O'Leary and Andrew Chaikin. Cambridge University Press and Sky Publishing Corporation, 1981.

VOYAGER 1 ENCOUNTER WITH SATURN. *Science*, Vol. 212, No. 4491; April 10, 1981.

THE CHEMISTRY OF FLAMES

FLAME STRUCTURE. R. M. Fristrom and A. A. Westenberg. McGraw-Hill Book Company, 1965.

FLAMES: THEIR STRUCTURE, RADIATION AND TEMPERATURE. A. G. Gaydon and H. G. Wolfhard. Chapman and Hall, 1970.

FUNDAMENTALS OF COMBUSTION. Roger A. Strehlow. R. E. Krieger Publishing Co., Inc., 1979.

CHEMICAL KINETICS OF HIGH TEMPERATURE COMBUSTION. W. C. Gardiner, Jr., and D. B. Olson in *Annual Review of Physical Chemistry*, Vol. 31, pages 377–399; 1980.

QUASARS AS PROBES OF THE DISTANT AND EARLY UNIVERSE

QUASI-STELLAR OBJECTS. Geoffrey Burbidge and Margaret Burbidge. W. H. Freeman and Company, 1967.

THE CLUSTERING OF GALAXIES. Edward J. Groth, P. James E. Peebles, Michael Seldner and Raymond M. Soneira in *Scientific American*, Vol. 237, No. 5, pages 76–98; November, 1977.

QUASARS: OBSERVED PROPERTIES OF OPTICALLY SELECTED OBJECTS AT LARGE REDSHIFTS. Malcolm G. Smith in *Vistas in Astronomy*, Vol. 22, pages 321–362; 1978.

OBJECTS OF HIGH REDSHIFT. George O. Abell and P. J. E. Peebles. D. Reidel Publishing Co., 1980.

COLOR VISION IN FISHES

VISION IN FISHES. Edited by M. A. Ali. Plenum Press, 1975.

THE EVOLUTION OF PHOTOPIC VISUAL PIGMENTS IN FISHES: PART III. W. N. McFarland and F. W. Munz in *Vision Research*, Vol. 15, No. 10, pages 1071–1080; October, 1975..

THE ECOLOGY OF VISION. J. N. Lythgoe. Oxford University Press, 1979.

HENRY A. ROWLAND

THE PHYSICAL PAPERS OF HENRY AUGUSTUS ROWLAND. The Johns Hopkins Press, 1902.

SOME EMINENT ANDOVER ALUMNI: 19. HENRY AUGUSTUS ROWLAND. 1848–1901. Scott H. Paradise in *The Phillips Bulletin*, Vol. 25, No. 1, pages 24–32; October, 1930.

THE EDUCATION OF AN AMERICAN SCIENTIST: H. A. ROWLAND, 1848–1901. Samuel Rezneck in *American Journal of Physics*, Vol. 28, No. 2, pages 155–162; February, 1960.

ROWLAND AND THE NATURE OF ELECTRIC CURRENTS. John David Miller in *Isis*, Vol. 63, No. 216, pages 4–27; March, 1972.

ROWLAND'S MAGNETIC ANALOGY TO OHM'S LAW. John David Miller in *Isis*, Vol. 66, No. 232, pages 230–241; June, 1975.

THE AMATEUR SCIENTIST

THE GRANULARITY OF SCATTERED OPTICAL MASER LIGHT. J. D. Rigden and E. I. Gordon in *Proceedings of the IRE* (Institute of Radio Engineers), Vol. 50, No. 11, pages 2367–2368; November, 1962.

LASER SPECKLE AND SURFACE ROUGHNESS. W. T. Welford in *Contemporary Physics*, Vol. 21, No. 4, pages 401–412; July/August, 1980.

Introducing the engineer's personal computer.

Improving the engineer's productivity is another way HP computers help manufacturers sharpen their competitive edge.

The new HP 9836 is not just a general-purpose computer that's been fine-tuned for engineering. It's a full-function workstation developed expressly for the jobs you deal with most in the lab or on the bench: design, computation and testing.

Yet for all its power, the 9836 is more than just a "personal computer" —it's *personalized*. For over 12 years at HP, we've been building personal computers for engineers, scientists and other technical professionals. By working closely with them, we've been able to identify the kinds of features they need to help gather data, crunch numbers and interpret results. Quickly and easily.

That's why a large, easy-to-read CRT is part of the 9836. So are high-performance graphics and Pascal (in addition to BASIC and HPL). Even a special rotary control knob that can be used for editing programs, calibrating instruments or controlling motor speeds. Add to that a growing list of specialized engineering software and you'd think the 9836 was designed just for you!

At the heart of the 9836 is the powerful Motorola MC 68000 processor, which provides ultra-fast computation and data transfer. You can also take full advantage of the computer's eight I/O slots, choice of interface cards and additional memory boards (up to 2 megabytes) to personalize your system even further.

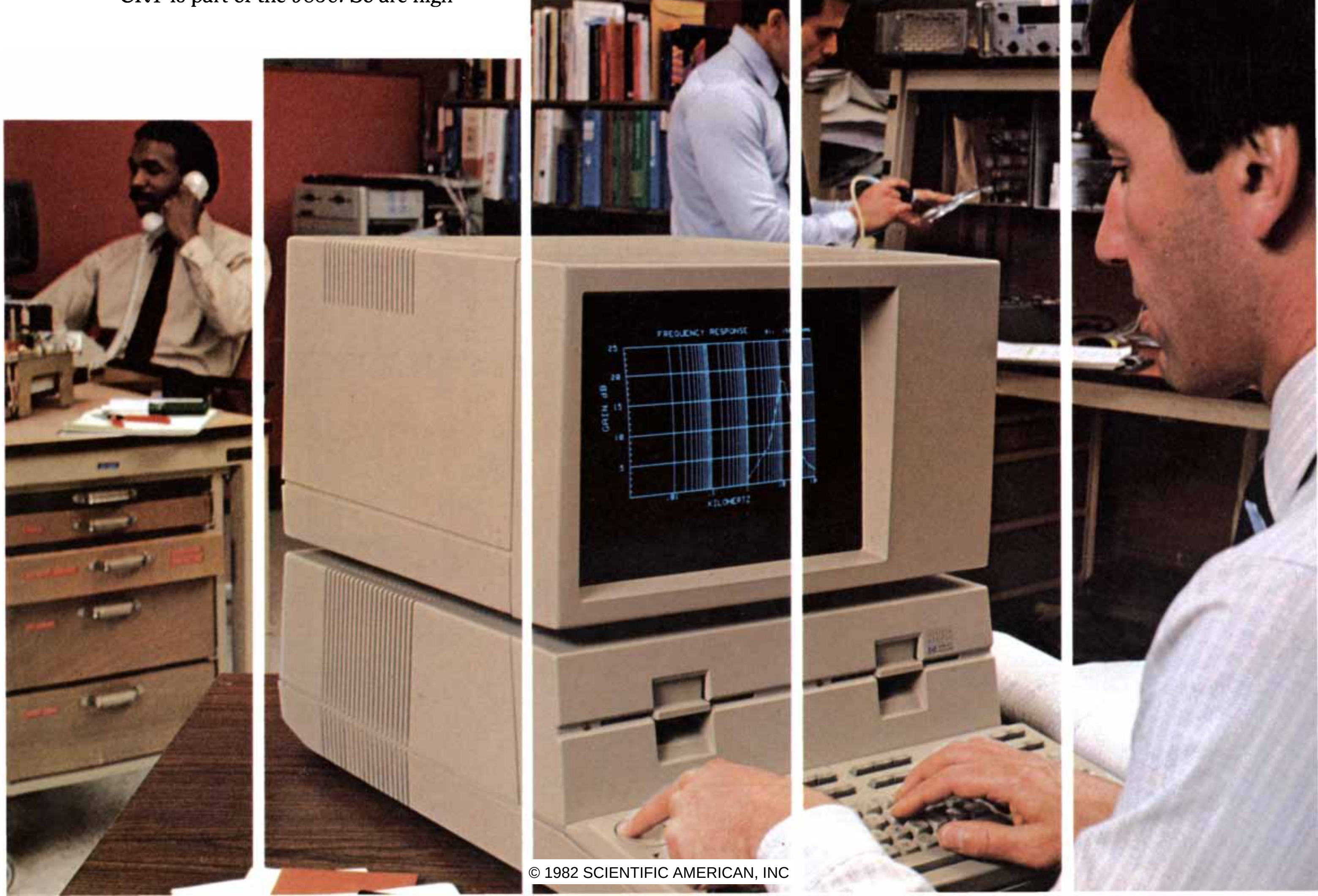
For example, you can plug in just the HP peripherals you need: printers, plotters, data tablets and high-speed discs (including the new 5″ Winchester). And with the built-in HP-IB standard instrumentation slot, you get access to the entire range of HP instruments.

Imagine the power of the 9836 on *your* desk. For a hands-on demonstration, contact your local HP sales office listed in the White Pages. Or come and see the 9836 in action at our Productivity'82 Seminars. (Watch your local newspaper for details.) You can also write Hewlett-Packard, Attn: Pat Welch, Dept. 56126, 3404 E. Harmony Rd., Ft. Collins, CO 80525.

When performance must be measured by results.

HEWLETT PACKARD

40202



A curiously tenacious that's saving millions

Hal Shaub discovered the molecule's properties at Exxon Research and Engineering Company.

Hal Shaub (Ph.D. Chemistry), Senior Research Associate in Exxon Research and Engineering Company, discovered some curious properties of a molecule in his work to develop fuel-saving motor oils. "It's a very tenacious molecule," Hal says, "sporting a pair of highly polar 'feet' that attach to positive and negative sites on metal surfaces."

Two Kinds of Friction

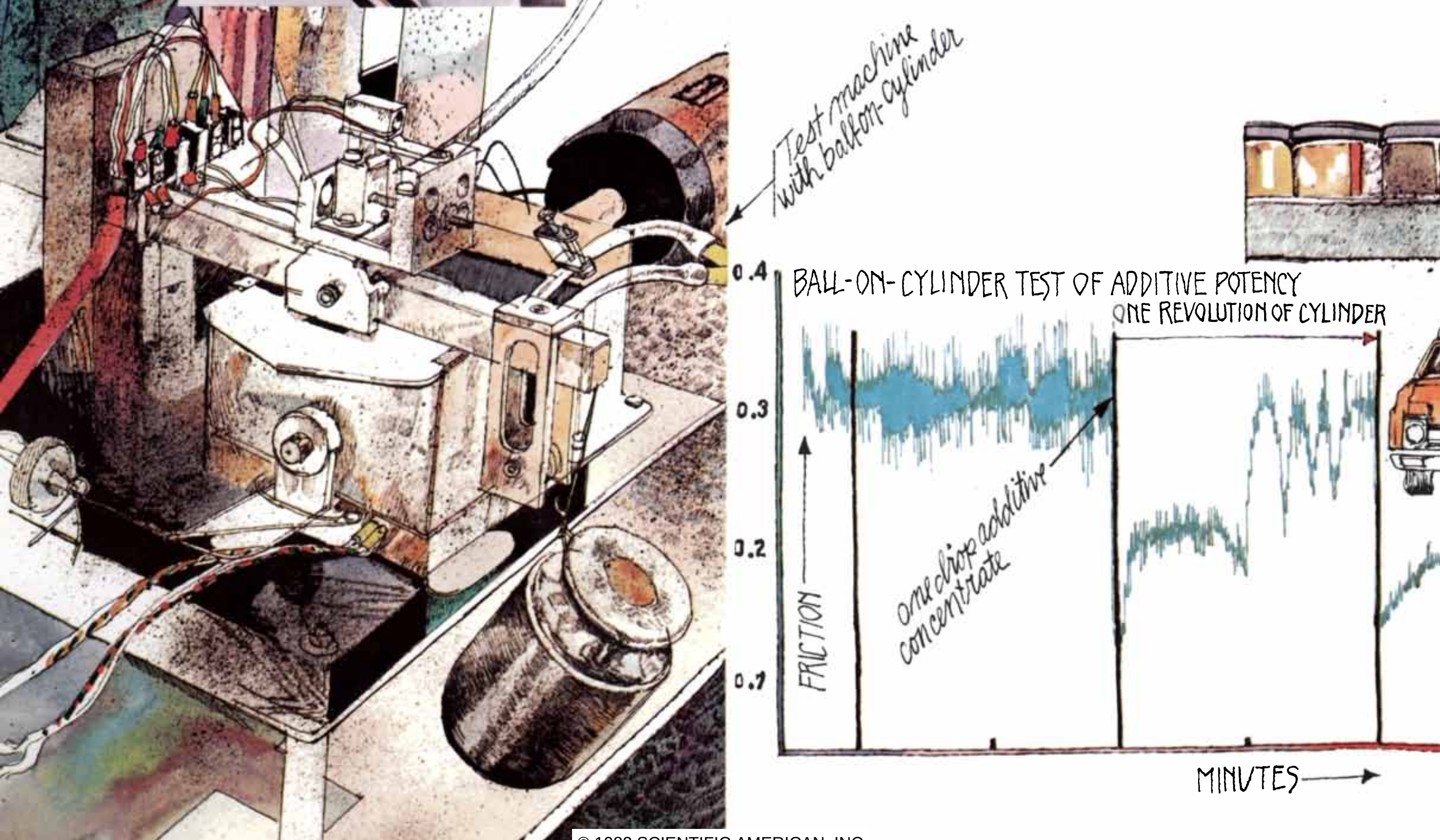
In a typical internal combustion engine, a considerable amount of fuel is consumed in overcoming friction. It has two sources: *rubbing* where lubricant film fails and metal-to-metal contact occurs, and *drag* caused by the viscosity of the lubricant itself. Friction can be reduced by lowering oil viscosity, but there is a point at which friction begins to increase due to failure of the lubricant film and resulting metallic contact. To Hal, this suggested developing sturdier lubricant films.

Hal uses a unique laboratory device to assess additives–Exxon's "ball on cylinder" test that simulates conditions in parts of the engine where lubricant films commonly give way. The test gives positive laboratory confirmation that the curious two-footed molecule reduces friction. But it has also raised puzzling questions about how and why.

Two Theories on How the Molecule Works

The additive that Hal discovered actually seems to chemisorb on steel surfaces, reducing metallic contact under thin lubricant film conditions.

One theory is that this chemisorption reduces adhesive wear, and low friction prevents sub-surface fatigue wear. So stresses exerted by the load cause plastic deformation of rubbing steel surfaces–resulting in smoother surfaces and less friction.



molecule from Exxon of gallons of fuel.

But Hal believes that when metallic parts contact and heat in the presence of the molecule, changes in the metal's chemistry occur. Specifically, the melting point may be lowered, causing surface irregularities to deform in a more uniform way–resulting in lower friction.

Hal adds that various additive chemicals can compete with the molecule for polar sites on the metal, can prevent it from getting to the surface, or both. So for the molecule to work its magic, special chemical techniques must be used to incorporate it into the motor oil.

Mileage Improvements Averaged over 4%

The complete additive technology, in a super premium motor oil, was assessed for overall engine performance in dynamometer and road tests, including a grueling taxi fleet test. Other special road fleet tests demonstrated improvements in fuel economy averaging over 4%, compared to conventional 10W-40 motor oils of the time.

The additive technology has been incorporated in Exxon's *Uniflo®* automotive motor oil since 1977, in heavy-duty oils for diesel engines since 1980, and is now in *Exxon Extra Motor Oil®*. The fuel savings resulting from consumer use of these fuel-efficient oils are estimated at millions of gallons per year.

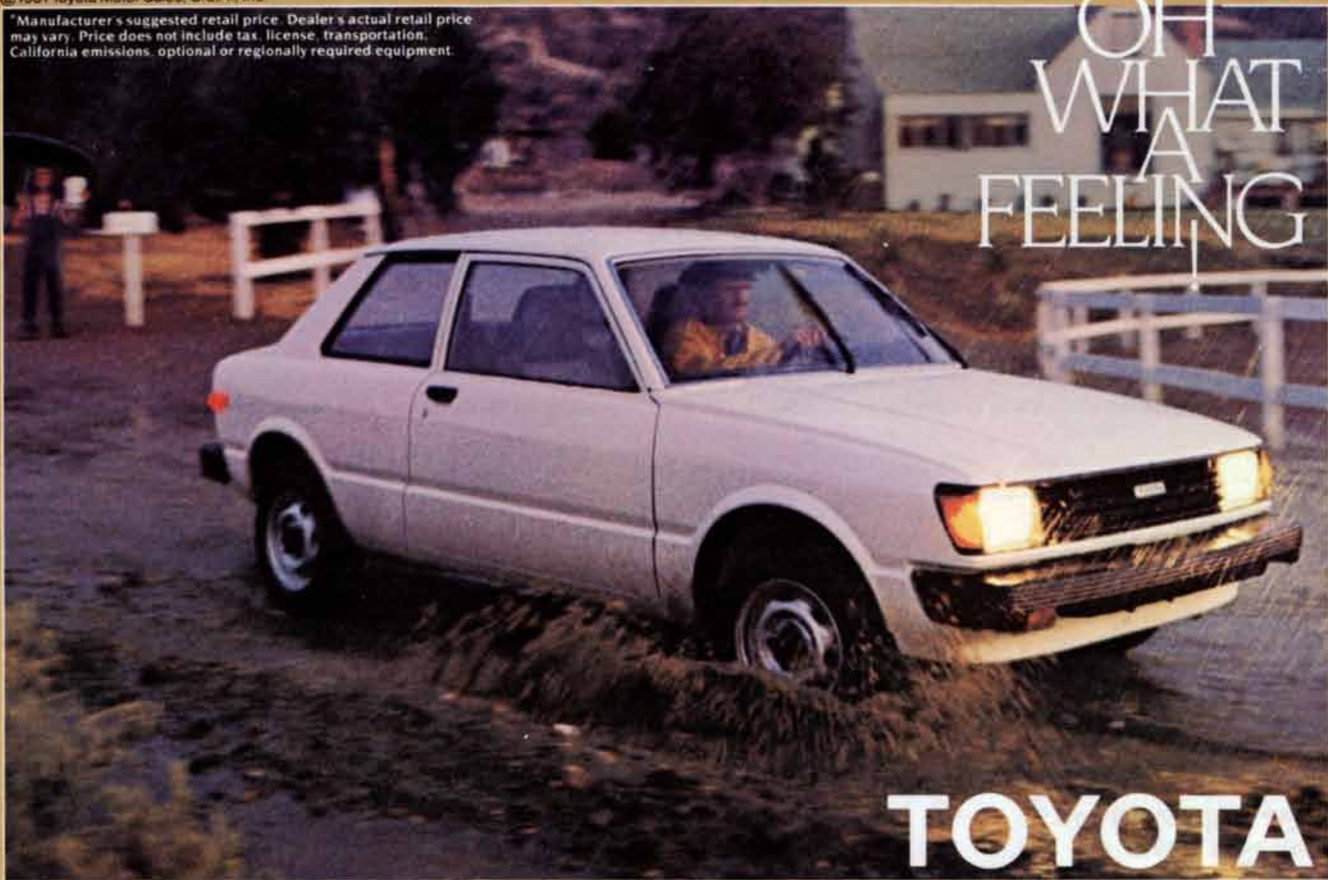
Meanwhile, Hal Shaub is continuing his work. "We think we can *double* the fuel economy improvements achieved to date," says Hal.

For more information on Hal Shaub's molecule and ER & E, write Ed David, President, Exxon Research & Engineering Company, Room 605, P.O. Box 101, Florham Park, New Jersey 07932.

Exxon Research and Engineering Company

Fuel-saving engine lubricants are just one example of technological innovation going forward on many fronts at Exxon Research and Engineering Company. A wholly owned subsidiary of Exxon Corporation, ER&E employs over 2,000 scientists and engineers working on petroleum products and processing, synthetic fuels, pioneering science, and the engineering required to develop and apply new technology in the manufacture of fuels and other products.

EXXON



THE $4998 FRONT-WHEEL-DRIVE COROLLA TERCEL. A CAR FOR ALL SEASONS.

Neither rain nor sleet nor darkness of the economy will stay this Toyota from the swift completion of its appointed rounds.

The 1982 Toyota Corolla Tercel 2-Door Sedan pulls you through with front-wheel-drive in any season. And a low Corolla price. Just $4998.*

Corolla Tercel gas mileage makes the going a little easier too. The Corolla Tercel is rated at 48 EPA Estimated Highway MPG, (36) EPA Estimated MPG. Remember: Compare this estimate to the EPA "Estimated MPG" of other cars with manual transmission. You may get different mileage, depending on how fast you drive, weather conditions, and trip length. Actual highway mileage will probably be less than the EPA "Highway Estimate."

The Corolla Tercel. Also available as a 3-Door Liftback. Or a practical 4-Door Sedan with automatic transmission. All built with Toyota's total economy to keep on going season after season.

The 1982 Toyota Corolla Tercel. The front-wheel-drive that really delivers!