SCIENTIFIC AMERICAN

RICE

$2.50

January 1984


Nissan is proud to sponsor
ABC-TV's telecast of the
1984 Summer Olympic Games
© 1981 ABC Inc.
"Star in Motion"
© 1980 L. A. Olympic Committee
COME ALIVE, COME AND DRIVE
THE FIRST
If we could give you the keys and an open road... this magnificent machine would capture your soul forever. It's a whole new world of driving where art and technology intersect. It's the Nissan 300 ZX. And it is major motion.
55 MPH
SPEED
TACH
Z

Test Drive the Leica R4.

We invite you to experience the perfection of the Leica® R4. We invite you to hold, inspect, and then test the Leica that has been acclaimed "...a miracle of modern technology." A limited production SLR of astonishing versatility, precision and beauty. With flawless Leitz® optics, two exposure measurement systems, five operation modes, and a series of advances unequalled in the history of photography.

The Leica R4 is Leica's proudest achievement. A photographic instrument designed for perfection, not price. Like every Leica since the beginning, it is lavished with attention to detail. Hand finishing. Hand assembly. Hand testing. No shortcuts or compromises. Because the Leica R4 is designed and built never to fail. And to have no limit to its useful life.

We invite you to experience the Leica R4. And through it enter Leica's world of superior photography. To see and appreciate for yourself the perfection of the photographic image that is uniquely Leica's.

You can test drive the Leica R4 now, at your local Leica dealer. And to show his appreciation for your visit and your interest, your Leica dealer will enter your name in a special drawing for the ultimate sports coupe—the BMW 633CSi.* We think you'll find this high performance German import, the perfect accessory to your new Leica.

For the name of your nearest Leica dealer write E. Leitz, Inc., 24 Link Drive, Rockleigh, N.J. 07647 or call (800) 223-0507, ext. 801. In New Jersey call (800) 223-0509, ext. 801.

And you could win a BMW 633 CSi to drive it home in.

*No purchase necessary. Entries must be made through your Leica dealer before January 31, 1984. The winner will be notified by March 30, 1984. Void where prohibited by law.

Established 1845 **SCIENTIFIC AMERICAN** January 1984 Volume 250 Number 1

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: $24 PER YEAR IN THE U.S. AND ITS POSSESSIONS; $33 PER YEAR IN ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.

"What do they mean our training program is outdated? We all went through it!"

Back in Calvin Coolidge's day, American business moved along at a nice comfortable pace.

Life was simple. The future was solid. Even if you were new on the job, you could get the hang of it just from the lectures and workbooks in the company training program.

Of course, all that was yesterday. The tougher the challenges we're up against today, the better we'll have to train our people to meet them.

And now, thanks to a company called Wicat, the technology of training has finally come into its own.

HOW TO CUT TRAINING TIME AND COST IN HALF.

Simple. Use the power of computers to provide intense, individualized study for each trainee.

Some of the world's largest corporations put a Wicat learning station in front of each trainee. And train them better, in less than half the time. And at less than half the cost.

DEVELOP COURSES OVERNIGHT WITHOUT PROGRAMMING.

Until now, computer-based training systems demanded sophisticated programming skills. Wicat doesn't.

For example, it took a skilled programmer six weeks to develop one course on PLATO™ (the other well-known training system). To develop the same course on the Wicat system, it took a non-programmer less than two days.

LASERDISCS VERSUS LECTURES.

There's nothing like realism to make training stick. Which is why Wicat was the first to incorporate the extraordinary graphics, sound, and simulation capability of the laserdisc into computer-based training.

THE TOTAL TRAINING SOLUTION.

At Wicat, we design and manufacture our own hardware. In fact, we spent four years in development, to make this the most productive training machine possible.

Our WISE authoring system sets new standards for ease of course development. And our large staff is ready to design custom courseware for any applications.

We even offer software for program management; everything from scheduling to progress reporting to testing.

Best of all, you can get started for under $1,000 a month.

Call us at (801) 224-6400 for details. Or write to Wicat Systems, P.O. Box 539, 1875 South State Street, Orem, UT 84057.

We'll show you how up-to-date a training system can be. When it's designed by a company that really means business.

WICATsystems™

PLATO is a trademark of Control Data Corp

SCIENTIFIC AMERICAN

CORRESPONDENCE

Offprints of more than 1,000 selected articles from earlier issues of this magazine, listed in an annual catalogue, are available at $1.25 each. Correspondence, orders and requests for the catalogue should be addressed to W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, UT 84121. Offprints adopted for classroom use may be ordered direct or through a college bookstore. Sets of 10 or more Offprints are collated by the publisher and are delivered as sets to bookstores.

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of SCIENTIFIC AMERICAN for the flat fee of $1.25 per copy of each article or any part thereof. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 21 Congress Street, Salem, MA 01970. Specify CCC Reference Number ISSN 0036–8733/84. $1.25 + 0.00.

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

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

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 5969, New York, NY 10017. The date of the last issue on your subscription is shown in the upper right-hand corner of each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (if convenient, on a mailing label of a recent issue) as well as the new one.

Name ______________________

New Address

Street ______________________

City ______________________

State and ZIP ______________________

Old Address

Street ______________________

City ______________________

State and ZIP ______________________

THE COVER

The painting on the cover shows an Asian field worker planting rice in a paddy that is covered with about 10 centimeters of water. One of the distinctive characteristics of rice is that it can grow in waterlogged soil because of an air-passage system that moves oxygen from the leaves to the roots. Indeed, some species of rice can elongate by as much as 25 centimeters per day in flooded areas and can grow in water as much as five meters deep (see "Rice," by M. S. Swaminathan, page 80). The rice genus is *Oryza,* and the Asian species, which has a large number of varieties, is *O. sativa*. Some two billion people rely on rice for a substantial part of their diet, and as a result the crop takes up some 145 million hectares (358 million acres) of land. Even so, the yield would have been outstripped by the demand had it not been for the development in recent years of high-yield varieties of rice resistant to diseases and insect pests.

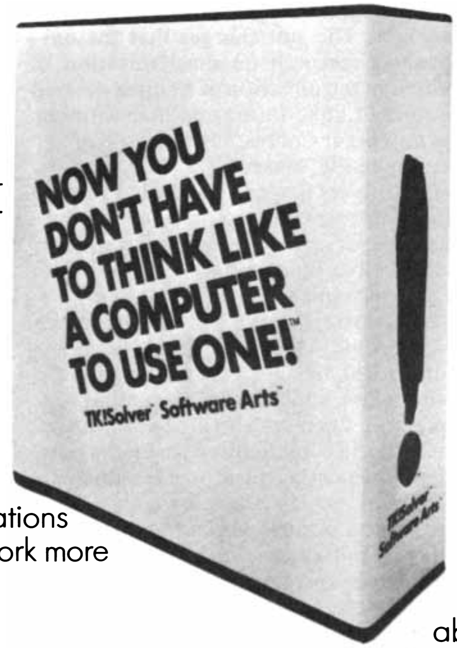
THE ILLUSTRATIONS

Cover painting by Teresa Fasolino

Page	Source
39–47	Jerome Kuhl
49	René Canul, Comisión Federal de Electricidad (*top*); Brian R. Wolff (*bottom*)
50	Brian R. Wolff (*top*), Todd Pink (*bottom*)
51	Todd Pink
52	Michael Matson, National Oceanic and Atmospheric Administration (*top*); Arlin J. Krueger, Goddard Space Flight Center (*bottom*)
53	Todd Pink
54	Alan Robock, University of Maryland, Michael Matson, National Oceanic and Atmospheric Administration, and Todd Pink
56	Todd Pink
59–60	Lawrence Berkeley Laboratory
61–63	Walken Graphics
64	Lawrence Berkeley Laboratory
65–66	Walken Graphics
80	Jurgen Schmitt, The Image Bank
82	Patricia J. Wynne
83–84	Andrew Tomko
85	Charles Moore, Black Star (*top*); Andrew Holbrooke, Black Star (*bottom*)
90	Andrew Tomko
91	Patricia J. Wynne
92–93	Andrew Tomko
95–100	Patricia J. Wynne
103–109	Ian Worpole
110	Michael Beer, Johns Hopkins University, Theo Koller, Swiss Federal Institute of Technology, and Masayasu Nomura, University of Wisconsin at Madison
111–112	Ian Worpole
116–120	Gabor Kiss
121–122	N. J. A. Sloane, Bell Laboratories, and Gabor Kiss
123	John Horton Conway, University of Cambridge, N. J. A. Sloane, Bell Laboratories, and Gabor Kiss
124	N. J. A. Sloane, Bell Laboratories, and Gabor Kiss
125	John Leech, University of Stirling, and Gabor Kiss
127–129	The Granger Collection
130	Colonel Richard Gimbel Aeronautics History Collection, U.S. Air Force Academy Library
131–136	The Granger Collection
139	Jearl Walker (*top and bottom*), David A. Katz (*middle*)
140–143	Michael Goodman

TK!SOLVER™

THE WORLD'S FIRST EQUATION PROCESSOR FOR PERSONAL COMPUTERS™



The TK!Solver™ program is a professional tool that allows you to work easily with equations. Whether your problem is a simple formula or a model consisting of many equations, TK!Solver can help improve your productivity. Once the equations are written, enter the known values, press the ! key, and TK!Solver gives you the answer.

Engineers, scientists, architects, financial analysts and planners, educators, researchers, and other professionals who use equations and mathematical models can work more creatively with TK!Solver.

TK!Solver provides you with:

BACKSOLVING

If the programs you use now require you to rewrite the same equation to solve for different unknowns, TK!Solver can dramatically improve your productivity. Enter your problem once and then solve for the unknowns no matter where they are in your equation.

ITERATIVE SOLVING

If TK!Solver can't solve an equation directly, take an educated guess at the answer. Type the ! key and the TK!Solver program starts with your guess and performs repeated approximations to converge on the answer.

LIST SOLVING

Given a list of input values, TK!Solver automatically calculates the equation for every value in your list. For example, if you want to know how different interest rates will affect monthly loan payments, enter a list of interest rates and let TK!Solver calculate the payment amount for each value.

UNIT CONVERSIONS

Any type of unit conversion – Fahrenheit to Celsius, meters to feet, dollars to deutschmarks, newtons to dynes – can be made without altering your equations. Just define the numerical relationship between two units of measurement and the TK!Solver program automatically converts the variable value to the unit you specify.

TABLES AND PLOTS

Quickly generate tables and plots of your results on your screen or printer.

If you can define a problem mathematically, the TK!Solver program can help you find the answer. It's a flexible problem-solving tool that you can apply easily to a variety of mathematical problems.

AVAILABLE NOW

You can run the TK!Solver program on the IBM® PC and XT and compatible machines, the Digital™ Professional™ 350, the Digital™ Rainbow™ 100, the Wang Professional Computer, Apple® //e, and on the following personal computers using MS™-DOS: TI Professional Computer, GRiD Compass Computer,™ Canon AS-100, Eagle® 1600, Toshiba T300, and the Zenith Z-100.™

Software Arts™
The creators of VisiCalc® and TK!Solver™

LETTERS

Sirs:

In his article "Labor-intensive Agriculture" [SCIENTIFIC AMERICAN, October, 1983] Philip L. Martin of the University of California at Davis gives one pause on the question of scientific objectivity. He writes: "The University of California is defending itself in a lawsuit brought by representatives of farm workers. The suit charges that the university's research on mechanization is biased in the direction of helping the operators of large farms and undermining farm-worker unions." Is this type of research really biased? Dr. Martin goes on: "Universities can head off some of the controversy by making clear the importance of mechanization in preserving a profitable fruit and vegetable industry." Mechanization, therefore, will help growers to preserve (if not increase) their profits. Mechanization is to be accompanied by the restructuring of the production of commodities "that depend on alien workers" to eliminate these workers (and thus weaken the base of many existing farm-worker unions). This will benefit the growers. Should they invest in such research? This possibility is not even suggested. The Government is to create a payroll tax on the farm worker's wages (not the grower's profits) to set up a special trust fund that "would levy a tax on the wages earned by alien farm workers." Dr. Martin thinks $150 million a year could be raised in this way—quite a lot for university researchers to dip into. The research will lead to a reduction in jobs for undocumented workers without reducing in any way the social causes that drive millions of people to come to this country in search of these jobs. These proposals can only help the growers and their university hangers-on while leaving unaddressed the real human problems of many millions of working people, "alien" or "nonalien."

THOMAS RIGGINS

New York, N.Y.

Sirs:

In "Labor-intensive Agriculture" Philip L. Martin writes: "Without mechanization [in fruit and vegetable agriculture] the U.S. must both accept an isolated, alien-dominated labor force for seasonal handwork and erect trade barriers to keep out produce grown abroad at even lower wages." But if produce grown abroad is cheaper (if foreign farm workers need the money more than U.S. and Mexican ones), the U.S. will not need a domestic product and thus will *need* neither trade barriers nor its own seasonal work force.

Martin's isolationist presupposition that the U.S. must be a world to itself in both ethnos and trade, together with his misuse of "American" to exclude Mexicans, shows clearly that the Protestant Republic is still alive, even among scientists. Pity.

JOHN A. WILLS

Pasadena, Calif.

Sirs:

Both writers miss my point: there is an inevitable shakeout coming in the labor-intensive fruit and vegetable industry. The U.S. is most likely to keep producing crops that can be handled mechanically, since our edge is high technology, not low wages. Public policies that accelerate this inevitable shakeout will strengthen our agricultural economy, create jobs abroad and reduce pressures to migrate illegally.

Mr. Riggins misunderstands the fact that, with or without alien workers, labor-intensive agriculture in the U.S. will be mechanized because countries with lower wages are using American-developed technologies to expand their production of fruits and vegetables for distant markets. American farmers will continue to produce crops that can be mechanically harvested; hand-harvested crops might migrate abroad, creating jobs and reducing incentives to migrate illegally to the U.S. The trust fund is one way to help equalize the costs of the legal alien workers expected by immigration reform during this transition (otherwise alien workers are cheaper than American workers). The trust fund could also generate money to upgrade farm jobs and retrain displaced farm workers.

Mr. Wills's letter includes a non sequitur. American farmers today produce *and* export labor-intensive crops. Foreign producers are taking away traditional export markets and will eventually export cheaper citrus fruits, tomatoes and strawberries to the U.S. At that time—sooner for tomatoes, later for strawberries—farmers will argue that they cannot pay higher wages because of low-wage foreign competition or that the U.S. must exclude the products grown with cheap labor abroad. In either case the U.S. faces an unpalatable choice: an isolated farm labor market *or* protectionist trade policies that keep out foreign produce and encourage illegal immigration.

PHILIP L. MARTIN

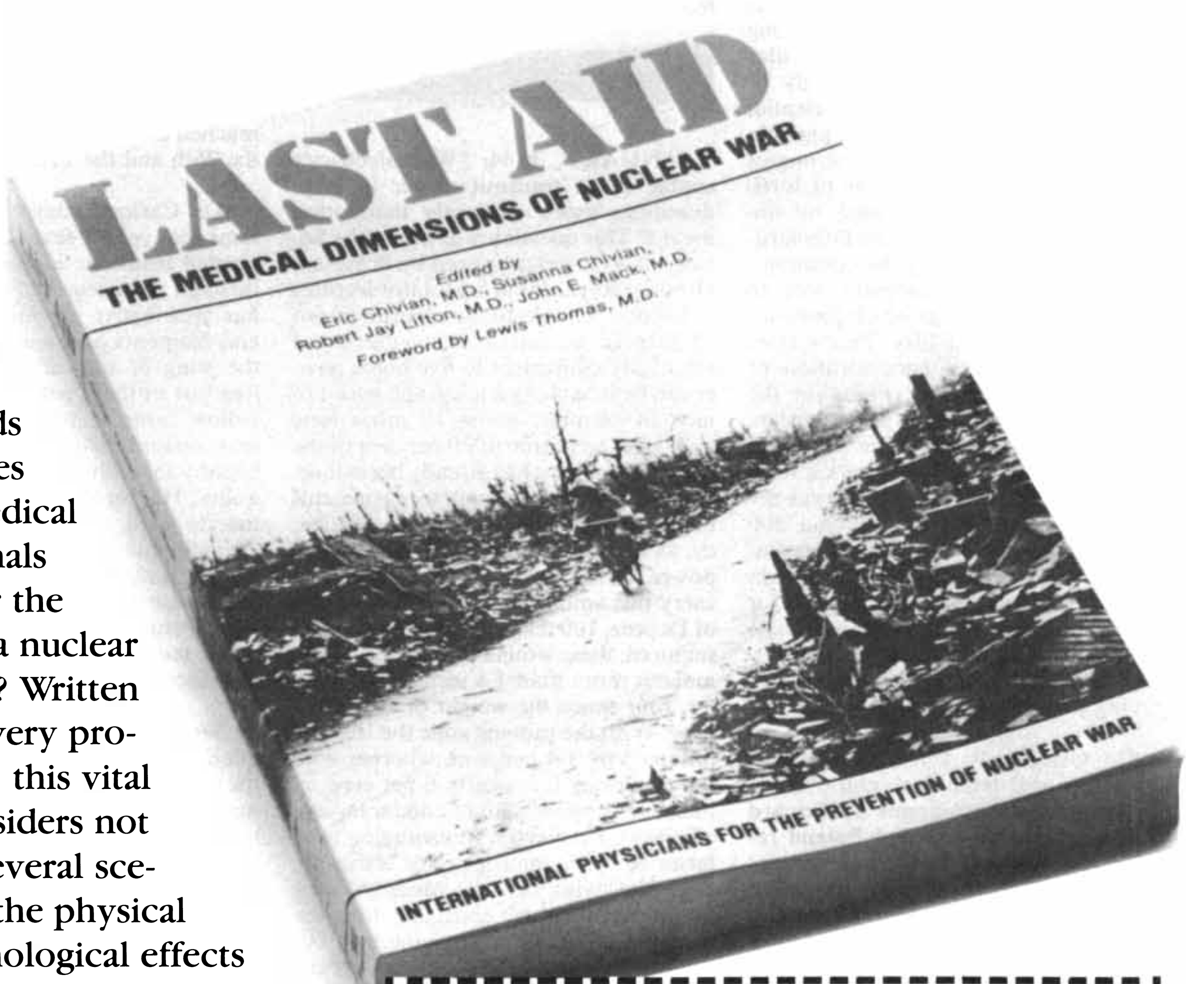
University of California
Davis

"To ignore this passionate presentation is to turn our backs on the most urgent issue of our time."

—Walter Cronkite

*"***Last Aid** *is a no-nonsense description of the horrifying medical consequences of nuclear war from the leading experts in the field. This invaluable and timely work should be read by everyone who cares about the future of humanity."*

—Coretta Scott King



What kinds of tragedies would medical professionals encounter the day after a nuclear holocaust? Written by those very professionals, this vital book considers not one but several scenarios of the physical and psychological effects of nuclear war on the global population. A reaction to government leaders who speak rationally of a conflict than can have no rational basis, *Last Aid* offers a compelling assessment of the most critical issue of our time.

Yes, please send me a copy of *Last Aid*.

☐ I enclose a check for $10.95 (paper), plus $1.50 to cover postage and handling. (California, New York, & Utah residents add appropriate sales tax.)

Charge my ☐ MasterCard ☐ VISA

Account # ______________________ Expires __________

Signature ______________________________

(All credit card orders must be signed.)

Name ______________________________

Address ______________________________

City ______________ State __________ Zip __________

W. H. Freeman and Company
4419 West 1980 South
Salt Lake City, Utah 84104

ISBN paper 0-7167-1435-3

0117

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JANUARY, 1934: "In recent years France, taking example by the lessons learned in the World War, has created an extensive system of fortifications along her frontiers. Some of the projects built for this purpose consist of strong forts while others are most spectacular in that they are composed chiefly of subterranean structures. The essential points of the system, built on a gigantic scale and thought to be the strongest ever evolved, consist of a line of fortified casemates interconnected by underground galleries safe from bombardment. At intervals along the communicating galleries the casemates rise to the surface on shafts in which there are electrically operated lifts. This system will allow a sudden concentration of troops in an emergency. Plans for the 'Maginot line,' so called after the late Minister of War, M. André Maginot, who was in charge of the works, were first submitted in 1925, when it was decided to make a stretch of about 200 kilometers as far as possible impregnable. It is believed by those who maintain the thesis that adequate preparation for war operations is the best insurance against the outbreak of war that this system of defense will go far toward preserving peace in Europe."

"On October 9 American observatories received from the central office for astronomical telegrams at Harvard the message, 'Belgium and Poland report great meteoric shower in progress.' Wherever the skies were clear they were assiduously watched, but before the news had reached our side of the world the display was over. All across Europe, from Russia to Spain, a magnificent shower was observed. Gerasimovic at Poulkovo reports a rate of more than 100 meteors per minute, Witkowski at Poznań a maximum rate of 10 per second, some with brightness of the zero magnitude—equal to Vega or Arcturus. This is the finest meteoric display since the great fall of Leonids in 1866. Meteor swarms of this kind are always associated with comets, and the body responsible in this case was at once identified as the periodic comet discovered by Giacobini in 1900 and independently at a later return by Zinner."

"A process by which files of newspapers may be preserved for posterity on photographic safety film was described recently by Charles C. Case of the Eastman Kodak Company. The process is made possible by the development of a camera that can photograph more than eight full-size newspaper pages on a strip of film $1\frac{3}{8}$ inches by 12 inches and a month of 50-page papers on a single reel less than four inches in diameter. The film can be read in newspaper offices on a simple viewing device that will enlarge the tiny page images from the film up to half again the size of the original newspaper page. The new miniature-image process is also useful for public libraries, which store large quantities of newspapers."

SCIENTIFIC AMERICAN

JANUARY, 1884: "Will electricity enable us to transmit power in large quantities more efficiently than other means? This question was put to the Society of Arts and answered by Professor Osborne Reynolds in his Cantor lectures as follows. Thanks to the experiments of M. Deprez, we can say that a current of electricity equivalent to five horse power can be sent along a telegraph wire 1/6 inch in diameter some 10 miles long with an expenditure of 29 per cent of the power, because it has already been done. Compared with a moving wire rope, this means it falls short in actual efficiency, as the Messrs. Hems send 500 horse power with a 3/4-inch flying rope. To carry this amount, as in the experiment of Deprez, 100 telegraph wires would be required; these wound into a rope would make it more than 1.4 inches in diameter, four times the weight of the Hems rope. With the moving rope the loss per mile is only 1.4 per cent, whereas with the electricity it is nearly 6 per cent; so that, as regards weight of conductor and efficiency, the electric transmission is inferior to the flying rope. Nor is this all. With the flying belt the Messrs. Hems found the loss at the ends, in getting the power into and out of the rope, was $2\frac{1}{2}$ per cent, whereas in M. Deprez's experiment 30 per cent was lost in the electric machinery alone, which is very small as such machinery goes. No account is here taken of the loss of power in transmission to and from the electric machinery. Taking the whole result, it does not appear that more than 15 or 20 per cent of the work done by the steam engine could have been applied to any mechanical operation at the other end of the line, as against 90 per cent that might have been realized with wire rope transmission."

"M. Ferdinand de Lesseps has communicated an interesting note on the tidal waves of earthquakes to the French Academy of Sciences. On the 27th of August last, after 4:00 P.M., the sea level at Colón, on the Atlantic side of the Isthmus of Panama, began to oscillate, as was shown distinctly by the marigraph of the Inter-Oceanic Canal Company. In amplitude the oscillations equaled the usual tidal rise but succeeded one another at intervals of $1\frac{1}{2}$ hours instead of 12 hours, as in the case of the normal tides. The curve of the marigraph showed that between 3:30 P.M. and 1:30 A.M. the sea oscillated eight times with a rise of from 0.30 to 0.40 meter. The movement began with an ebb of the water, as if a hollow had been in the sea, and gradually diminished after 1:30 A.M. on the 28th, till 11:00 P.M. on the succeeding night. M. de Lesseps can connect this phenomenon only with the volcanic eruptions in the Straits of Sunda, near Java, which began on August 25th and reached their height during the nights of the 26th and the 27th of August."

"Dr. Carlos Finlay of Havana maintains that yellow fever may be communicated from one individual to another through the agency of mosquitoes. He has seen under the microscope spores and filaments of a particular nature on the sting of one of these insects that had just bitten a patient suffering from yellow fever, and he thinks the germs may undoubtedly be introduced into a healthy individual by the bite of a mosquito. He recalls the fact that these insects were remarkably numerous in Philadelphia at the time of the great yellow fever epidemic in 1797, and he states also that the same conditions of temperature are necessary for the life of the mosquito as are necessary for the existence and spread of yellow fever."

"We are pleased to find that increased attention is being paid to the question of the physical training of young and growing girls. The Swedish physical exercises have found general favor, while many games and athletic pursuits are now allowed which formerly were proscribed by prudish schoolmistresses and timid mamas. When the girl is naturally healthy, little is wanted but to encourage ordinary systematic exercise being taken daily. This should consist of certain gymnastic exercises, which are to be practiced each day as part of the school work, supplemented by such games as lawn tennis, rounders, golf, etc. Swimming is an exercise that every girl should indulge in, and it ought to be taught systematically at all our girls' schools. Until recently dress proved a great barrier to free exercise of the limbs and body, but the introduction of a more sensible costume for the playground will in the future, it is to be hoped, remove the disadvantage. The costume in use consists of a short skirt of blue serge, draped with a crimson scarf, blue jersey, short trousers and long stockings."

Product of France. Made with fine cognac brandy 80 proof. ©1983 Carillon Importers, Ltd., N.Y.
Grand Marnier Liqueur
Triple Orange
Paris France
Produce of France
Beyond the best is the only place to be.
Grand Marnier.

THE AUTHORS

JOHN STEINBRUNER ("Launch under Attack") is director of foreign-policy studies at the Brookings Institution. He got a B.A. at Stanford University and went on to earn his Ph.D. in political science from the Massachusetts Institute of Technology. He taught briefly at M.I.T. before moving to the Institute of Politics at Harvard University. There he became associate professor of public policy and assistant director of the Program for Science and International Affairs before joining the faculty at Yale University as associate professor of political science. He left Yale in 1978 to take up his current job. Steinbruner is author of *The Cybernetic Theory of Decision: New Dimensions of Political Analysis* (Princeton University Press, 1974).

MICHAEL R. RAMPINO and STEPHEN SELF ("The Atmospheric Effects of El Chichón") are geologists who are collaborating in a study of volcanic eruptions and their effect on world climate. Rampino writes: "I received a B.A. from Hunter College and went on to earn a Ph.D. in geological sciences from Columbia University. From 1978 to 1980 I held a National Academy of Sciences research associateship at the Goddard Institute for Space Studies of the National Aeronautics and Space Administration. Since 1980 I have been a research associate in the department of geological sciences at Columbia working at the Goddard Institute and have also been visiting assistant professor at Dartmouth College. At Goddard I became involved in an effort to make a model of global climate and specifically to predict the effects of explosive volcanic eruptions on the atmosphere." Self was born in England and got a B.S. at the University of Leeds. He continued his study of geology at the Imperial College of Science and Technology, which awarded him his Ph.D. in the subject. He did field work in New Zealand and served on the faculty at Arizona State University before taking up his current job in the department of geology at the University of Texas at Arlington.

WM. C. McHARRIS and JOHN O. RASMUSSEN ("High-Energy Collisions between Atomic Nuclei") are nuclear chemists who share an interest in the subject of their article. McHarris was educated at Oberlin College and the University of California at Berkeley; his Ph.D. in nuclear chemistry was conferred by Berkeley in 1965. After obtaining his doctorate he joined the faculty at Michigan State University, where he is currently professor of chemistry and physics. Rasmussen got a B.S. from the California Institute of Technology before earning his Ph.D. in chemistry at Berkeley. He joined the faculty at Berkeley in 1952 and has been there since with the exception of the period from 1969 to 1973, which he spent as professor of chemistry at Yale University. He has also served as visiting professor at the Nobel Institute of Physics in Stockholm and as National Science Foundation senior fellow at the Niels Bohr Institute in Copenhagen.

M. S. SWAMINATHAN ("Rice") is director general of the International Rice Research Institute in the Philippines. His career has combined work in genetics with the application of scientific results to agricultural practice. Born in Kumbakonum, a temple town in southern India, he received a B.S. from the Agricultural College at Coimbatore before beginning his studies in genetics at the Indian Agricultural Research Institute (IARI). He did graduate work in the Netherlands and at the University of Cambridge; his Ph.D. was awarded by Cambridge in 1952. He then moved to the U.S. to take a job at the University of Wisconsin. In 1954 he returned to India and the IARI. In 1972 he was appointed director general of the Indian Council of Agricultural Research, a post he held until 1979, when he became Secretary of Agriculture and Rural Development in the Indian government. He was appointed director of the International Rice Research Institute in April, 1982.

CHARLES J. COLE ("Unisexual Lizards") is curator in the department of herpetology at the American Museum of Natural History in New York. He got a B.A. at Wesleyan University and an M.A. at the University of Kansas before obtaining his Ph.D. at the University of Arizona. After earning his doctorate he joined the staff of the American Museum. In the work on unisexual lizards described in his article he was assisted by his son, Jeffrey A. Cole, a student at Louisiana State University at Baton Rouge. The work is supported by grants from the National Geographic Society and the National Science Foundation.

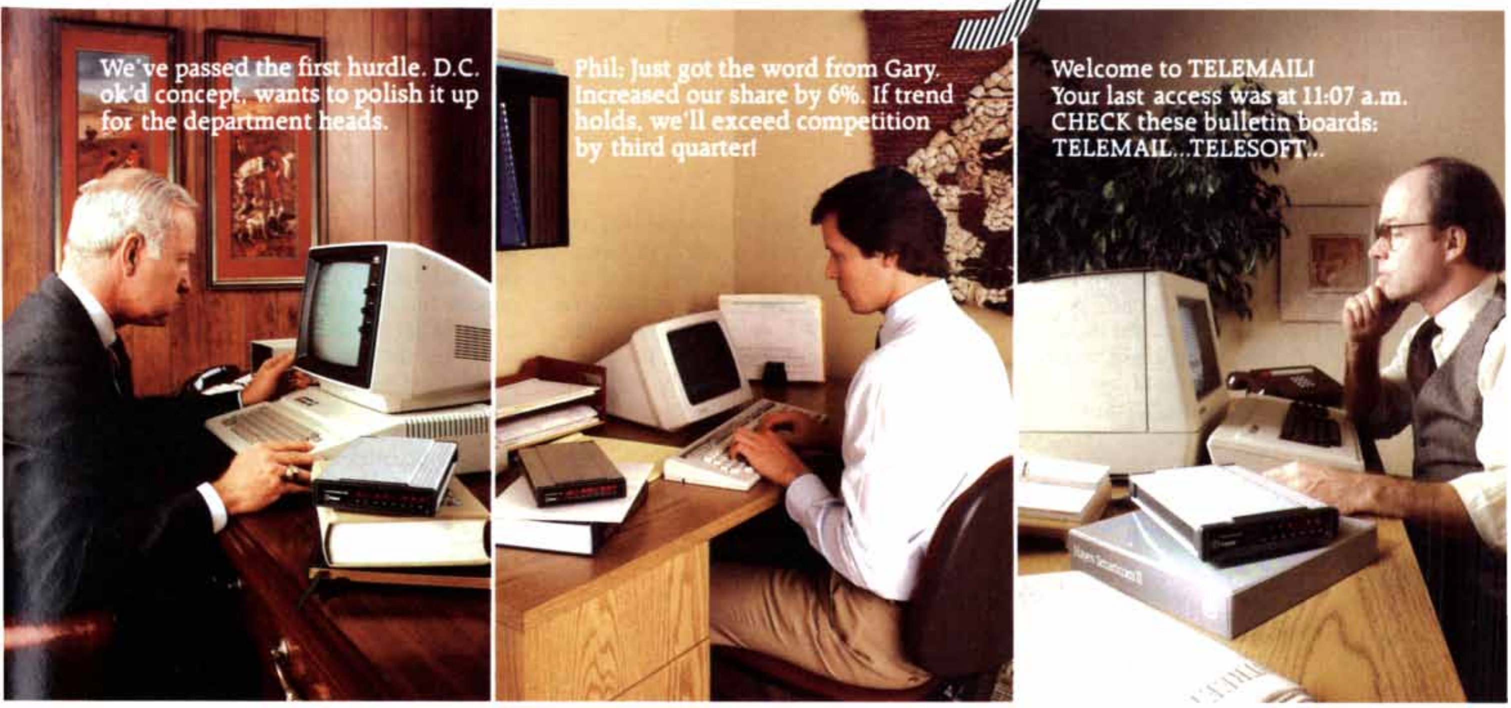
MASAYASU NOMURA ("The Control of Ribosome Synthesis") writes: "I received my college education during the miserable post–World War II period in Japan. I majored in microbiology and fermentation chemistry in a department of applied science at the University of Tokyo. I selected the department to get a job in industry and not to become a scientist. After I started to take applied-science courses, however, I realized that I had more interest in basic science than in its applications. Around that time we were just starting to hear news of the development of molecular biology. The news excited me, and in 1957 I came to the U.S. as a postdoctoral fellow. I spent the next three years working in the laboratories of Sol Spiegelman at the University of Illinois, James D. Watson at Harvard University and Seymour Benzer at Purdue University. I returned to Japan in 1960 to become assistant professor at the Institute of Protein Research at Osaka University. After three years I made the difficult personal decision to leave my own country and commit myself to doing science in the U.S. I took up a job in the department of genetics at the University of Wisconsin at Madison. In 1970 I moved to the Institute for Enzyme Research at Wisconsin, with appointments in the department of genetics and the department of biochemistry."

N. J. A. SLOANE ("The Packing of Spheres") is a member of the technical staff at Bell Laboratories. A native of Wales, he received his education in Australia and the U.S.; his B.A. is from the University of Melbourne and his M.S. and Ph.D. in electrical engineering are from Cornell University. He became assistant professor of electrical engineering at Cornell in 1967 and stayed there until moving to Bell Laboratories two years later. Sloane's scientific work has concerned coding theory, communication theory and combinatorial mathematics. Outside the laboratory his main interest is rock climbing.

ARTHUR F. SCOTT ("The Invention of the Balloon and the Birth of Modern Chemistry") was emeritus professor of chemistry at Reed College at the time of his death in 1982. He was graduated from Colby College with a B.A. in 1919. His Ph.D. in chemistry was awarded by Harvard University in 1924. He joined the Reed faculty in 1923, moving to Rice Institute three years later. He returned to Reed in 1937, ultimately becoming chairman of the chemistry department, a job he held until 1965. During World War II he served as acting president of the college. His initial scientific interest was in the determination of the atomic weights of elements. In the 1950's he became interested in the chemistry of the radioactive elements and helped to set up the first radiochemistry laboratory in a Portland hospital. From 1962 to 1964 he was director of special projects in science education for the National Science Foundation. Toward the end of his life his interests turned to the history of chemistry. At the time of his death he had written a manuscript on the subject of the present article. The manuscript was prepared for publication by Joel Keizer of the University of California at Davis, a former student of Scott's.

Hayes™

Your computer's telephone.



Whether they're collaborating on competitive strategies or accessing data bases for valuable information, more and more corporate managers are communicating with each other. With networks, data bases and mail services all over the country. And for thousands of personal computer owners, we are providing the communications link.

Hayes Smartmodem. Think of it as your computer's telephone. Just plug Smartmodem into any phone jack, and watch it send messages to and from your computer. Over ordinary phone lines, and without using the telephone receiver.

Operating with rotary, Touch-Tone® and key-set systems, Smartmodem will dial, answer, and disconnect calls. Automatically.

Choose your speed; choose your price. The lower-priced Smartmodem 300™ is ideal for inter-office communications and local data swaps, transmitting at 300 bits per second. For longer distances and larger volumes, Smartmodem 1200™ communicates at speeds up to four times faster.

Both are compatible with many computers (*any* computer with an RS-232C port), and a variety of communications programs. However, if you have an IBM® Personal Computer, Hayes also provides the perfect companion software.

Smartcom II™ communications software. For the IBM PC.

We've spent a lot of time developing it, so you can spend less time using it.

Smartcom II makes quick work of the communications process—reducing to mere keystrokes what would take a lengthy series of entries on other programs. Because Smartcom II is menu driven, you simply respond to the screen's prompts. There are no complicated procedures to learn or troublesome instructions to unravel. Success comes swiftly. Simply!

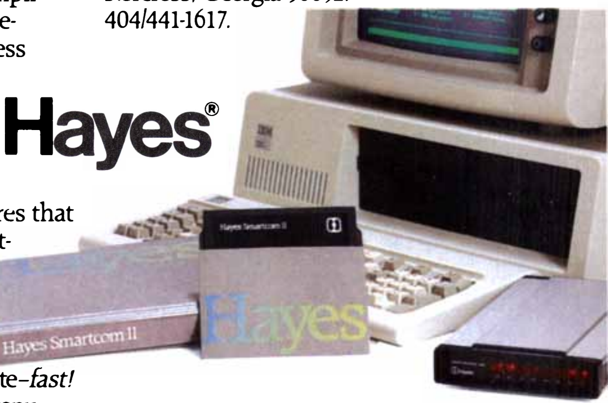
And now, a complete plug-in telecomputing system for the IBM PC. Smartmodem 1200B.™ All the great features that have made our freestanding Smartmodem a hit. On one, convenient plug-in board. And packaged with its own software. Smartcom II. Together, they're all you need to communicate–*fast!*

Backed by the experience and reputation of Hayes. Our record speaks for itself. Five years of solid leadership in the microcomputer industry. Nationwide availability through retail computer stores. Trouble-free factory service and call-in assistance. A limited two-year warranty on all hardware. And the most efficient telecomputing system available. Anywhere.

To appreciate the capabilities of Hayes Smartmodem and Smartcom II fully, see your dealer for a demonstration. Then get beyond "desktop" computing. Get a telephone for your computer.

Hayes Microcomputer Products, Inc.
5923 Peachtree Industrial Boulevard,
Norcross, Georgia 30092.
404/441-1617.

Hayes®



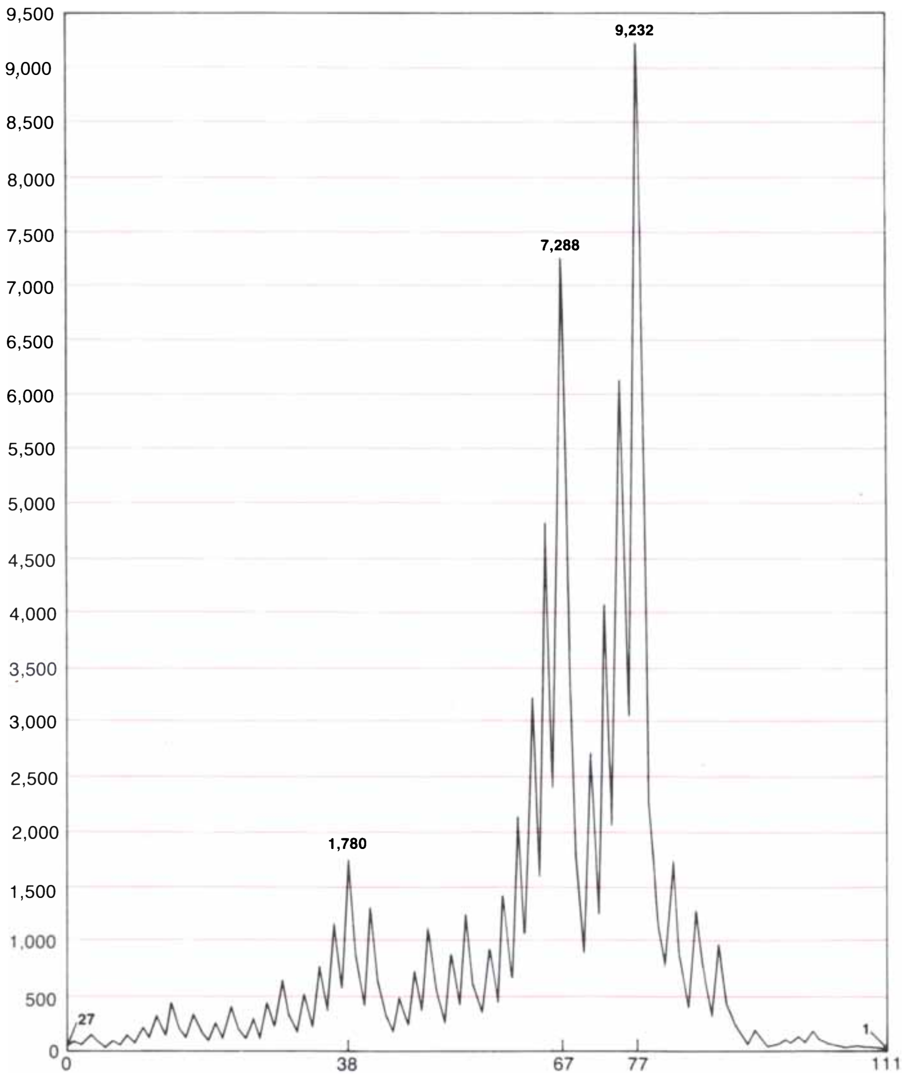
Smartmodem 300, Smartmodem 1200, Smartmodem 1200B and Smartcom II are trademarks of Hayes Microcomputer Products, Inc. Telemail Electronic Mail Service is a service mark of GTE Telenet Communications Corp. TeleSoft is a trademark of TeleSoft, Inc. Touch-Tone is a registered service mark of American Telephone and Telegraph. IBM is a registered trademark of International Business Machines, Corp. ©1983 Hayes Microcomputer Products, Inc.

COMPUTER RECREATIONS

On the ups and downs of hailstone numbers

by Brian Hayes

Three steps forward and two steps back: it is not the most efficient way to travel, but it seems certain to get you there in the end. A curious unsolved problem in the theory of numbers puts that conclusion in doubt. The problem can be stated as follows. Choose any positive integer (any whole number greater than zero) and call it *N*. If the number is odd, triple it and add 1, or in other words replace *N* by $3N + 1$. If the number is even, divide it by 2, replacing *N* by $N/2$. In either case the result is the new value of *N* and the procedure is repeated. After many iterations do the numbers tend to grow larger or smaller? Do they converge on some particular value or diverge toward infinity? How long does it take to settle the "fate" of a number?



The sequence of hailstone numbers beginning with 27

For any given value of *N*, answering these questions calls for nothing more than simple arithmetic. For example, if *N* is 27, an odd number, the next value is $(3 \times 27) + 1$, or 82; it is followed by 41 and then by 124. Evidently there will be many ups and downs in this series of numbers; the value goes up whenever *N* is odd and down whenever it is even. The reader is invited to extend the series to see where it leads.

The difficult task is not evaluating the series for a given *N* but finding a general solution, one that applies to all possible values of *N*. As yet no general solution has been devised. A great many numbers have been tested explicitly, and they all follow the same pattern, but no one has been able to prove that every number conforms to the pattern. It is hardly the most important unsolved problem in number theory, but it is one of the most irksome. The procedure is easy to describe and to carry out, but it is remarkably difficult to understand what is going on.

The problem illustrates well both the utility and the limitations of the digital computer as a mathematical instrument. To explore beyond the smallest integers some mechanical aid to computation is needed, but almost any computer will do, even a programmable calculator. On the other hand, extending the calculation to a significantly larger range of numbers is practical only with the most powerful computing machinery. When it comes to the very deepest questions, it is not certain any computer can be of help. For the most part the computer is a tool of "experimental" mathematics: it generates examples and counterexamples. Discovering a principle in the peregrinations of *N* seems to call for theorem proving rather than number crunching.

When the transformation rule is applied repeatedly to an arbitrary number, what outcome can be expected? Here are three naive hypotheses:

The first argument runs thus: There are equal numbers of odd and even integers, and so in any long series of calculations odd and even values of *N* should come up equally often. When *N* is odd, it is increased by a factor of 3 (and a little more), but when *N* is even, it is decreased by only a factor of 2. Hence the value of *N* after many iterations should increase without limit. On the average the value should increase by $(3N + 1)/2$ per iteration. For large values of *N* that is essentially $3/2\ N$.

The second hypothesis relies on the notion that what goes up must come down. This line of reasoning begins with the observation that whenever the calculation happens to yield an exact power of 2, the series of numbers immediately cascades back down to a value of 1. (When any power of 2 except 2 itself is divided by 2, the result is necessarily an even number, so that the descending branch of the calculation is invariably selected.) There are infinitely many exact powers of 2 among the infinite counting numbers, and a calculation that is continued long enough is certain to alight on one of them. Very large values of N might well be reached in the course of a calculation, but eventually there must be a crash.

The third argument is similar in form to the second but leads to a different conclusion. Note that whenever the calculation changes direction, such as when an odd number is encountered after a series of even ones, it reenters territory it has been in before. Indeed, in wandering up and down the number line it can return to a finite domain of numbers arbitrarily often. Eventually it can be expected to stumble onto a value it has visited before, and once that happens the entire future of the calculation is fixed. Because the procedure for choosing a next step is fully deterministic, any duplicated value of N must lead into a loop that will thereafter be repeated endlessly.

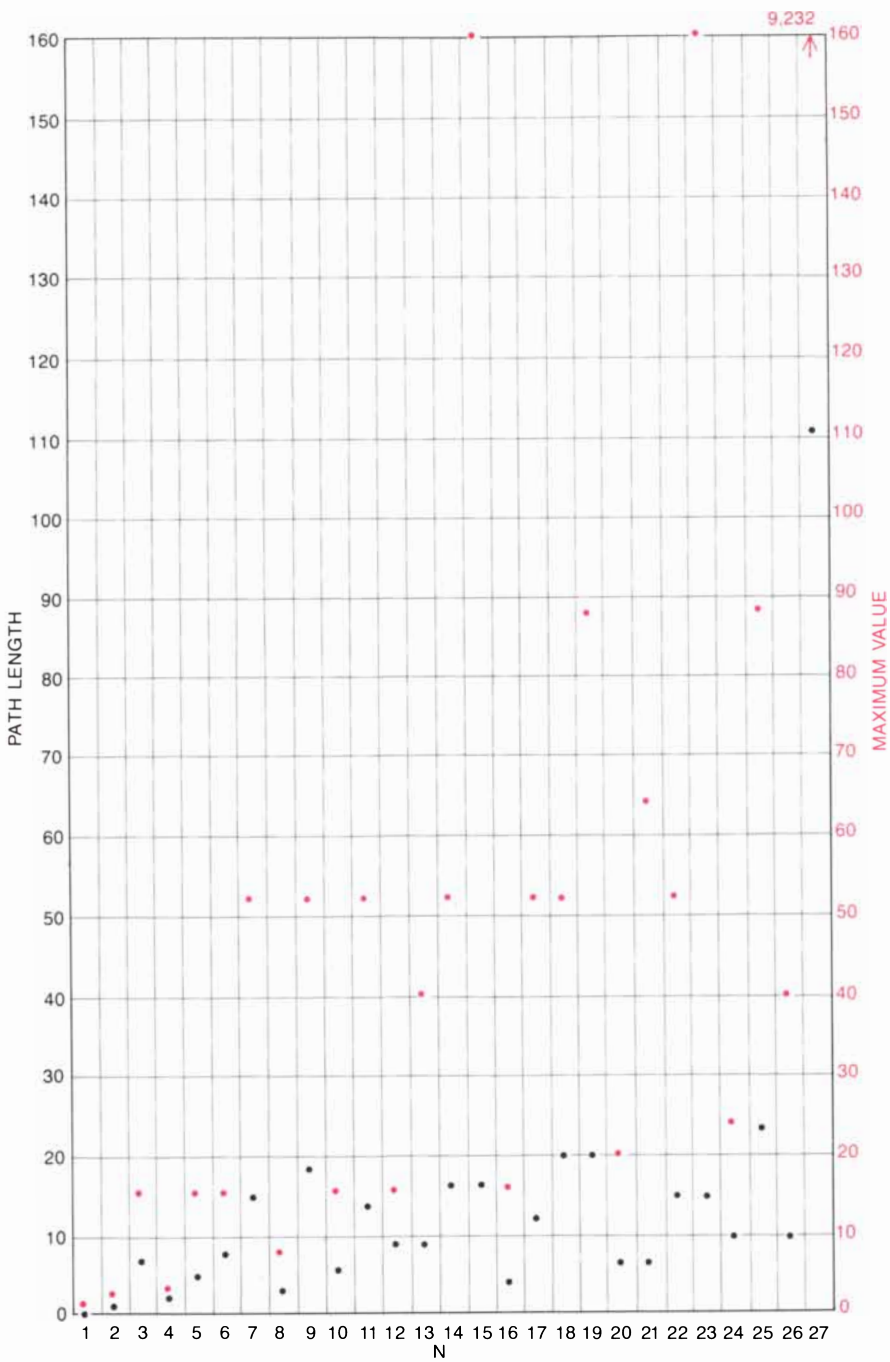
The three hypotheses presented here should not be taken too seriously. They cannot all be right. Some of their premises are definitely open to question. In particular, all three theories rely on a probabilistic analysis, but the series of numbers generated by applying the rule is not a random one. What does mathematical experiment have to say about the matter?

The place to begin the calculation is at the beginning, with 1. It is an odd number, and so the instructions call for multiplying it by 3 and adding 1. The result, 4, is even and is therefore divided by 2, yielding another even number; dividing by 2 again brings the calculation back to 1. Hence with the first computation two of the speculative theories cited above are given handsome support. As the crash hypothesis predicts, the calculation stumbles on a power of 2; it does so after just one iteration. As the cyclical theory predicts, the calculation becomes trapped in an endless loop; the values 4, 2 and 1 will be repeated indefinitely.

Among all the counting numbers 1 is very special: it is the first and the smallest. The results obtained when N is equal to 1 may therefore be atypical; before reaching any conclusions one ought to check further. Since the fate of 2 and 4 are known already from the calculation for $N = 1$, the obvious candidate is 3. It is odd, and so the next value is $(3 \times 3) + 1$, or 10. Dividing by 2 yields 5, and then multiplying by 3 and adding 1 gives a result of 16. Again a power of 2 has turned up, and the series cascades through $N = 8$ into the 4–2–1 loop.

After examining the first four natural numbers the trend seems clear, and yet there is still reason for doubt. In the calculations made so far two quantities of interest stand out: the highest value of N reached during a calculation and the path length, which I shall define as the total number of iterations needed to reach a value of 1. For 1 itself the maximum value is 1 and the path length is zero. For 2 the peak is 2 and the length is 1. For 3 the maximum is 16 and the length is 7. The example of 3 suggests that the maximum value reached and the length of the series can be much larger than the initial value of N, and so perhaps the function will turn out to be unbounded for some values of N.

Consider again the series generated when the initial value is 27. As noted above, the first three numbers are 82, 41 and 124, but two successive divisions bring the series back down to 31. Hence after five steps almost no progress has been made. As the calculation continues, however, the three-steps-forward, two-steps-back mechanism gives rise to a series of oscillations of ever larger amplitude. New peaks are reached at 142, 214, 322 and 484. There are further setbacks (at step 19 the value has dropped to 91), but the trend continues to be upward. The calculation passes through 700, through 1,186 and through 2,158



The path length **(black)** *and the maximum value reached* **(color)** *for the first 27 integers* and by the 77th iteration has reached the substantial value of 9,232. It seems we are on our way. As it turns out, however, the path ends at 1 after a total of 111 steps, never having risen higher than 9,232. (The complete path is shown in the illustration on page 10.)

Calculations of the kind I have just traced have been made for all the integers in an extremely wide range. Nabuo Yoneda of the University of Tokyo has tested all values up to 2^{40}, or 1.2×10^{12}. In every case the result has been the same: after a finite number of steps the series subsides into the 4–2–1 loop, where it must stay forever. Among the first 50 integers 27 has the longest path back to 1 (although 41 and 31 are not much shorter and reach the same peak value, for reasons that should be apparent from the information given above). No positive integer has been found to generate a series that goes off toward infinity, and no loops other than the 4–2–1 loop have been found. Nevertheless, the conjecture that all positive numbers conform to the same pattern remains without a secure theoretical basis.

N	PATH LENGTH	MAXIMUM VALUE
1	0	1
2	1	2
3	7	16
6	8	16
7	16	52
9	19	52
18	20	52
25	23	88
27	111	9,232
54	112	9,232
73	115	9,232
97	118	9,232
129	121	9,232
171	124	9,232
231	127	9,232
313	130	9,232
327	143	9,232
649	144	9,232
703	170	250,504
871	178	190,996
1,161	181	190,996
2,223	182	250,504
2,463	208	250,504
2,919	216	250,504
3,711	237	481,624
6,171	261	975,400
10,971	267	975,400
13,255	275	497,176
17,647	278	11,003,416
23,529	281	11,003,416
26,623	307	106,358,020
34,239	310	18,976,192
35,655	323	41,163,712
52,527	339	106,358,020
77,031	350	21,933,016

Sequence of longest paths up to N = *100,000*

N	PATH LENGTH	MAXIMUM VALUE
1	0	1
2	1	2
3	7	16
7	16	52
15	17	160
27	111	9,232
255	47	13,120
447	97	39,364
639	131	41,524
703	170	250,504
1,819	161	1,276,936
4,255	201	6,810,136
4,591	170	8,153,620
9,663	184	27,114,424
20,895	255	50,143,264
26,623	307	106,358,020
31,911	160	121,012,864
60,975	334	593,279,152
77,671	231	1,570,824,736

Sequence of peak values up to N = *100,000*

The $3N + 1$ problem, as it is generally called, has a murky history, but it does not seem to be of great antiquity. Over the past 30 years or so it has turned up repeatedly in various university departments of mathematics and computer science, its comings and goings seeming to be as capricious as the advances and recessions of the numbers themselves. Jeffrey C. Lagarias of Bell Laboratories, who has recently looked into the origins of the problem and the prospects for solving it, notes that it may have been invented several times. In the 1930's Lothar Collatz, who was then a student at the University of Hamburg, investigated a class of problems that includes the $3N + 1$ problem, although the work was not published until many years later. In 1952 the British mathematician B. Thwaites independently discovered the problem, and a few years later it was invented yet again by Richard Vernon Andree of the University of Oklahoma at Norman.

Lagarias cites some 20 research articles on the $3N + 1$ problem and its generalizations, most of them published within the past 10 years, but the problem had circulated by word of mouth long before. Collatz' colleague Helmut Hasse introduced it at Syracuse University in the 1950's, and Stanislaw Ulam took it to Los Alamos and elsewhere. Shizuo Kakutani, who first heard of the problem in about 1960, reported to Lagarias: "For a month everybody at Yale worked on it, with no result. A similar phenomenon happened when I mentioned it at the University of Chicago. A joke was made that the problem was part of a conspiracy to slow down mathematical research in the U.S."

Another sustained attack on the problem, with an emphasis on computer-aided numerical calculations, was made in the early 1970's by a group in the Artificial Intelligence Laboratory at M.I.T. The problem is recorded as Item 133 in the group's informal (and unpublished) transactions, called HAKMEM, or "hackers' memorandum."

In its wanderings the problem has been known by many names. Calling it the $3N + 1$ problem does not seem entirely satisfactory, in that it gives undue attention to one half of the procedure and slights the other half. Of the various alternatives the one I find most congenial identifies the numbers generated from a given starting value as "hailstone numbers." The path the series follows is rather like the trajectory of a hailstone through a storm cloud, rising in updrafts and then falling under its own weight.

A computer program for calculating hailstone numbers can be written in a few lines of a higher-level programming language such as BASIC. Indeed, the central algorithm can be expressed in a single statement. In BASIC it might be

```
IF N MOD 2 = 0 THEN N = N/2
ELSE N = 3*N + 1 .
```

Here the first operation is one that people (but not computers) are capable of doing without explicit calculation: determining whether N is odd or even. N MOD 2 is a modulus operation, which computes the remainder when N is divided by 2. If the remainder is 0, the THEN part of the statement is executed and N is set equal to $N/2$; otherwise the ELSE part is executed, setting N equal to $3N + 1$.

A program in BASIC serves well enough for generating hailstone numbers from the first few hundred integers, but if more extensive calculations are undertaken, it becomes intolerably slow. The BASIC statement calls for a division (as part of the modulus operation), a comparison and then either a second division or a multiplication and an addition. Division and multiplication are time-consuming operations, particularly in a small computer system. There is much to be gained here by speaking directly to the central processing unit in its own language. All the division and multiplication operations can thereby be eliminated.

The illustration on the opposite page gives a schematic account of such a machine-language program. It is assumed that the value of N is initially in a register designated AX, which also serves as an "accumulator" where arithmetic operations are done. The value at the start of the procedure is the binary representation of the decimal number 27.

The first step is to save a copy of the initial value in another register, here labeled BX. The division operation is avoided by exploiting a property of the binary number system: shifting a binary number to the right one position is equivalent to dividing it by 2, just as shifting a decimal number to the right divides it by 10. In the course of the shift the rightmost digit (the units digit) is preserved in a one-bit storage location called the carry flag. Testing the carry flag determines whether the original number was odd or even, since in binary notation every odd number ends in a 1 and every even number ends in a 0.

If N is even, the calculation is now finished. The value remaining in register AX after shifting to the right one place is the quotient $N/2$. In this case, however, N is odd and further computations are needed. First the original value of N is recovered from register BX. Then, instead of multiplying by 3, the value is added to itself twice; even though this requires two machine instructions instead of one, it is done appreciably faster. The final step is to increment the number in AX by 1. In the instruction set of one microprocessor the entire procedure takes 20 cycles of the computer's clock when N is even and 18 cycles when N is odd. At a clock frequency of roughly five megahertz the program fragment could in principle be executed some 250,000 times per second. (A few more clock cycles could be saved, at some cost in program clarity.) The equivalent algorithm employing division and multiplication instructions takes 175 cycles for even N and 286 cycles for odd N.

In the illustration registers are shown as being eight bits wide and can therefore accommodate numbers no larger than 2^8, or 256. In most microprocessors the registers are actually 16 bits wide and can hold numbers up to 65,536. Even that limit is a severe constraint; a program employing 16-bit arithmetic could not calculate hailstone numbers beyond $N = 702$. Achieving a higher capacity requires multiple-precision arithmetic, in which a single number is split between two or more registers or memory locations. With 32 bits of precision numbers up to about four billion can be represented; 64 bits extend the limit to 10^{19}. Each increase in precision, however, exacts a penalty in speed.

The algorithm for calculating one value of N is only a fragment of a working program. In addition there must be some facilities for getting input values into the machine and for displaying results. A practical set of programs for exploring the hailstone numbers ought to do a good deal more. For example, it should be possible to print out the entire series of numbers generated by a given starting value, or to list the path length and maximum value associated with all the integers in a given range. Another program could be set up to search for integers yielding progressively longer paths or larger peak values. There are many other possibilities.

Variations on the $3N + 1$ formula employing different coefficients and constants are also worth exploring. R. William Gosper and Richard Schroeppel, when they were members of the HAKMEM group, investigated the $3N - 1$ problem and showed that it is equivalent to the $3N + 1$ problem with negative values of N. Every number they checked terminates in one of three loops; the longest loop begins at $N = 17$ and has a period of 18 steps.

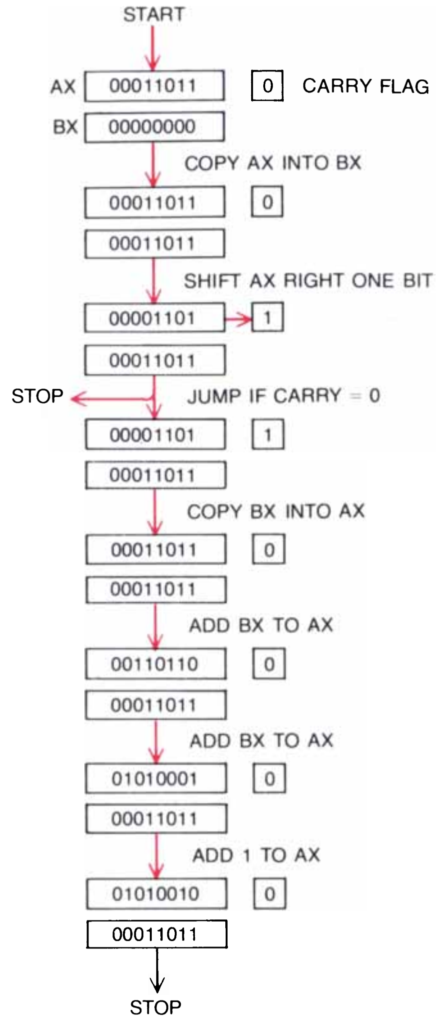
A program whose only aim is to search for numbers that do not fall into the 4–2–1 loop can be greatly streamlined. If numbers are checked in succession beginning with 1, only odd numbers need to be examined. Any even number is immediately reduced by half, and so the path it generates would already have been detected. For similar reasons there is no need to follow the path of a number all the way to 1; once the value of N falls below the initial value the candidate can be dismissed. Still more effective rules for narrowing the search have been developed by William H. Henneman, a student in the HAKMEM group who is now at Boston University.

Although no proof has yet been discovered, a hint of an explanation may lie in a heuristic argument more refined than the three naive hypotheses given above. There it was noted that in any stage of the calculation N has an equal probability of being multiplied by 3 or divided by 2, leading to the suggestion that the value should tend to increase by a factor of 3/2 per iteration. Lagarias points out, however, that one-fourth of all the integers are divisible not only by 2 but also by 4; one-eighth of them are divisible by 8, one-sixteenth by 16 and so on. Taking into account divisions by all possible powers of 2 yields a prediction that N should decrease by a factor of 3/4 per iteration. The empirical evidence supports the prediction.

Even if it turns out that all positive integers fall into the 4–2–1 loop, the hailstone numbers offer an abundance of curiosities. Perhaps the most intriguing properties of the numbers are conspicuous patterns in the distribution of path lengths and peak values. If a number as small as 27 can keep the ball in the air for 111 steps and reach a height of 9,232, one might well expect that the path length and the peak value would grow rapidly as N increased. Actually the path length grows very slowly; the increase in the maximum value is faster, but it is also quite erratic.

Among the first 100 integers the longest path is 118 steps (at $N = 97$); among the first 100,000 integers the longest path is just 350 steps (at $N = 77{,}031$). Thus increasing N by a factor of 1,000 increases the path length by a factor of only 3; the relation appears to be a logarithmic one. The record maximum of 9,232 set at $N = 27$ is not exceeded until $N = 255$, which reaches a peak of 13,120. New maximums are recorded at quite irregular intervals. The hailstone sequence for $N = 77{,}671$ reaches the extraordinary height of 1,570,824,736.

It is easy to see that the peak value reached in a hailstone calculation must invariably be an even number. It can also be proved that only an odd value of N can set a new record for maximum height (with the possible exception of $N = 2$). In the case of numbers that set new records for path length there is no

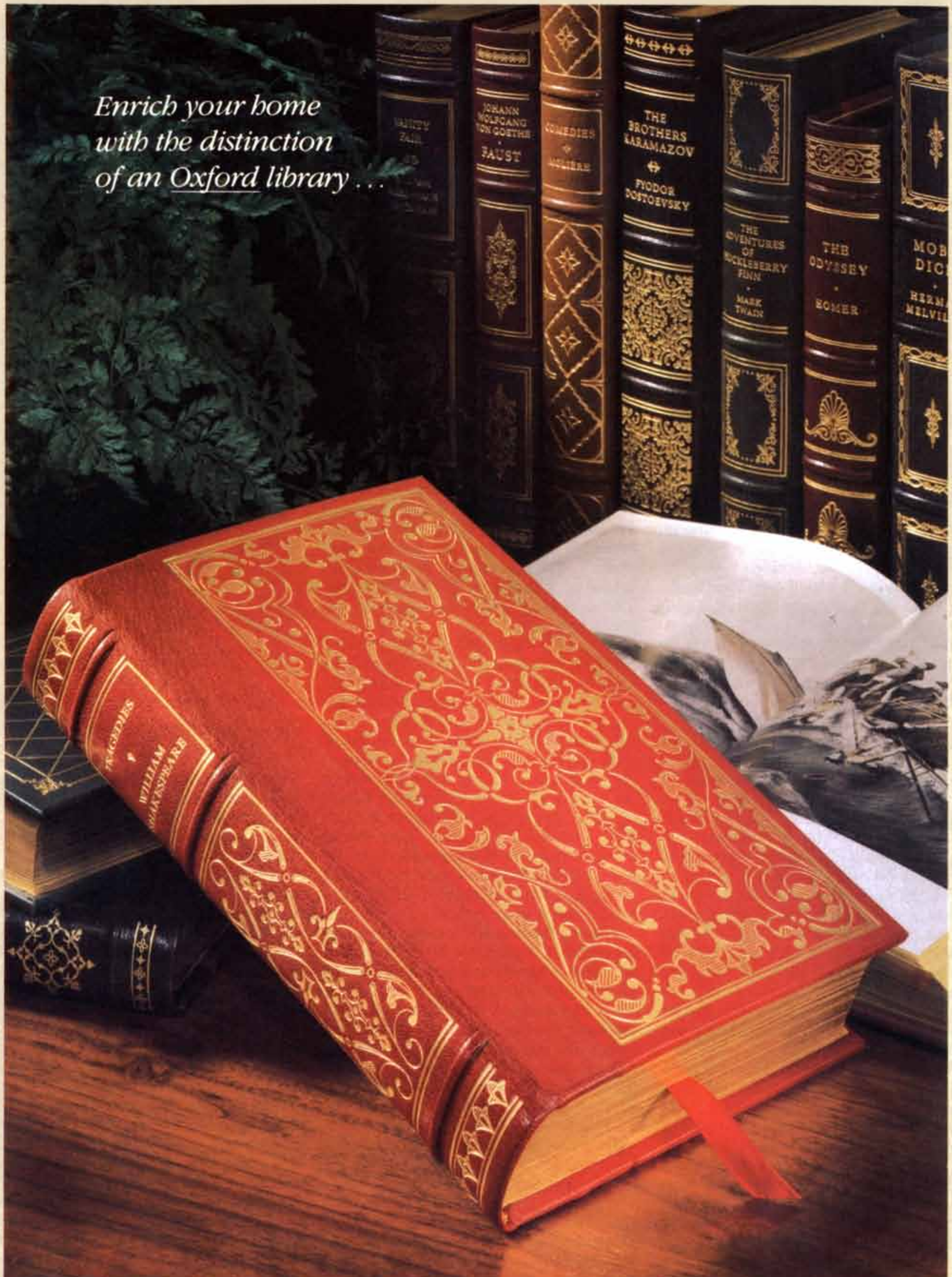


The hailstone algorithm

theoretical argument I know of that requires them to be either odd or even. Nevertheless, among the first 100,000 integers path-length records are set almost exclusively by odd values of N.

A listing of the path length and maximum value for a range of numbers has a frustrating mixture of regularity and disorder: it is definitely not random, but the pattern resists interpretation. For instance, certain maximum values are much commoner than others and far too common to be explained by any statistical process. The outstanding example is 9,232, the number first reached at $N = 27$. Of the first 1,000 integers more than 350 have their maximum at 9,232.

The distribution of path lengths is equally peculiar. Every possible length can be produced (by the successive exact powers of 2), but again some numbers appear far more often than others. Moreover, both the path lengths and the maximum values show a strong tendency to form clusters. In 1976 Fred Gruenberger of California State University in Northridge published a list of such clusters; the largest was a string of 52 consecutive numbers that all have the same path length. Can two consecutive val-


Enrich your home
with the distinction
of an Oxford library . . .
FAUST
THE BROTHERS KARAMAZOV
FYODOR DOSTOEVSKY
THE ODYSSEY
HOMER

Definitive and authoritative – the one great private library brought together by the experts of Oxford University

The Oxford Library of the World's Great Books

In quarter-leather volumes of classic distinction — the leather inlaid with 22 karat gold.

Excellent value at $24.50 per volume.

OXFORD, ENGLAND. For the first time in the history of England's oldest university, a board of eminent Oxford dons has come together to select 'The Oxford Library of the World's Great Books.'

From out of the many thousands of books written over the centuries, the board has chosen a collection of fifty universal literary masterpieces: the supreme works of the world's most brilliant writers. A source of inspiration that stands at the core of our civilization — the indispensable heritage of future generations.

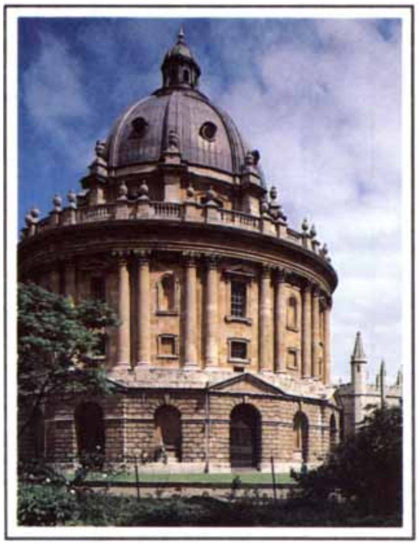
And 'The Oxford Library' will now be published in a manner befitting its historic status. Each book in this magnificent Collector's Edition will be a splendid example of the traditional English quarter leather volume. The leather inlaid with 22 karat gold ... the entire collection as exciting as the literature it contains.

A collection as worthy as it is beautiful

This is a collection like no other: The Oxford Library of Great Books. The definitive private library with which to surround oneself and one's family — brought together by some of Oxford's most brilliant minds.

Here are the towering novels — Tolstoy's *Anna Karenina,* Melville's *Moby-Dick,* Dickens' *Great Expectations.* The grand epics — Homer's *Odyssey* and Milton's *Paradise Lost.* The tragedies of Shakespeare. The comedies of Molière. The immortal works of Dante, Swift, Shelley, Keats ... Conrad, Mann, Joyce, Yeats ... and other great authors from Chaucer to F. Scott Fitzgerald. And to heighten the pleasure of reading each book, each will contain a valuable introduction, commissioned by the Oxford University Press exclusively for 'The Oxford Library.'

Furthermore, the Oxford University Press has authorized the world's foremost creator of heirloom volumes, The Franklin Library, to craft each volume in the Collector's Edition. And they will 'quarter-bind' each volume in the classic design tradition recommended at the turn of the century by the English Society of Arts.



Craftsmen of exceptional skill will bind the spine of each volume in genuine leather. They will carefully work the leather to achieve the handsome raised 'hubs' so prized by book collectors ... and will then inlay the leather with original designs of 22 karat gold.

Equal care and craftsmanship will be evident in the superbly embellished cover fabrics, the marbled endpapers, the richly gilded page-edges and the elegant ribbon bookmark bound into each volume. And even the papers for the text pages will be custom-milled by a costly process that enables them to retain their beauty for literally hundreds of years.

Each binding different ... to a purpose

Each sumptuous, quarter-leather volume will be strikingly *different* from the others — and not simply for the sake of being different. For each volume has been individually designed to perfectly complement its literary content.

Shakespeare's *Tragedies,* for example, is quarter-bound with crimson leather, befitting the splendor and richness of the literature. The book is inlaid with a motif recalling the famous "Azure Tooling" that distinguished many fine volumes of the 1600s. Even the typeface can trace its heritage to a special style that originated at the time of Shakespeare's First Folio.

Thus, when the complete collection is arranged in the owner's home, its variety will fire the imagination. The different leathers, each with its own pleasing grain, color and texture ... the beautiful fabrics, with their varied designs ... the fascinating patterns of hubbing and 22 karat gold inlay on the spines ... and all the different styles of *illustration,* including the work of such masters as Tenniel, Rowlandson and Doré.

Printed and bound only for subscribers

The books are available only by subscription — and will be issued on a monthly basis at $24.50 per volume. This is indeed *outstanding value* today, considering all the advantages that these extraordinary quarter-leather volumes have to offer their owner: The certainty of possessing the very greatest works of literature. The pleasure of seeing and handling and reading these beautiful, individually designed volumes whenever you wish. The satisfaction of knowing that this magnificent private library, by its very presence in your home, encourages younger members of the family to experience great literature and to share your own appreciation of excellence as a desirable goal in life.

Apply by February 29, 1984

In addition, the original issue price of $24.50 per volume will be guaranteed not to increase throughout the entire subscription period. Yet you take on no long-term commitment whatever. You will have the right to cancel your subscription at any time, upon 30 days' written notice — or return any volume for any reason, within thirty days, for your choice of replacement or a full refund.

Please note, however, that the collection is available *only* by subscription. And there is a timing requirement for subscribing. Be sure to return your application by February 29, 1984.

ues of N have the same path length and the same maximum? The question can be settled algebraically, but readers who prefer a numerical demonstration might want to examine the hailstone sequences for $N = 386$ through $N = 391$.

One illuminating way to look at the hailstone problem is to turn it upside down. Suppose it is true that all positive numbers ultimately fall into the 4–2–1 loop. They must then form an unbroken chain through which any number in the infinite counting sequence is connected to the bottom of the loop. Accordingly it should be possible to invert the hailstone function: to begin with 1 and apply the transformation "backward" in order to generate every larger number. If some number cannot be reached in this way, by following the river upstream, the number cannot yield 1 as its final value.

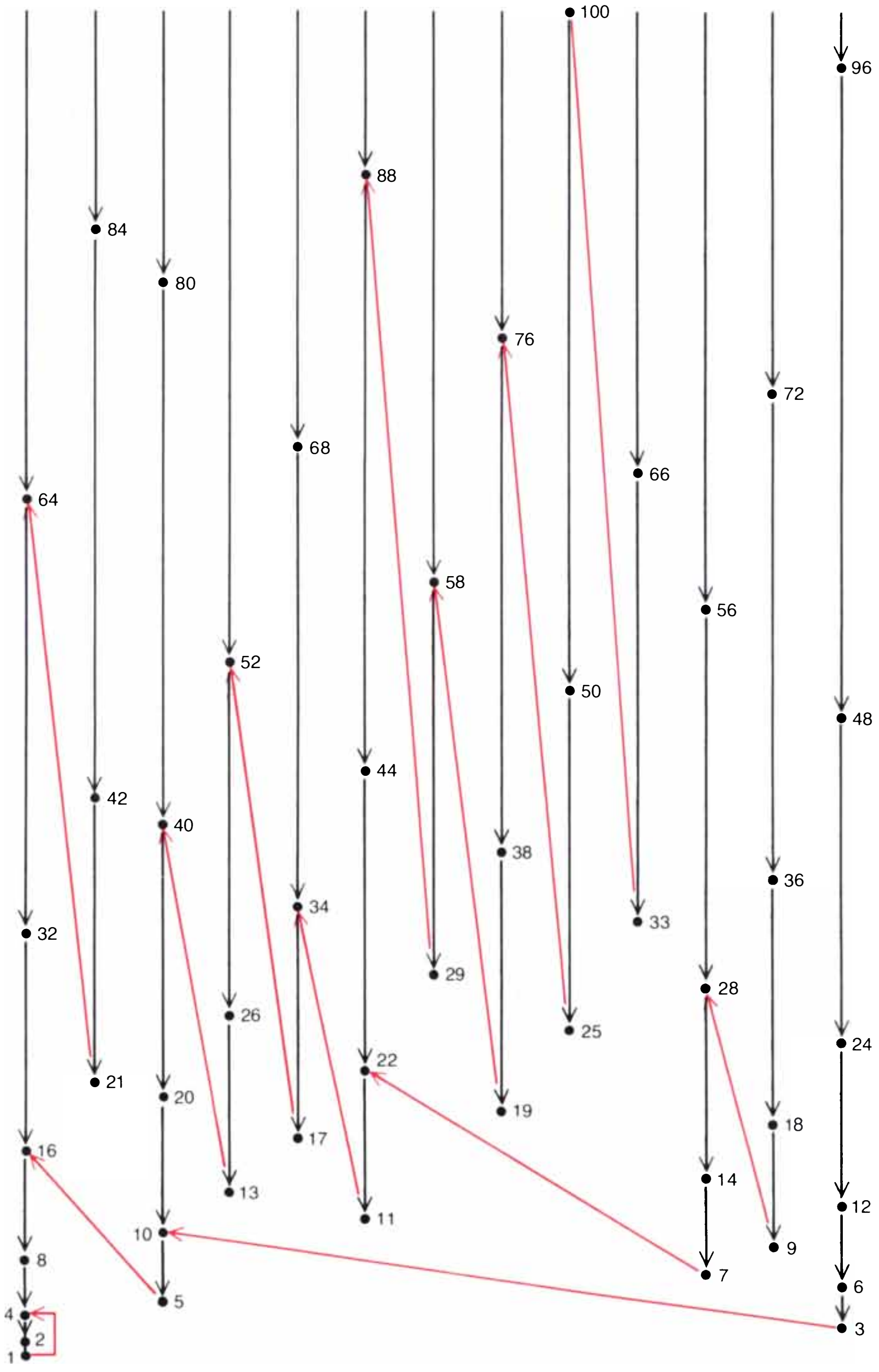
The method might well yield a general solution of the hailstone problem, if only it could be carried to completion. As it turns out, the procedure is not as straightforward as it seems. The normal hailstone function is deterministic: a value of N at any point in the calculation can have only one possible successor. If N is 40, for example, the next number can only be 20. When the path is traced in reverse, there are ambiguities. When the value $N = 20$ is encountered, it is known it could only have been generated from 40, which must therefore be the next value. At 40, however, the next value could be either 80 or 13; the stream splits, and both tributaries must be explored. There is a bifurcation at every number of the form $6K + 4$, where K can be zero or any positive integer.

A branching system of this kind can be traced only to a finite depth. A single branch must be followed until some predetermined limit is reached and then attention must be diverted to another branch. When the limit is set at 100, 13 branches are explored and 49 numbers are confirmed to be connected to the system of numerical rills and rivulets. With the limit at 1,000 there are 84 branches, but only 340 numbers are counted. A limit of 10,000 yields 1,065 branches, which pass through 4,235 numbers. Note that more than half of the numbers seem to lie in the interstices between the branches of the stream. As the limit is increased more numbers are included, but even more are missed. If the ramifications of the system could be explored to infinite depth, would all positive integers ultimately find a place in it? That is the big question yet to be answered.

In October two combinatorial problems were mentioned as being unlikely candidates for solution by "nonalgorithmic" methods with an electronic spreadsheet. A number of readers were quick to show that it can be done.

Spreadsheet solutions to the Tower of Hanoi problem were sent by David Behar, John B. Jones, Jr., George Arthur Miller, J. B. Sladen, Alun Wyn-jones and others. The techniques employed were similar. An algorithm in which odd-numbered disks circulate clockwise and even-numbered ones counterclockwise was disassembled so that a temporal sequence of instructions became a spatial array of them.

Behar, Miller and Sladen also solved the eight-queens problem. Here the main difficulty is the need to backtrack when a developing solution fails. I had supposed some record of previous unsuccessful attempts would have to be kept, but that is not the case. D. H. Fremlin, in a commentary on the spreadsheet approach to the problems, pointed out "a method that will generate all the moves in turn without any memory other than what is displayed.... The formulas are complicated and involve sequential algorithms if they are done in their natural forms, but all these 'look for' subroutines can be represented on the spreadsheet by local calculations."



Tree generated by inverting the hailstone function

The Computer in Education

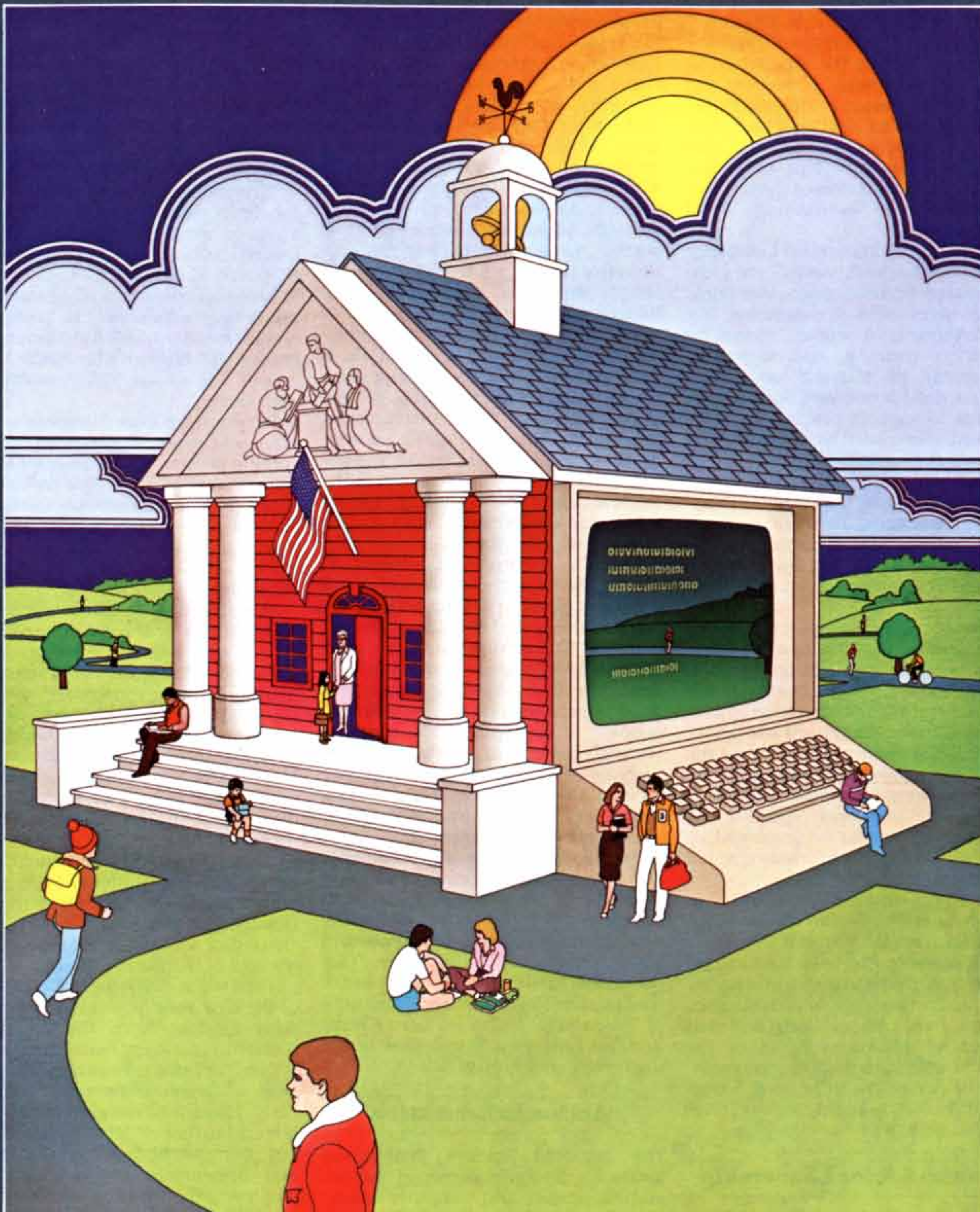


ILLUSTRATION BY ROBERT FROST

SCIENTIFIC AMERICAN

ADVANCED TECHNOLOGY CENTERS USA ADVERTISING SUPPLEMENT / SCIENTIFIC AMERICAN / JANUARY 1984

More than ten years ago Daniel Bell, in *The Post Industrial Society*, wrote "In the post-industrial society technical skill becomes the base of power and education the avenue to power...." Ten years later Bell's conclusions have been reinforced by the experiences of leaders in business, education and government. Last year three such men published *Global Stakes: The Future of High Technology in America.* They are Ray Stata, president of Analog Devices, Inc., and James Botkin and Dan Dimanescu, consultants, all of Massachusetts.

"As society shifts toward knowledge intensity," the book states, "the critical resource becomes not conventional capital investment in machinery, but the investment in people... money for education, training, and research... investment in physical equipment such as word processors or computer systems to support people's activities or, even more so, the software that operates the equipment."

Scholastic Aptitude Test (SAT) scores "reveal a 22% drop in the number of top-scoring mathematics students," *Global Stakes* documents. "In 1972, a total of 93,868 students had scores of 650 or above versus 73,386 in 1980...."

Congress has taken note. In March 1983, Senator Paul Tsongas of Massachusetts introduced the High Technology Morrill Act, "a bill to establish a national technology education grants program." Tsongas named the bill after the 1862 law that established the Land Grant colleges. Funding for the 1862 law came from assessments on railroad rights-of-way; funding for Tsongas's bill would be generated by assessments against oil leases granted on Federal lands.

Thus both industrialists and academics have sounded an alarm: American education isn't what it should be. Simultaneously the "information age" has arrived, partly in the form of highly capable, increasingly affordable microcomputers and similarly any discussion of education turns to the point of computers—and vice versa. How the two ought to be used is now the subject of a highly constructive national debate.

Private-Sector Leadership

Analog Devices, Inc. is one of the advanced-technology companies clustered about Boston's Route 128.

In a January 1983 speech before a Massachusetts High-Technology Council audience, which included President Reagan, Analog's president, Ray Stata, said, "One-third of our work force needs college degrees, more than half in science and technology.

"The simple fact is that our educational system today does not have the capacity to educate and train enough people, with the right qualifications, to sustain our industry's growth."

Analog is a manufacturer of components and systems for precision measurement and control, which is to say real world signal-processing. Over the three years of 1980-82, Analog and its employees contributed more than $1.25 million to 45 educational institutions. As Analog vice president for strategic planning, Graham Sterling states, "We've made it clear that our industry *cannot* do the whole job.. it simply won't get done without enhanced state and federal support."

The sense of "newness" of microcomputers, in or out of education, is well-founded, since few leading firms are as many as ten years old. Yet one major company, Texas Instruments, was founded back in 1930; at that time its main business was oil and gas exploration. Today the company is the world's largest manufacturer of semiconductor components, with a clear focus in education.

From a corporate standpoint TI shares the urgency of the authors of *Global Stakes.* In a 1982 address before the Council of Chief State School Officers, TI executive vice president William Sick pointed out that the firm, according to its own growth plans, will need large numbers of new technical people. "If current trends continue through the end of the decade," Sick demonstrated, "TI will have to hire 12% of all electrical engineering and computer science graduates. That clearly will be a problem...."

Bernie List is Texas Instruments vice president/corporate staff and manager of corporate education and training. "Our commitment to the educational market is an intense and highly personal one," List states. "Our president, Fred Busy, has been head of the Board of Regents at Texas Institute of Technology. That job's been a pipeline, for Fred, to a 'microcosm' of the university community.

Action in Education

The National Science Foundation funded a 50-state survey of mathematics, science and computer education initiatives. Over half the states reported they had established task forces or commissions on computer education. Thirty-two states reported state-level computer education programs. Marc S. Tucker is Director of the Project on Information Technology and Education in Washington. Tucker's two-year-old Project, funded by the Carnegie Corporation of New York, is intended to examine the role of technology, specifically computers, in education.

"No parent of the fifties ever felt that his or her children would badly damage career opportunities if they failed to master the 8mm film loop projector. They did not send their children to TV camp or buy them home-language laboratories. Something else is happening here," Tucker states.

"Scarcely any area of human activity has been more resistant to scientific analysis and technological change than education," wrote B.F. Skinner in SCIENTIFIC AMERICAN back in 1961. "There is no dearth of suggestions for improving education. In assigning certain mechanizable functions to machines, the teacher will emerge in his proper role as an indispensable human being."

Twenty years after Skinner's article, more and more "suggestions" are becoming actions. The Center for Children and Technology at New York's Bank Street College of Education was founded three years ago "to bring thoughtful research to bear on burgeoning electronic technology and its role in children's lives."

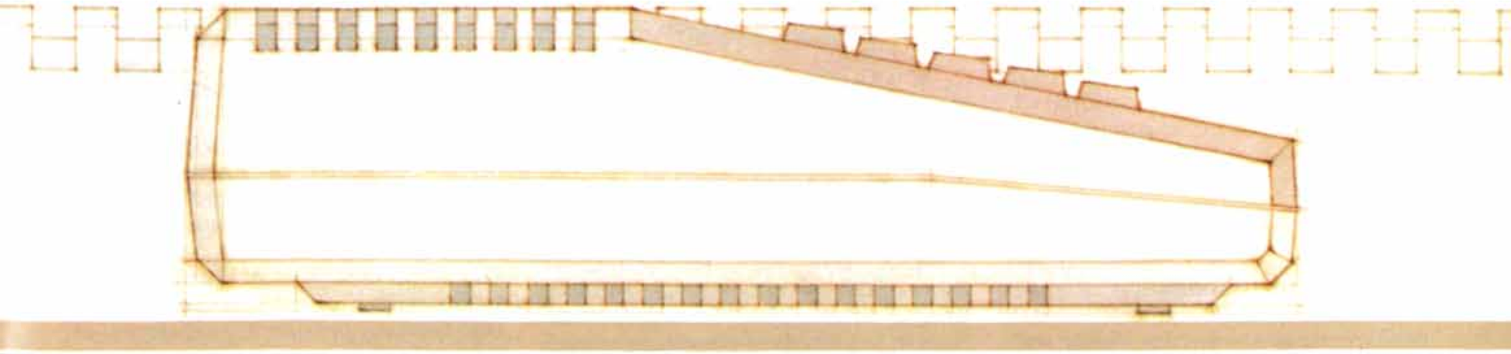
At the Bank Street Children's School, children cluster around three computers. At one, a pretty third-grade girl draws complex figures on a screen. At another, three youngsters use the school-devised, and now commercial, word processing program "Bank Street Writer" to develop a story about a train.

"There is a great deal of chaos and confusion," says Karen Sheingold, director of the Bank Street Center. "Some of the reasons why we exist are to sort it all out. Funders include the Xerox and International Paper Company Foundations as well as the U.S. Department of Education.

Another active program is New York's "School of the Future" project, originally started in 1980 under the auspices of the New York Academy of Sciences, with funding from the Atari and Richard Lounsbery Foundations.

Director Bonnie Brownstein is a disciple of Seymour Papert of MIT, an early advocate of computers for children and author of "Mindstorms, Children, Computers and Powerful Ideas." Ms. Brownstein points out that—whatever the desired end-result of computers-in-education may be—children are very much part of the process.

The computer responds at once to the actions of a child. Interaction at any level is educational in some manner. Ms. Brownstein observes, "To place a point on a monitor screen, you must have two number elements: the

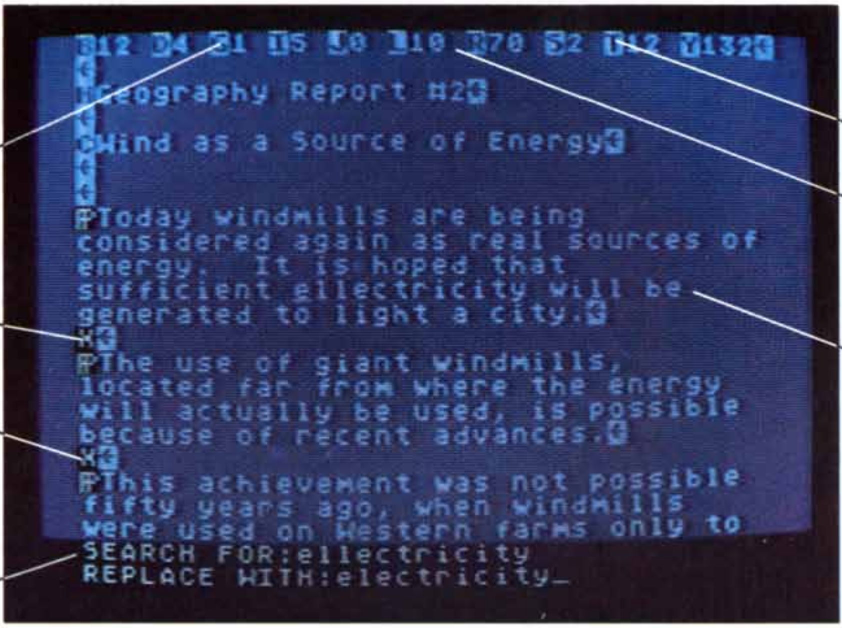


Prints with built-in format or lets you create your own: center, underline. Boldface, elongated, proportional and condensed print.

Inserts and deletes characters, lines or blocks of text.

Duplicates and moves blocks of text.

Searches for and/or replaces words or phrases, one at a time or all at once.



Changes line spacing, margins, print type, paragraph indents anywhere in the document.

Instantly reformats.

Enters text easily, with word wrap, one main menu and helpful prompts.

Creates multi-line headers and footers.

Prints double columns, form letters, multiple copies of a page or document, at the touch of a key.

Numbers pages and creates section numbers automatically.

Corrects spelling (with extra software).

Chains and merges files.

Has "print preview": you can check your copy before it's printed.

You can't find a friendlier, more powerful word processor at twice the price. New AtariWriter.™ Under $100.

Now you can do multi-featured word processing at home, simply. At a family budget price.

Our ROM-based cartridge technology means you can use new AtariWriter on any ATARI® Home Computer (even 16K) for personal and business correspondence, term papers, committee reports, mailings, etc.

It also lets you choose between cassette and disk storage systems.

One very special AtariWriter feature: you can correct as you write, without switching back and forth between Create and Edit modes.

And our memory buffer offers an "undo" command to let you change your mind, and restore text you've just deleted.

Check into our remarkable AtariWriter, and our choice of letter quality and dot matrix printers, also reasonably priced, at Atari dealers. Call 800-538-8543 for dealer nearest you. In California, call 800-672-1404.

You'll do more with Atari Home Computers.

X and *Y* coordinates. Five-year-olds understand that. It's revolutionary."

Much of the national debate on the computer's role in education centers on the extent of the "revolution."

The American Federation of Teachers (AFT) in Washington produced a resolution at its 1982 annual convention dealing with computers in education. "Schools have an obligation to prepare students for the highly technological age in which they will live," reads the resolution's preamble.

The National Education Association has contracted with a private company, Cordatum, Inc., to help evaluate and certify microcomputer educational software. "We will set standards and guidelines for what educational software should look like," says Dr. Larry J. Fedewa, executive director for the project, introduced in May 1983."

In Dr. Fedewa's assessment, it is teachers themselves, who most strongly influence purchase decisions regarding course materials. Where standards are lacking, teachers have learned to consult each other. One result: the Minnesota Educational Computing Consortium (MECC).

"MECC arose in 1973 to try to put some order in the 'pockets' of computing all over Minnesota," spokesperson Shirley Griffing relates. "Our State is fortunate in having highly concerned private-sector influences like Honeywell and Control Data and 3M.

MECC now evaluates and offers thousands of educational software program packages. "We supply materials to hundreds of educators outside Minnesota. It's all word of mouth," says Ms. Griffing.

"Orders come in from all over the U.S., plus Germany, Japan, New Zealand, Belgium, even Botswana," she concludes. "Botswana! I had to look it up."

Sarah F. Klein, past president of the National Science Teachers Association, considers the computer's role in society as "a way of life." She points out that "A young person's first exposure to the working world is often a job employing computers—cashiering in a supermarket, or waiting on tables, or pumping gasoline. It would be an injustice to future generations if we didn't recognize the value of computers in the learning process."

Computers in the Grades

McGrath, Alaska is approximately 300 miles northwest of Anchorage; its nearest villages are Crooked Creek and Red Devil. In part because of its remoteness, McGrath has four to five times the computer-to-student ratio as the nation as a whole, according to Dan Shanis, president (and owner) of Educational Services of Alaska. The main reason for the large volume of microcomputers in Alaska's schools, he adds, is money: oil money, distributed by the Alaska Department of Education, which funds local school budgets.

McGrath and other remote Alaskan areas are a sort of prototype for the possible effects, nationwide, of a massive infusion of government-administered money into education. "We've no hard data, but I can tell you the program has been a real success," says Shanis. "First of all, computers are intrinsically interesting, and most kids respond to them. Secondly, we've pushed for programs which really improve thinking skills—not just programs which regurgitate textbooks in question-and-answer format. Finally, we've revolutionized our approach to 'special' education: gifted students have more to work with than they can handle, and remedial students have simple, rewarding tools to help them keep up."

The Apple Education Foundation, established in 1979, is perhaps the major positive private-sector influence on enhancement of learning via microcomputers. The Foundation's grants have exceeded $1.5 million in support of its Charter, which states "The primary goal of the Foundation is to improve the results of the educational process."

Steve Jobs, chairman of Apple Computer, has committed his company and the Foundation to leadership in educational enhancement. This commitment is largely market-driven, as estimates of increased microcomputer installations in schools range from 150,000 to 500,000 per year in just the next three years. The Foundation's objectives and activities treat the existence of microcomputers as a platform on which to build improved basic learning skills and methods, to enhance "learning how to learn."

In 1974, when Cupertino's economy was based largely on grain storage, Allan B. Ellis wrote in *The Use and Misuse of Computers in Education*, "... thinking about the computer's role in education does not mean thinking about computers; it means thinking about education." Thinking about education—and acting on it—is the Foundation's business.

Few companies have evolved as rapidly within the educational-applications market as WICAT Systems, Inc., of Orem, Utah. WICAT was formed by educators, in 1977, as a non-profit research group; its name is an acronym for World Institute for Computer-Assisted Teaching. Founders predicted the impact of microcomputers in education; WICAT was to develop high-quality educational software for the forthcoming technology.

WICAT's efforts were hampered by the quality of its products. In its early years, hardware with capacity to run WICAT software was too expensive. In response, WICAT engineers and managers designed, built, and marketed their *own* hardware, based on the then-new high-capacity Motorola 68000 microprocessor. WICAT, Inc., the non-profit corporation, smoothly became profit-making WICAT Systems, Inc. Its products are designed to complement the technology of 1985 and beyond.

WICAT founder and current board chairman D.H. Heuston states, "Our objective is simply defined: literacy, for every child in the world."

The Jefferson County, Colorado school district is representative of hundreds of communities which have undertaken large-scale microcomputer programs. It is representative as well in that its area, adjacent to Denver, has a high concentration both of university people and advanced technology industry. As the culmination of a 1981 plan, the District spent $2.3 million for 1,695 microcomputers for its 114 schools. Each school will have 15 machines in a computer laboratory of not more than 30 students. 25 manufacturers were evaluated on 41 price and performance criteria. Apple computer was selected.

Among computer hardware manufacturers only Tandy Corporation, of Fort Worth, maintains its own retailing network—the national "Radio Shack" chain, familiar as a source of high-technology equipment long before the advent of personal computers.

"Tandy has a special and useful knowledge of the educational market," says education division director Bill Gattis. "Our National School Bid Department processes requisitions and requests for bids from all schools who approach us, often through our local Radio Shack stores. Thus we know what they buy. And it's interesting to read educators' public comments that simple 'drill-and-practice' programs are inferior to more interactive literacy and needs-application software, because 'drill-and-practice' material is overwhelmingly what they're buying."

Diane LeBold is editor of "Commodore—The Microcomputer Magazine," a bi-monthly published by Commodore Business Machines, the big Pennsylvania microcomputer manufacturer. Commodore claims to have more of its machines in schools worldwide than any other hardware maker. In the magazine's May 1983 special education issue, LeBold describes a central aspect of Commodore's marketing philosophy: "When we talk about computer education there's one important thing we have to remember—there are real live kids at the other end of our talk."

Commodore approaches those "real live kids" in a businesslike fashion, indeed. As Commodore's 1982 sales climbed 63% to over $300 million, its profitability remained at over 13%. The firm's entrepreneurial president, Jack Tramiel, describes its education marketing approach as "unlocking the Ivory Tower." One of its newest and already successful products, the Commodore 64™, sells for less than $250 and is compatible with peripherals and software designed for earlier Commodore models. Commodore's latest software/"courseware" catalog weighs nearly 8 1/2 pounds and runs to over 800 pages.

David Rosenwald, director of education sales, says Commodore is anticipating an education market where overall planning, teacher training, and softwear availability have reached high levels. "To provide computer literacy, two kids per machine is the maximum practical ratio. A school is a business; it can't afford to pay more than it has to, to produce its product. In literacy applications, kids need to learn how the microcomputer works and how to use it."

Commodore has constructed a national network through which educators can communicate with the Company and each other. Hundreds of schools have been established as Commodore Education Resource Centers; they share information among themselves and with their schools, and are aware of materials published by Commodore, including names of clubs or "users' groups" which specialize in Commodore equipment and software. Commodore's dealer organization complements its approach to the education market. Parents and others seeking a single Commodore microcomputer have a range of retailers to choose from. Educators, who require added values like training and system-design advice, may access with specialized dealers such as Fisher Scientific, who are trained and equipped to provide solutions to computer-related learning needs. Such dealers can deliver any configuration of Commodore products, and can provide training for teachers and administrators as well as for the staff responsible for repairing school-owned equipment.

"Computers can, if used correctly and with the right software, enhance the learning of any subject. A good (software) program will offer variety... the computer should teach the student in the same way a teacher would." These might be the words of an academic professional; they are in fact those of Tracey Cullinan, a 15-year-old Los Altos student, one of 20 young people who comprise a Youth Advisory Board commissioned by Atari, Inc. of Sunnyvale, California.

Atari is totally committed to learning enhancement via computer—preferably with its own four models ranging (in price and feature content) from the 600 XL to the recently-unveiled 1450XLD, all with full color and sound built in. The new 1450XLD includes a telecommunications connection and a synthesizer permitting the computer to translate text into its own speech: it can literally talk to its operator.

"As an educational medium—in the classroom or in the home—the computer's potential is like a great dream," says Linda Gordon, senior vice president of the Atari Educational Division. "Our job is turn the dream into reality."

Atari has a two-level strategy to achieve this objective: the first level, focusing on communication, experiment and training, involves The Atari Institute for Educational Action Research; the second level, concerned with product and courseware supply and service, is the responsibility of Gordon's marketing group. Both build on a specific Atari benefit—wide acceptance, among young people, that Atari leads in developing challenging, exciting products and programs.

The Atari Institute was founded in 1981. It promotes the concept of "life-long learning" as a matter of Atari corporate policy. Institute program manager Sandra Williams says, "Literacy as a concept is tied, to an extent, to what teachers learned in school. The traditional definition of literacy is out-of-date. Even geography no longer constrains the spread of the 'new' literacy."

Williams explains, "We have a program called 'Sister School Project,' for example. The project links bigger schools which are technologically experienced with smaller schools just starting up. Teachers—and students—can get help or exchange ideas via telecommunications facilities built-in to Atari equipment." Between "Sister Schools," traditional roles and limits are minimized; a student at one school can help a teacher at another. Teachers are able to reach easily beyond their own classroom walls—even beyond their school districts. The Atari Institute provides partial funding and initial instruction. It then stands back and observes.

Some of the Institute's observations are discussed with youth advisors, corporate managers, and—ultimately—with Atari Educational Group personnel. Ideas are given project status and professionally "followed up." Bob Hall, director of educational division sales, demonstrated a videodisc-based software catalog for teachers. "The disc contains actual excerpts from hundreds of Atari software programs," Hall explained at a demonstration. "Teachers can select titles from an accompanying list, and then personally 'call up' a one-minute preview, just as it would appear on an Atari monitor." The videodisc facility may well aid teachers in helping parents and other community members understand Atari's role in their children's classrooms.

"We're not ashamed to call a lot of what we do 'Edu-tainment,'" says educational division marketing manager Mark McCrackin. "The market tells us that learning can be fun. We think listening to the market makes sense."

Courseware

Educational software, or "courseware," is most tangible in the form of individual cassette tapes and plastic discs, and accessible material like PLATO, containing the thoughts, information and graphics implanted by authors, artists and educators. The talents and skills of these people, packaged and produced and marketed by software "publishers," defines the influence of courseware. Of course students can, and are frequently anxious to, create and "run" courseware or other software programs they have designed themselves.

The Children's Computer Workshop, according to director of research

Leona Schauble, "creates educational media that are irresistibly fun." Programs available from CCW include: "Growbots" for pre-teens. This creates an agricultural environment which a child tries to cultivate (factors such as climate and pests are built-in). "Timebound," which involves science and chronology in a quest to "trap" a mad inventor and "Bagasaurus" simulating word-choice to improve dictionary/thesaurus skills.

"Programs are created by a team of professionals," explains Ms. Schauble. "On the team are programmers, for the technical side; artists and designers; and education specialists and psychologists. Our producers, all from education, mediate among the team. Finished products are evaluated by our eleven full-time researchers, and by about 150 children."

Few companies have the tradition and experience in understanding "family values" as The Reader's Digest Association does. Family values are subjective and difficult to quantify. By any definition they include high quality education, and Reader's Digest products (the famous monthly magazine, plus books and records) are generally perceived as constructive and useful.

Reader's Digest—through its new Microcomputer Software Division in Pleasantville, New York—has already developed programs for school and home markets. Some, like "The Chambers of Vocab™", use traditional game devices like mazes to reinforce children's interests while encouraging them to build language skills.

"There's always been a sense that tedium and learning go hand-in-hand—that young people who enjoy the education process are a bit strange," says Reader's Digest division director Richard Scott. "That's a dangerous thing to say. But our experience has *proved* that people like the occasional logic puzzles and 'vocabulary-builder' features in our magazine. And if any sort of computer-oriented education practice is to extend into the home, it will have to compete for the student's leisure time—with what the child considers as fun."

Program development manager Ellen Smith agrees. "As a parent, I've learned to deal with the real world," she says. "My 20 years of experience have convinced me that you have to sell the benefits of education—not to yourself or other adults, but to the children.

"We have a new product called 'Micro Habitats™' in our Video Construction Set Series. You can see where we're going just from the title: high-resolution color-background environments, including underwater, jungle, and outer space. Pre-animated shapes—sharks, lobsters, even a submarine—may be moved and placed by the child. Each has a sound effect, and as environments are created, the sounds build and mix.

"'Habitats' won't influence SAT scores the way our four disk, 30-workbook 'Problem Solving Strategies™' courseware will. But critical thinking is the central theme of both products, and it can happen anywhere."

Sunburst Communications, a software publisher also located in Pleasantville, markets "The Factory," a program noted as exemplary in several education trade publications. This ingenious program requires the user (grade levels, like many truly interactive programs, are from 4 to Adult) to correctly place, rotate and perform various operations in order: test a machine, build a factory, then make a product. Educational objectives are to develop inductive thinking, visual discrimination, and spatial perception, and to understand sequence, logic, and efficiency.

No one knows the number of independent publishers/producers of educational courseware microcomputer programs. Neither Atari nor Apple executives in education marketing know the number of programs available for the products they manufacture. These facts *are* known: there are a great many publishers and a great many programs; there soon will be a great many more; finally, a few highly-motivated publishers of high-quality programs have been identified, and are achieving rapid and substantial acclaim.

"Pinball Construction Set" is published by Electronic Arts of San Mateo. It permits players to create their own pinball games, and to control scores, sounds, colors, and weight and speed of the "ball," on either Apple or Atari computers. Beyond its entertainment value, the program permits practice in operations common to many microcomputer activities.

Two other Electronic Arts products, "M.U.L.E." and "Seven Cities of Gold," combine the features of games and simulations. Both programs were authored by an Ozark Mountain-based team comprised of brothers Dan and Bill Bunten, Alan Watson, and Jim Rushing. "M.U.L.E." alone represents 3500 man-hours of writing and experimentation. It is a game, whose object is to amass money, land, and goods by various means. Within the games format are clear, practical examples of such economics/mathematics basics as supply and demand, economies of scale, the learning curve theory of production, the law of diminishing returns, and the classic prisoner's

Readings

The Computer Primer is a 481-page *handwritten* book by Ann Cavanaugh, a Connecticut teacher, and published by Trillium Press of New York. Ms. Cavanaugh's handwriting is excellent: "The first aids to calculation were fingers.... A caveman might have used three fingers to describe the three deer that got away on his hunting trip. The word 'digits' is derived from the Latin word 'digitis' which means fingers." The book ($12.50) is used with students eight years old and up. "It's a kids' book," says Trillium editor Tom Kemnitz, "but..."

The Apple Guide To Personal Computers in Education compresses much information into 48 pages (including over seven pages of color artwork by Don Weller). $1.95.

I Speak Basic To My Apple (or Atari or TRS-80 or Commodore PET or other microcomputer; book is "system-specific") comes from Hayden Book Company of Hasbrouck Heights, N.J.

Cost of a classroom course, with materials for teacher and 20 students: about $160.


3M
diskettes
Box of 10 8 inch
DS,DD,
1024/F
double side
double density
1024 byte format
3M
diskettes
Scotch
DC 300XL data cartridge
Unforgettable.

dilemma.

Electronic Art's president Trip Hawkins, a former Apple executive, says his products are designed to provide "a new category of experiences. We have new artists at work in a new medium. The hardware, with its elaborate hidden machinery, is the scaffold on which we work. The products we create must be simple (to operate), 'hot' (in the McLuhan context), and 'deep' (highly interactive). Whether we fit into a classroom environment is moot; the learners will find us."

Ann Piestrup is chairman of The Learning Company, an independent courseware publisher in Portola Valley, California. Even though Piestrup's firm is just four years old (like many of the children who derive benefit from its products), it has established a national reputation for high quality. One of The Learning Company's programs, "Rocky's Boots," permits children to begin understanding both the engineering basis of computers and the manner in which they are operated. Briefly, the program displays simplified circuit-components which children learn to move and connect. Using color, sound, and gamelike scoring features, operable "systems" may be designed. Success and achievement are rewarded by an animated "Rocky Raccoon" character who appears and dances on the computer monitor screen. The *New York Times* wrote last summer, "... if you have an Apple and kids, not getting 'Rocky's Boots' is like having a toaster and not buying bread."

One phenomenon of computer-assisted education is the privacy of the relationship between a child and a truly interactive, success-directed courseware program like "Rocky's Boots." Even the most caring and sensitive parent can barely gauge the involvement, fascination, and delight obtained by a child, as high-quality courseware operates. As more high quality courseware develops, and children become "literate" in a way never before experienced, parents often will find themselves distanced from their children's learning experiences. It will cause a new and constructive "coming of age" for both generations.

College and Beyond

Most schools employing microcomputers at the K-12 grade levels use them in tutorial projects, in gifted-child programs, and for organized classes in which the computer is primarily the object, rather than the medium, of instruction. At the college/university level microcomputers and/or computer terminals are much more integrated into the learning curriculum. Technical and engineering colleges/universities have wide-ranging computer facilities and extensive student demand for their use; several colleges, including Stevens Institute of Technology, Drexel University and Carnegie-Mellon University require microcomputer ownership of all entering freshmen.

Digital Equipment Corporation of Maynard, Massachusetts is the largest maker of minicomputers. DEC education systems are well-suited to central networking. "I think the school districts will eventually make it through host computers," says Bob Trocchi, product group manager for education, "with a number of computers tied into a central unit, such as a DEC PDP-11. It's just not possible to manage more than two or three micros in a classroom," he says. "You would have to deal with dozens of floppy disks and exuberant youngsters."

DEC has started an interesting innovation at Indiana University. For a small fee, students in some dormitories may use DEC-supplied stand-alone computers for word processing. "It has been most successful," says Trocchi. "Dorms without the machines are complaining."

Data General Corporation of Westboro, Massachusetts, was founded in 1968. Its 1982 sales of over $800 million ranked it 339th on last year's *Fortune* 500. Along with Data General's sales growth has come a growing involvement with education on three main levels: first, the company's need for more qualified people to augment its own staff; second, its development of the new Desktop Generation™ computer line for broad-reach school-system administrative use and interconnection to its own and IBM compatible systems; third, its entry into the professional-level small computer market with products designed in part for student use.

Planning Data General's approach to the education marketplace is the responsibility of Peter Jessel (Ph.D., MIT; M.B.A., Wharton). "We respond to what we see," he says. "For years our products have had strong application in university-level situations. But now we see the same kind of work being done by freshmen—even by high-schoolers—that was once put only before upperclassmen. The computer learning process is migrating downward, in terms of student age."

Data General has, for two years, been developing a distributed data processing system potentially serving school districts in New York State, under the auspices of the Southern Tier Regional Computer Center. The Center is located at the Broome-Delaware-Tioga BOCES (Board of Cooperative Educational Services) in Binghamton. The system permits wide use of the IBM 4300 mainframe computer shared by all the schools for administrative, instructional, and research services. Students' access to scores of installed computers and terminals brings them high-quality programming software, plus the databases and other resources created over the past several years and restricted hitherto to the mainframe system.

The Center has proposed to Data General that an expanded system be created as a controlled model, for possible application among similar large Centers across the country.

"When we first examined the potential of the elementary/high school market for Data General products, back in the seventies, we found confusion," says manager of industry marketing Bob Orr.

"Now we are beginning to see sophistication: realistic, consistent objectives coming from education professionals who are planning-wise and product-wise. We can employ consultative selling and be productive as long term business partners."

Future Quality

Dorothy W. Blake, coordinator of planning for Media Resources and Utilization for the Atlanta Public Schools, spoke before the House Committee on Education and Labor in January of last year. Here are selections from her remarks:

"... I speak from the experience of a mother, a grandmother, the wife of a school principal, the wife of a former science teacher, and, after three decades in the school business, I consider everybody who works as a part of the instructional team as those who have answered a very high calling...

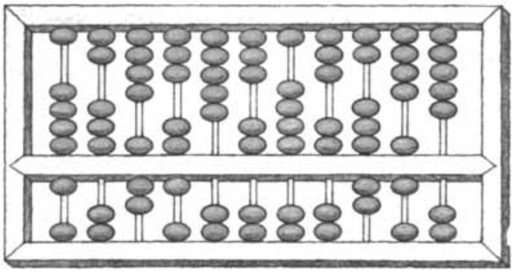
"Elementary and secondary school students as well as their teachers must learn to use newer technologies, such as microcomputers and/or computer terminals, not only to solve mathematics and science problems, but also to gain further access to information needed for their education and for their work and for daily living.

"The machines, and the young people, are here to stay. They match each other in their high quality and great potential."

This special report on "The Computers In Education" was written by Peter J. Brennan and produced by Development Counsellors International Ltd. of New York. Original artwork/graphics: Sherin & Matejka, Inc.

NOTES ON THE ORIGINS AND CAUSES

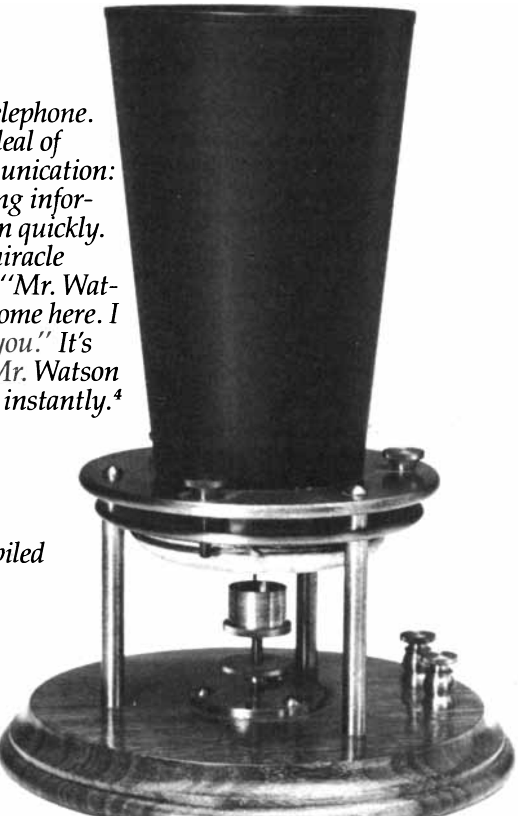
Paper. First century AD, China. Information becomes a more abundant resource. But now has paper itself become too abundant? The plague of the 20th century office: Clutter.[1]



The abacus. Early hardware. Until recently it was, in the hands of a skilled operator, actually faster than a computer. Note: "Skilled operator."[2]

The keyboard. Next to shorthand, the fastest writing system known to man. Question: If you could transcribe your thoughts faster, would you think faster? Had Cervantes the benefit of even average typing skills, would he have written another book?[5]

The telephone. The ideal of communication: Sharing information quickly. The miracle isn't: "Mr. Watson, come here. I need you." It's that Mr. Watson came, instantly.[4]



The encyclopaedia. By definition, all current and essential information compiled and made accessible to the non-specialist. A fourteen-volume learning tool.[6]

[1]It is to be noted, in general, that Digital Equipment Corporation's computer systems place particular stress on the elimination of many routine administrative chores involving paper, such as student registration, class scheduling, budget accounting.

[2]As recently as ten years ago it was the consensus that there would gradually come into being a sort of "computer priest class," who alone would consult these electronic oracles. Today, thanks to accessible systems–Digital's VAX, in particular–computer literacy is fast becoming accepted as a basic skill, like reading.

[3]An alternative to the postal system is developing between computer-connected parties: Electronic mail on the VAX system. VAX computers at different branch-campuses routinely exchange and transmit even lengthy written communications. Finding a stamp, a chronically irritating chore, becomes a thing of the past.

[4]Imagine a twist: What if Watson had been out? (Or his phone had been busy!) What if you wanted to ask a colleague in London a question. Would you call him at 1 A.M. his time? Or would you use a terminal to relay your question (via DECmail) to his terminal–guaranteed it would get to him?

[5]Cervantes was sometimes an inattentive writer. Had he the benefit of what any freshman can have today–a computer to "converse with"–his terminal might have flashed him "...Hold it! Two pages ago you said that person's name was spelled..." On the vast potential of student time-sharing, consider that with current VAX technology it is perfectly possible for a motivated student to sit down at a terminal for a minute between classes and simulate the moon's gravitational system.

[6]The encyclopaedia, if taken one more step, becomes computer-based education: An infinitely patient computer conducting a student through a simulated chemistry experiment step by step. Let him *see* the consequences of, say, adding water to sulphuric acid. Obviously, this requires a computer terminal screen with exceptional graphics capability, like Digital's GIGI terminal, and Courseware Authoring System.

[7]A telling comparison: To get out an *annual* col-

Commodore 64 Magic Desk I

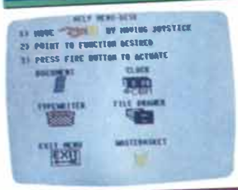
Only From Commodore—The Excitement and Simplicity of Magic Desk!

Only Commodore brings you the magic of MAGIC DESK... the next generation of "user friendly" software! Imagine using your computer to type, file and edit personal letters and papers *without learning any special commands!* All MAGIC DESK commands are PICTURES. Just move the animated hand to the picture of the feature you want to use (like the TYPEWRITER) and you're ready to go.

The MAGIC DESK Typewriter works just like a real ELECTRIC TYPEWRITER ...and it's COMPUTERIZED. All the filing is *electronic*. Excellent sound effects and screen animation make typing fun, whether you're typing letters, reports or memos...and the built-in filing feature makes MAGIC DESK useful for keeping names and addresses, home inventory lists, insurance information and more.

Your COMMODORE 64, COMMODORE DISK DRIVE and MAGIC DESK are an unbeatable combination. Filing operations are automatically linked to your Commodore disk drive—but you don't have to know any commands—just "file" the pages you type in the file cabinet and your text is automatically saved on diskette. There are 3 file drawers with 10 file folders in each drawer and 10 pages in each folder.

To PRINT a page you've typed, just "point" at the picture of the printer and your pages are automatically printed on your COMMODORE PRINTER or PRINTER/PLOTTER. If you want to erase what you've typed, the WASTE-BASKET under the desk lets you "throw away" pages. There's even a DIGITAL CLOCK which helps you keep track of time while you're typing.



Not only is MAGIC DESK easy to use... it's hard to make a mistake! Just press the COMMODORE key and one of several "help menus" appears to tell you exactly what to do next. Special messages show you how the various picture commands work and help you when you make a mistake. Help messages also show you how to use the printer, filing cabinet, digital clock and wastebasket.

Experience The Magic At Your Local Dealer.

commodore COMPUTERS

First In Quality Software

BOOKS

A paper orrery, a sumptuous atlas of the world, microelectronic visions, nuclear Armageddon

by Philip Morrison

OUR SOLAR SYSTEM CALENDAR OF 1984, by John Gregory Hanses. Box 30005, Winona, Minn. 55987 (poster with pins, $5.95 postpaid). A thousand golden guineas was what that "worthy and ingenious artificer Mr. John Rowley" of London was paid in about 1712 for his new, beautifully decorated clockwork instrument. It was a princely sum; the desirable new second edition of Sir Isaac Newton's *Principia*, a thick quarto volume, nicely bound, came out that year at just one guinea. The noble commissioner of that device, Charles Boyle, the fourth earl of Orrery in Ireland, grandnephew of the Sceptical Chymist Robert Boyle, nonetheless got more than he could have expected. Ever since then such an instrument has been called an orrery, as Rowley had named his first example in gratitude to his patron.

What the earl bought was an elegant model of the solar system driven by a hand crank, with a clever gear train that moved an ivory earth with its well-paced moon around a central brass sun, all the model objects turning properly on their axes. The idea was not truly a new one; the device goes back to Ole Roemer and Christiaan Huygens. But orreries grew grander throughout the century, particularly in London, and examples of these Enlightenment splendors, gilded Newtonian systems of the world crowded with planets and their moons, grace many a museum.

Rowley's Mark I is now to be seen in South Kensington; the Science Museum bought it from the Boyle family only a decade ago for 30 times its original price. The name orrery is now rather loosely given to any such Copernican model of the planetary system, the more since we ordinarily stretch the apter word planetarium to describe a very different kind of geocentric demonstration theater.

Clockwork orreries are still made by a few clever artificers, and they are still pricey. The solar-system calendar listed here, however, is actually a bargain orrery done in paper, the planetary motions as they would be viewed by an observer far to the north of the disk of the planets, stationary with respect to the sun. Nearer software than hardware, the model orbit segments the planets will traverse in 1984 are mapped in white on a night-black poster sheet of heavy stock, meant to be mounted on a wall or a tabletop. The sheet needs a backing of some stiff material, say corrugated cardboard, that can accept the colored map pins supplied. The orbits are drawn in dots around a big yellow pin for the central sun, each dotted path carefully labeled with the dates appropriate to successive positions along the orbit. The paths of the four inner planets are more or less to scale (Venus and Mercury require winding spirals, to allow more than one circuit in an earth year), but the sluggish outer planets must be crowded in to fit on the large page. Neptune would be eight feet out if its orbit size were scaled to the hand-size orbit of the earth.

Instead of clockwork it is the user who must place and advance the planet pins in the stately year-long ritual celebrating the rhythms of reality. The reward is clear: easy recognition of the planets in the sky. The pins once in place map directions in the solar system accurately enough for knowing planet watching, supplying the sense of giddy interplanetary depths and the satisfying feeling of realizing just how one need look out to find the planets glowing at dawn, at dusk or at night, from one spinning planet among the others. Imaginative users might be able to put themselves as stargazers on Mars. Our moon is a much needed feature of this paper orrery; every full and new moon is marked along the earth's orbit, and a small white pin can be fastened to the earth pin in such a way as to show the changing direction of the moon as well as the directions of the sun and planets. (Earth spin is provided only mentally, cued by the evident direction of sunrise and sunset for an observer who places himself properly on the pinhead.)

Utter beginners will perhaps prefer an ordinary geocentric sky calendar, a month-by-month display to match our earthbound view directly. Watchers with a little experience who take pleasure in the true celestial geometry of three dimensions, however, will find the paper orrery a steady delight. The solar system, flat disk though it is, is no paper-thin plane; here that convention means the eclipses and the close planetary approaches are not fully modeled. This is not the best year for such little spectacles; perhaps the end of January will offer about as admirable a jeweled pair as at any time this year, bright Venus and Jupiter cheek by jowl in the predawn sky, although distant along the line of sight. There are other data on the chart, such as the star backgrounds around the ecliptic. Particularly useful are times of new moon given to the minute, indispensable for playing the game of sighting the shy young crescent. We miss a marker for the moon's nodes, fixing the eclipse seasons.

Hanses is a lifelong amateur astronomer. After a varied career in education he is now teaching young people and their parents as a one-man traveling circus of astronomy, whose spirit is plain enough from this excellent publication. He renders your sun sign, for instance, not as a mere magical affix of your birth date but as a reality to be observed. There is evidence that he is conspicuous in downtown Winona at the Christmas season, a Santa whose warmth is no more feigned than his honest beard or his orbits.

THE NEW INTERNATIONAL ATLAS. Rand McNally & Company ($100). In 1969 the forerunner of this atlas was reviewed here. In a 1982 edition it remains the best modern world atlas of reference; several versions appeared earlier, the first few without "new" in the title. Changes are to be seen, although they are not major ones. The eight or 10 maps added are mostly western European regions not shown earlier at the one-millionth scale, from Scotland to Sicily and over to Wien. The maps of 60 cities around the world at uniform scale are somewhat updated. Most conspicuously parks and airports dwindle or appear, and suburbs earn larger typefaces as the cities spread; there are many new roadways, one out to Fire Island and one behind a long barrier beach south of Rio. But no new cities are mapped. A new 30-page section of topical maps is here; its source is the *Encyclopaedia Britannica*. Those maps are a medley: historical, political claims, economic data, climate, languages, transportation and the like. The most important of these topics were already mapped in the geographical lead essay of the 1969 edition, which is now dropped. Population estimates for 1982 are listed for all countries, for urban centers of more than 50,000 and for well-known towns.

Comprehensive maps are the heart of every atlas; those here are outstanding, with their readable and convincing shaded relief for landforms. It is the intensive application of that technique, a modern improvement in cartographic art, that singles this atlas out from all others. The information content of the work is no less praiseworthy. Since 1969 the main maps have been updated; Zimbabwe and Belize are clear, along with Thanh-pho Ho Chi Minh (once Saigon), the Aéroport Charles de Gaulle and the Tanzan railway. The chief map changes since 1969 are interestingly listed. The lengthy index is here, reset on new paper stock, and the few awkward gatefolds have been omitted, even if a map page has to be repeated. The polyglot instructions that justified the title (cartographic publishers from a number of countries shared the work from the first) now appear in a fifth tongue, Portuguese. One can recommend this atlas with enthusiasm to any home or library that can hope for such a luxury, but the 1969 version remains entirely valid. Inflation in price is the biggest change of all.

STATE OF THE ART: A PHOTOGRAPHIC HISTORY OF THE INTEGRATED CIRCUIT, by Stan Augarten. Foreword by Ray Bradbury. Ticknor & Fields, 52 Vanderbilt Avenue, New York, N.Y. 10017 ($19.95). The internal-combustion engine, the mushroom cloud and the integrated circuit span this century of technological novelty. The three dozen photographs of such circuits presented here chronologically, each with a swift, chatty, well-informed one-page gloss, amount to an introductory history of this remarkable technology, closer in some ways to a second strain of life than to any other artifact of our kind. Most of the photographs are at some fortyfold magnification, an atlas of stained tissues of these new organisms. The couple of dozen pseudo-neural networks we see resemble meticulous maps of well-ordered cities in their variety of geometric texture, here parks, there tenements, there rail yards. Those images go back only to 1970. The work of the 1960's showed simpler textures, still mere surface designs under tenfold magnification, the patterns more like Persian rugs or fine old wallpapers. Only the first few archetypes look three-dimensional, the wires and droplets of wax or solder visible as works of the hand, the entire little structure arrayed in space. Jack Kilby's "first bona fide IC," from 1958, evokes nothing so much as a microscopic Dagwood sandwich.

Today's crisp planar world began in 1959; a single early planar transistor is shown as it was made at Fairchild by Jean Hoerni and Robert Noyce, who together grasped how to diffuse the needed impurities into the silicon single-crystal base to make the functional transistor interface, how to insulate it with a silicon dioxide layer and then how to furnish it with leads of evaporated aluminum. Although the device is flat enough, it is really elaborately multilayered. The rest of the story is shown by example; year by year the control of feature dimension and hence of the possibility of viable intricacy grows. The random-access memories expand from 256 bits to 1K and now, a decade later, to 288K. Here are interesting sidelines as well, such as the first eight-bit charge-coupled device, connected capacitors instead of transistors, its 64K-bit kin now masters of light sensing from the checkout counter of the supermarket to the prime-focus cage of the 200-inch telescope on Palomar.

There are microprocessors famed in the marketplace, from the Zilog Z-80 of 1976, a version of it in every Pac-Man, to the Motorola 68000 in the older Radio Shack computers and a 1981 32-bit Hewlett-Packard microprocessor, wired with tungsten instead of aluminum, the equivalent of 450,000 transistors on a square a quarter of an inch on a side. Speech synthesizers, analogue-digital converters and erasable memories are successive surprises sprung by companies large and small. Materials and devices change now and again, a faster response time begins to appear, architectural changes begin to divide the functions, but the chief overall effect is ever smaller features and hence ever more complex products at manageable yields and lower prices. The photographs are all-American, although the microscope used is commonly a Zeiss. Indeed, the images are somewhat deceptive: they are colorful, often beautiful. What one sees is the product of a differential-interference optical technique, with the illuminating light itself usually a strong color. Without such optical staining one would see mainly a dullish aluminum surface.

There is a useful introduction at a simple technical level. It tells of logic gates and junction transistors, it sketches the forming of a typical IC layer by layer and it includes a block diagram of a microprocessor broken down into a dozen functional segments. One analogy of the author's is worth citing: "Computers are composed of nothing more than logic gates stretched out to the horizon in a vast numerical irrigation system." Still, the best of the text is in the glosses for the photographs, where even the knowing are likely to learn something intimate about each product, its origins, its designers, its impact, its fate.

The Zilog page is a good example: "Zilog's founding is a typical Silicon Valley story. A group of entrepreneurially minded engineers quit their jobs." Those jobs were at Intel. Intel was established by three men who had left Fairchild, which was started by 12 engineers who came from Shockley Transistor, which was founded by William Shockley, who had left for the golden West from the three-man group who at Bell Laboratories had made the first transistor. The gallium arsenide chips and Josephson junction ultrafast cryogenic switches shown here in a foretaste of the future come out of no upstart shop but from the giants Rockwell and IBM. There may yet be another two or three orders of magnitude ahead in usable chip complexity; the human brain seems five orders beyond that, but it is soaking wet and must run very slowly at low voltage.

The author is a technical journalist from Silicon Valley itself. He knew the old valley in the yeasty pioneer days, say six or eight years back. He has turned visual historian with flair.

A HISTORY OF STRATEGIC BOMBING, by Lee Kennett. Charles Scribner's Sons ($15.95). NUCLEAR HOSTAGES, by Bernard J. O'Keefe. Houghton Mifflin Company ($14.95). THE WIZARDS OF ARMAGEDDON, by Fred Kaplan. Simon and Schuster ($18.95). The elaboration of myth seems an indispensable part of human response to the challenge of novelty. The word is not used here in the sense of the mythical, an antonym to the real; rather it implies the domain of the decisive metaphor, within which the memorable, the significant and the personal can form cogent symbols. Simple matter-of-factness is often present, but it becomes secondary.

Such a myth has grown up around the quandary set for us by the coming of nuclear warfare in 1945. A few phrases evoke the rich and familiar tale: Einstein's letter to FDR, the Faustian scientists and the rude general who seduces them, the Italian navigator into the New World, the mushroom cloud in the desert, the agony of the Japanese cities and the terrible aftermath of radiation. Our fear is patently not that of prop-driven air war over Japan; although grim, it was endured, even through two nuclear explosions. The danger today is in contrast only potential, but it is catastrophic. Not novel in kind, it is unprecedented in scale: nuclear warheads are ready not in ones and twos but in the tens of thousands.

These three quite different books, all of the year 1983, offer material for the construction of needed myth, not of the sowing of nuclear war but of its distant origins and of its ripening to a deadly

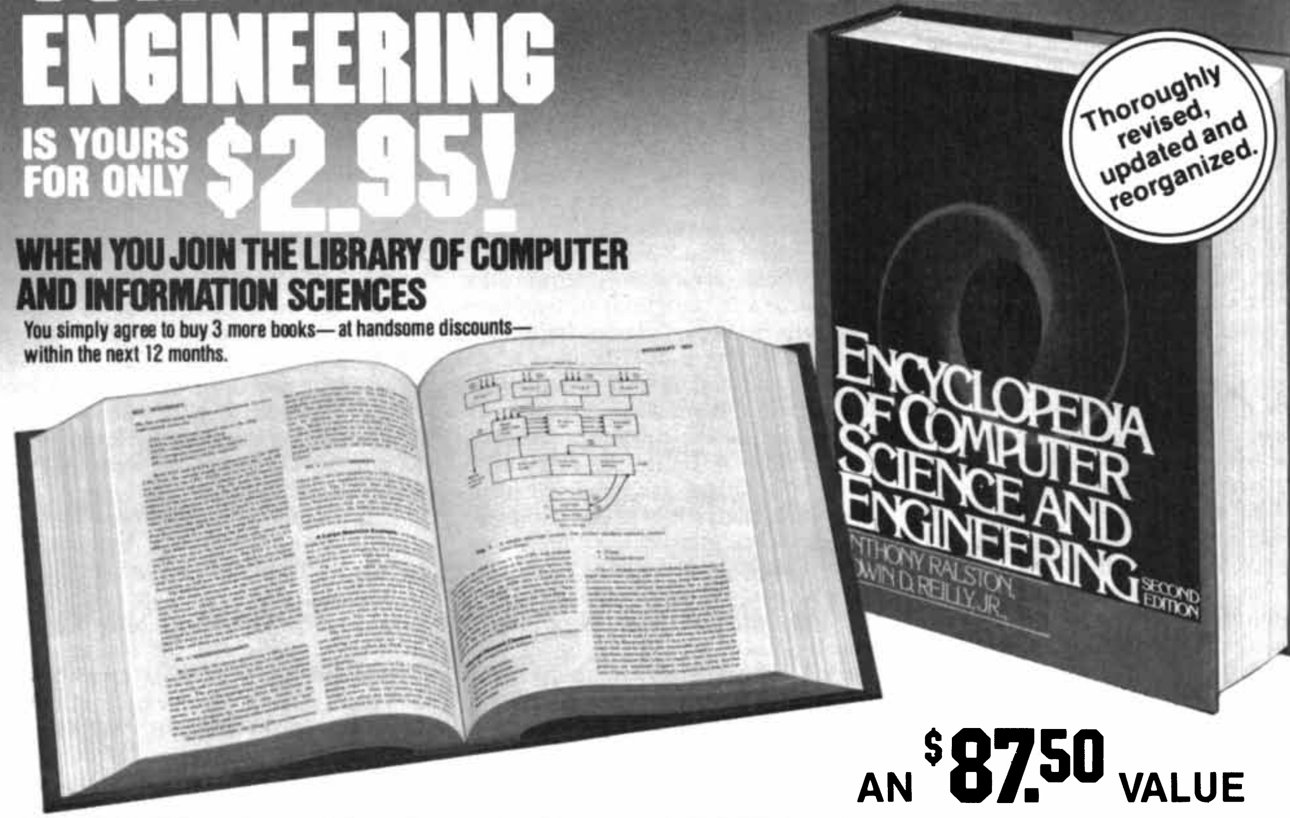
THE MOST COMPLETE COMPUTER INFORMATION RESOURCE

THE NEW ENCYCLOPEDIA OF COMPUTER SCIENCE AND ENGINEERING

IS YOURS FOR ONLY $2.95!

WHEN YOU JOIN THE LIBRARY OF COMPUTER AND INFORMATION SCIENCES

You simply agree to buy 3 more books—at handsome discounts—within the next 12 months.



AN $87.50 VALUE

This up-to-date Second Edition contains:

- **Over 1,670 pages of revised, expanded, and updated text**
- **550 articles on virtually every aspect of the computer sciences**
- **Written by 301 distinguished authorities**
- **Profusely illustrated with over 500 photos, and over 250 diagrams, graphs and charts**

Indispensable to consultants, business people, data processing professionals, and enthusiasts, the ENCYCLOPEDIA OF COMPUTER SCIENCE AND ENGINEERING is a veritable data base of information on:

- Hardware
- Software
- Programming languages
- Artificial intelligence
- Computer applications
- Personal computing

and much more!

The Encyclopedia is organized to make finding and using its wealth of information an ease. Articles are alphabetically arranged and are cross-referenced to related articles and to specific subject matter. The clear and expanded appendices include abbreviations, acronyms, special notation and terminology, as well as numerical tables, the mainstay of applied technologies. A complete 5,000-term index contains references to sub-categories, doubles as a computer science dictionary, and is an invaluable tool for locating specific information.

Praise For the First Edition:
Called "Impressive ... comprehensive ... well done" by *Datamation*, and "... a real treasure cache" by *Business Management*, the new Second Edition promises to eclipse *Computer Management's* statement on the original of "There isn't another book like it." Send for your *free* 10-day trial.

The Library of Computer and Information Sciences is the oldest and largest book club especially designed for the computer professional. In the incredibly fast-moving world of data processing, where up-to-date knowledge is essential, we make it easy for you to keep totally informed on all areas of the information sciences. In addition, books are offered at discounts up to 30% off the publishers' prices. Begin enjoying the club's benefits today!

4 Good Reasons to Join

1. The Finest Books. Of the hundreds of books submitted to us each year, only the very finest are selected and offered. Moreover, our books are always of equal quality to publishers' editions, *never* economy editions.

2. Big Savings. In addition to getting the ENCYCLOPEDIA OF COMPUTER SCIENCE AND ENGINEERING for $2.95 when you join, you keep saving substantially—up to 30% and occasionally even more. (For example, your total savings as a trial member—including this introductory offer—can easily be over 50%. That's like getting every other book free!)

3. Bonus Books. Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings up to 70% off the publishers' prices.

4. Convenient Service. At 3-4 week intervals (16 times per year) you will receive the Book Club News, describing the Main Selection and Alternate Selections, together with a dated reply card. If you want the Main Selection do nothing and it will be sent to you automatically. If you prefer another selection, or no book at all, simply indicate your choice on the card, and return it by the date specified. You will have at least 10 days to decide. If, because of late mail delivery of the News, you should receive a book you do not want, we guarantee return postage.

If the reply card has been removed, please write to:

The Library of Computer and Information Sciences
Dept. 7-CD9 Riverside, N.J. 08075 to obtain membership information and an application.

Scientific American 1/84

harvest still unreaped. Professor Kennett's book is the coolest; its detachment is one of time. He is a military historian who has singled out of the two centuries of human flight the story of its use for "direct attack on the enemy state with the object of depriving it of the means or the will to continue the war." His is a concise treatment for the general reader, scrupulous and broadly seen, with the detail essential to support the chronicle but without technicalities. It treats mainly of the two world wars of our century; only a prelude can be found in war before 1914, and the book closes with the events of 1945.

The root of air bombardment of cities are found, Kennett says, in a half-absurd campaign against Venice, Queen of the Adriatic. In the summer of 1849 Austrian armies laid siege to the city. Their guns could not reach their targets from either land or sea; the lagoon was too wide and too shallow. Some 200 small hot-air balloons were released to drop cast-iron bombs filled with gunpowder as the balloons drifted on well-chosen winds over the rooftops along the canals. The little experiment failed; the defenders beset by hunger and cholera hardly noticed the effort. Perhaps one bomb did explode on the Lido; the others fell into the waters. In 1852 one eyewitness wrote: "Many discoveries which we laugh at as childish and fantastic later vindicate themselves."

World War I saw casualties caused by war from the air, only thousands out of the millions who died in the war. But the military prophets of air power emerged then, after the stick-and-wire bombers had generated public fear. World War II was not the same. The bombers had grown powerful and in the end ruthless. "None of the belligerents began the war with plans for an air assault against civilian populations—and this includes Nazi Germany. They were in fact all anxious to avoid such an air war." They were all uncertain of the response of their own citizens to a rain of fire, explosive and poisons. They nonetheless found uses for bombs against other targets, and step by step "those who controlled the bombers gradually unleashed them on the enemy population." Tens of millions of civilians lost their homes, and the better part of a million their lives, to the attack of "independent air power." Few voices protested anywhere. The historian, still detached, nonetheless senses the growth of "a certain uneasiness."

There is sufficient reason. Bernard O'Keefe, then a young engineer as a naval ensign, was a participant at Los Alamos, with a responsible part in the implosion bomb, all the way past takeoff to a visit to Japan. Then he left for the campus, very briefly. By 1947 his new company had a laboratory in an old garage in Boston; he was off-campus for good, preparing specialized instrumentation for atomic tests. He recounts directly and simply his own experiences in test after postwar test as the bombs grew. He attended the strange test on Frenchman's Flat in 1953, when the army "wanted to do everything themselves" and fired a 17-kiloton warhead out of the "enormous piece of machinery" that is a 280-millimeter gun. The 25 officers present simply made ready to watch from a small trench. One pulled the lanyard on the great gun, and they all jumped into the trench. Six miles away the warhead duly and blindingly exploded into a mushroom high in the air. "I was aghast." No tactical explosion has been set off in the open air since.

O'Keefe tells something of many other tests he has seen, but none holds more drama than Bravo of Operation Castle in 1954, the first test of a deliverable megaton thermonuclear weapon. He was one of the firing party, on an island in the atoll of Bikini 20 miles from the shot, within a sand-covered heavy concrete bunker. It was too close: the ground shock rocked the bunker, the power failed and the heavy fallout held them prisoner in the dimly lighted shelter. The helicopters found it safe to come for them after a few hours; their hurried eight-minute drive to the copter pad subjected them to only a tolerable dose of radiation. The ships fled from the contaminated lagoon at once, but the tuna fishermen 75 miles downwind were less lucky.

O'Keefe is now chief executive officer of the big technical firm that grew out of his little instrumentation business. He is no reformer but a man of experience and good sense. For him the nuclear-arms race is over, only "pushings and haulings" can remain. We and the Russians alike are held in mutual bondage. He sees slow reconciliation as the way out, first economic, then political. More weapons will not suffice, nor will defenses. Not every reader need be persuaded by his prescriptions, but his is warm and eloquent witness to the reality of weapons into the 1960's, before the tests burrowed underground.

Fred Kaplan is a young political analyst out of M.I.T. His book is oral history, current in the 1980's, the substance of more than 150 interviews with the principals of his story, supported by a raft of documents in the open or newly declassified archives. The sparseness of direct quotations from the interviews somewhat weakens the impact of the study. There are no nuclear physicists at work, no Fermis or Tellers. The theater is not the cable-strewn laboratory or the distant test site or mission control; it is the committee room and the private office, sometimes in the Pentagon, more often in Santa Monica, Omaha or McLean in Virginia. The issues are not timing circuits and excess neutrons but when, where and how to use, or indeed to withhold, what type of nuclear weapon and how many.

The story unfolds chronologically, from the days of 1945, or a little before, to a climax in about 1968, after which a kind of coda ends in the past two years. More than half of the text concerns itself with the men and women of the Rand Corporation, the famous think tank formed at Douglas Aircraft to study air war under a research grant of 10 million unspent wartime dollars direct from General Hap Arnold. Rand now dwells in two buildings near the beach in Santa Monica. That focus remains tight throughout the coldest years of cold war, from the start in about 1946 to the days of John Kennedy, when the venue shifts closer to the Secretary of Defense and when systems analysis, often still done by Rand alumni, now whiz kids, became a part of Pentagon power at first hand.

It is a story of tides. From 1946 to nearly 1960 the tide rises to flood. The targets of the bombs, bigger and bigger bombs on faster planes, are the population centers. General LeMay will deliver any bombs with the Strategic Air Command; the tone was apocalyptic, and the style of it became plain to us all, come Dr. Strangelove. In 1960 the tide ebbs. Russian test missiles fly freely, and thermonuclear weapons are being built not only to the Teller-Ulam design but also to that of Dr. Sakharov.

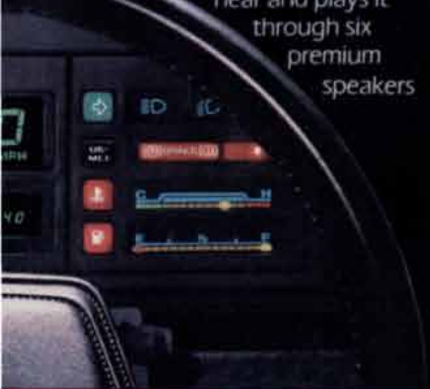
Still the SAC plans grow baroque. A city the size of Hiroshima was targeted with a 4.5-megaton weapon, with three 1.1-megaton weapons as a backup. SAC General Thomas Power displayed a graph showing that tens of millions of Chinese would be killed by fallout from the heavy U.S. attack on the U.S.S.R. foreseen in the Single Integrated Operational Plan 62, with no weapons dropped on China. At the final endorsement before the Secretary of Defense in the last Eisenhower month, Marine General David Shoup could not sit still: "Sir, any plan that kills millions of Chinese when it isn't even their war is not a good plan. This is not the American way." (The cited source is handwritten notes given to Kaplan by a participant, plus interviews.)

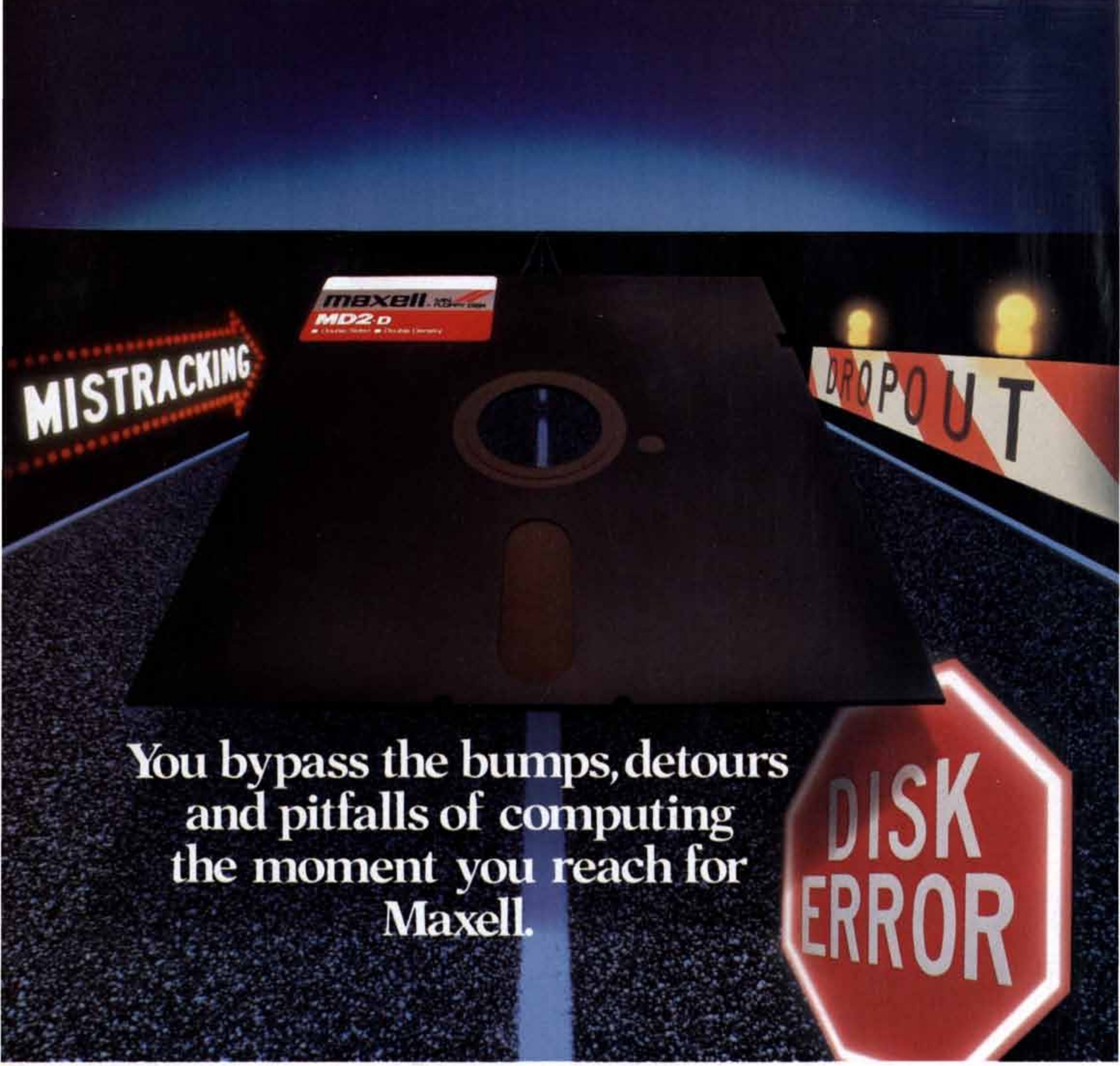
A new doctrine was worked out from earlier gropings, largely by another Rand veteran, William Kaufmann, now a political scientist at M.I.T. It is too dangerous to retaliate massively against the U.S.S.R., say for a mere attack on NATO Europe. There is another road; after all, the artilleryman may shoot some of his shells at the opposing batteries, not at the troops. Counterforce could punish a Russian attack condignly by inflicting heavy but solely military losses. The Russians would not themselves open fire against our cities to retaliate: it would be too dangerous. They might pick off the rest of SAC, but in the end, outgunned by our larger forces, they would quit. The dead in both countries need be only some millions, not a hundred million or two.

By 1964 the ebb tide too had encountered its difficulties. Counterforce did not even, as the whiz kids hoped, slake the thirst for a larger arsenal; there was now no limit to the number of targets foreseen, even if they called for smaller warheads on more accurate missiles. Moreover, the doctrine assumed a reasoning adversary, and it suggested another race for advantage. As Russian strength grew it could become their side that would levy blackmail against our own fears of city attack.

Counterforce was saved for the 1970's by the new powers of MIRV, since shared by the U.S.S.R. Now the launchers do not much exceed the McNamara numbers, but there are more warheads than anything the Joint Chiefs might once have asked for. The creed of Santa Monica has three articles: nuclear war can be planned and executed with some precision; the vulnerability of SAC is the major threat; we need a strong counterforce. By the end of the Carter term, through and around the famous planning document PD-59, "every tenet of the Rand philosophy was set in place as official U.S. policy."

In the Reagan years the doctrine remains, spiced by command attacks. The MX missile is now placed in silos as Peacekeeper, with assertions about the vulnerability of land-based missiles forgotten in the interest of counterforce. Just as in SIOP-62, ours is a policy of preemptive nuclear attack as a deterrent to conventional war. The milieu, however, has changed; although doctrine is now all but canonical, its application to the real world has met with skepticism. So far the taxpayers have ordered only 21 Peacekeepers, a tenth of the Carter plan. Since 1980 the citizens of many countries have come to question these options and their simplistic logic. It looks increasingly certain that the abstract doctrines of nuclear war, little affected by facts, are not so much reasoned as they are rationalized to fit unstated institutional imperatives. A reader completes the three books with the growing conviction that our survival has become an uncertain contest. On one side is the slowly but surely deepening public understanding; on the other is 60 years of the ready ingenuity and evergreen zeal of the United States Air Force.





The road to floppy success is paved with Gold Standards.

Maxell can speed your success in computing. Helping you avoid the traps that can block the way to the information you've stored. After all, our disk has outpaced every other in performance tests. And earned a lifetime warranty.

Consider this: disks travel through a disk drive where heat builds up. And up. Only Maxell designed its protective outer jacket to defy 140°F. So the disk keeps its shape. And keeps your information on track.

How good is Gold? Maxell's the disk that many floppy drive manufacturers trust to put new equipment through its final paces.

And the unique way we pack our oxide particles and bind them together means quality for the long run.

Dropouts? Disk errors? Just pass them by. You're on the Gold Standard.

maxell®

IT'S WORTH IT.

Maxell Corporation of America, 60 Oxford Drive, Moonachie, N.J. 07074 201-440-8020

Established 1845 SCIENTIFIC AMERICAN January 1984 Volume 250 Number 1

Launch under Attack

This strategy calls for launching nuclear weapons on warning that attacking weapons are on their way. Pressures for its adoption are symptomatic of an increasing instability in the nuclear balance

by John Steinbruner

The decision by the Reagan Administration to install 100 MX missiles in existing Minuteman missile silos raises serious questions about the current policy of the U.S. with respect to its strategic nuclear forces. The theoretical vulnerability of the Minuteman force to a preemptive strike by increasingly accurate Russian intercontinental ballistic missiles (ICBM's) had long been advanced as the primary justification for proceeding with the MX program. After an exhaustive review of more than 30 alternative basing modes for the MX, however, the Administration, acting on the advice of the President's Commission on Strategic Forces (the Scowcroft commission), in effect acknowledged that none of the proposed schemes could guarantee the ability of the MX force to absorb a full-scale nuclear attack by the U.S.S.R. and retaliate successfully thereafter. Given that judgment, a plan that calls for deploying the first contingent of the new missiles in the unimproved launch facilities of the very missiles the MX was designed to replace seems illogical, to say the least.

Actually the situation may be worse than that. The 10-warhead MX is a much more threatening weapon than the smaller three-warhead Minuteman III, currently the most lethal ICBM in the American arsenal. Indeed, taking into account planned improvements in accuracy, it can be estimated that each MX missile will have a total destructive capacity between 16 and 24 times that of a Minuteman III missile. All other things being equal, the deployment of MX missiles in Minuteman silos would appear to increase the incentive for the U.S.S.R. to attempt a preemptive strike against these installations. Hence the end result could well be a net reduction in the stability of the present balance of nuclear deterrence.

Will all other things remain equal, however? In defending the latest MX decision before a congressional committee last April, Secretary of Defense Caspar W. Weinberger and Chairman of the Joint Chiefs of Staff General John W. Vessey, Jr., indicated one thing that might not, namely the willingness of the U.S. to absorb an initial attack before committing itself to retaliation. Although the two officials refused to elaborate in public, their remarks implied that the U.S. might be planning to accompany MX deployment with an increased inclination to launch its ICBM's after acquiring reliable evidence that a Russian attack was under way but before its actual effects were felt. This policy, usually labeled "launch on warning" or "launch under attack," has been advocated by some military strategists as a solution to the perceived problem of ICBM vulnerability, and it has been criticized by others as a dangerous consequence of the vulnerability problem. Behind this difference of opinion lies one of the most difficult problems of contemporary national-security policy.

In examining this issue, which has a direct bearing on the likelihood of nuclear war, it is helpful to distinguish among several levels of analysis. On one level there are theoretical arguments about nuclear strategy, supported by strictly technical calculations. On another level there are some practical judgments that have been made in applying these theories to the U.S. strategic-force structure. Finally, there are several underlying military realities that are rarely acknowledged in public discussions of national-security policy. The problem under analysis changes character quite significantly when it is viewed from these different perspectives.

At the level of theoretical argument the problem of ICBM vulnerability is clearly and compellingly defined. Standard calculations have been worked out by U.S. strategists to estimate the probability that a fixed ICBM launch facility—a silo—could survive a nuclear attack. These calculations are based on a few technical factors, all of which can be measured with some degree of precision; they include the hardness, or resistance to blast damage, of the missile silos under attack and the number, explosive yield, accuracy and overall operational reliability of the attacking warheads. The characteristics of the most advanced warheads are such that any ratio of warheads to silos exceeding about 2.5 to one will result in a calculated rate of ICBM survival so low that it approaches complete destruction. At present the U.S.S.R. is believed to substantially exceed this ratio against the U.S. ICBM force. The U.S., by combining 550 Minuteman III missiles carrying three warheads each with 100 MX missiles carrying 10 warheads each, would have a comparable capability against the 638 silos that currently house all the advanced warheads of the Russian ICBM force.

According to the prevailing theory of nuclear deterrence, which assumes that the actions of both sides are governed by refined rational assessments, this mutual vulnerability of the ICBM forces creates a threat of limited nuclear war and the possibility of political coercion based on such a threat. Under the pressure of a severe crisis, it is supposed, one side might deliberately initiate an attack directed solely at the vulnerable ICBM installations of the other side in order to partially disarm the opponent and to establish military superiority. The theoretically rational victim, having lost its readiest and most precise offensive force, would be unable to respond in kind, would be deterred by the prospect of greater damage from retaliating against other military and civilian targets and would therefore seek political accommodation to prevent a further extension of the war. By this logic mutual ICBM vulnerability appears to offer a significant if partial victory to whichever side is able to muster the will to exploit the situation.

A systematic policy of launch under attack is designed to provide an answer to this problem. Technical calculations roughly comparable in their scientific basis to those used to define the problem of ICBM vulnerability suggest the feasibility in principle of detecting the initiation of an opponent's attack and of launching some fraction of the threatened ICBM force before the attack could be completed. The most sophisticated variant of the idea, advanced by Richard L. Garwin, calls for a system of infrared sensors in space to detect the attacking missiles early in their boost phase and to measure the scale of the attack. The information would be communicated to ground stations in the U.S. and confirmed by appropriate national authorities. An emergency communications system based on an existing relay satellite or a specially launched rocket would then broadcast orders to launch a significant but strictly limited fraction of the threatened ICBM force according to previously established plans.

According to Garwin, there could be enough redundancy in the sensors, communications links and information-processing routines of such a system to protect it against direct attack and to drive to extremely low levels the theoretical probability of a false detection of attack and a violent response. Under most circumstances the mechanisms for the final launch of U.S. missiles would be shut off, thus giving the system a capacity for continuously checking its diagnostic performance while (again in theory) absolutely preventing any retaliatory response due to random failure. For security emergencies a range of internal controls would be prepared as options, allowing the system to operate either automatically (following the initial authoritative confirmation of the attack) or with varying degrees of human intervention to provide additional protection against undesired retaliation.

Within the theoretical framework defined above such a system could present the potential aggressor with enough uncertainty to eliminate the incentive to initiate an attack offered by the vulnerability of ICBM installations. Since any reasonably large probability of a successful launch under attack would presumably have this effect, the potential victim could accept, say, a 20 percent chance of failure to retaliate against actual attacks and still maintain strong deterrence. Strictly as a matter of prevailing theory and supporting technical calculation, then, if one believes in the problem of ICBM vulnerability itself, there appears to be no compelling reason not to believe in an appropriately designed system for launch under attack as a solution.

At a more practical level of assessment there is good reason to doubt the validity of both the problem and the solution. The organizational procedures that govern the operation of strategic weapons make it unlikely that either the preemptive strike on ICBM silos or the protective response of launch under attack could be executed according to theoretical calculations. Nevertheless, in the development of U.S. strategic policy very different judgments have been made. Whereas the problem of silo vulnerability has been accepted as a real issue demanding a practical solution, the safe implementation of a launch-under-attack policy has been considered to be so impractical that systematic efforts to develop the necessary capability have not been attempted.

The practical problems of the attack and the response arise from the same source: a conflict between the demands of peace and those of war that can be seen in technical terms as a matter of balancing the risks of two types of error. In deploying offensive forces for the purpose of deterring war the dominant peacetime objective is necessarily that of preventing accidental or unauthorized use of those weapons (referred to by strategic analysts as Type I errors); preventing errors of this type is a matter of maintaining negative control. If war were actually to break out, the execution of retaliatory attacks to achieve some meaningful military purpose would become the dominant objective, and it would then be important to minimize failures to launch weapons against preassigned targets (Type II errors); preventing errors of this type requires positive control. Negative and positive control inherently conflict; enhancing one diminishes the other to an extent that is determined by the details of the command arrangements for particular strategic-weapons systems.

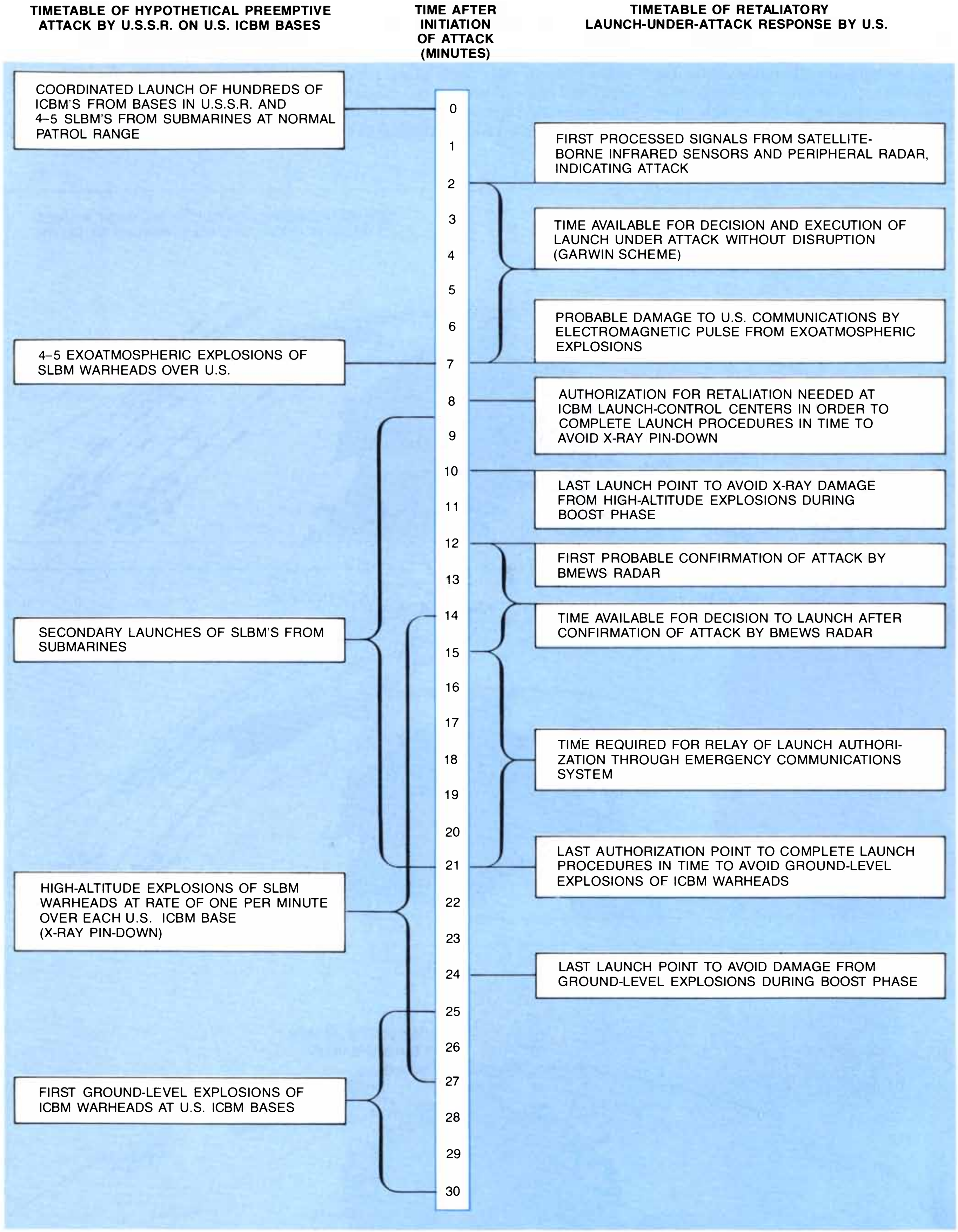
Moreover, this conflict is significantly sharpened in practical terms by the widespread decentralization of control that has been made necessary by the great destructiveness of nuclear weapons. A modern strategic arsenal cannot be entirely controlled through any single location or even through a small number of locations lest these points be identified and preemptively destroyed, thereby incapacitating the entire force. In recognition of this fact the U.S. has dispersed the physical ability to fire nuclear weapons among hundreds of military officers at numerous locations, some of them mobile. From all public indications the U.S.S.R. has done the same.

In the U.S. negative control is ensured through a variety of physical constraints and organizational procedures, making it impossible for any one individual to fire nuclear weapons and improbable that the necessary combination of people (ranging from two to 10 or more depending on the circumstances) could be organized in the absence of plausible political authority. Under this dispersed-control arrangement timing becomes a significant factor in nuclear-weapons operations. In general it is quite difficult to preserve negative-control procedures and also to achieve very precise timing of offensive operations.

Tightly controlled timing is necessary for an attack on ICBM silos in order to minimize interference between warheads exploding in the same area. For a typical U.S. Minuteman base the attack on individual silos must be sequenced from south to north, and the relative timing of individual warheads attacking the same silo must be carefully controlled in relation to the yield and the accuracy of the warheads. Moreover, the task must be accomplished the first time it is actually attempted on a scale of between 2,000 and 3,000 individual warheads. Errors in timing could leave the results of the attack well short of its theoretical potential.

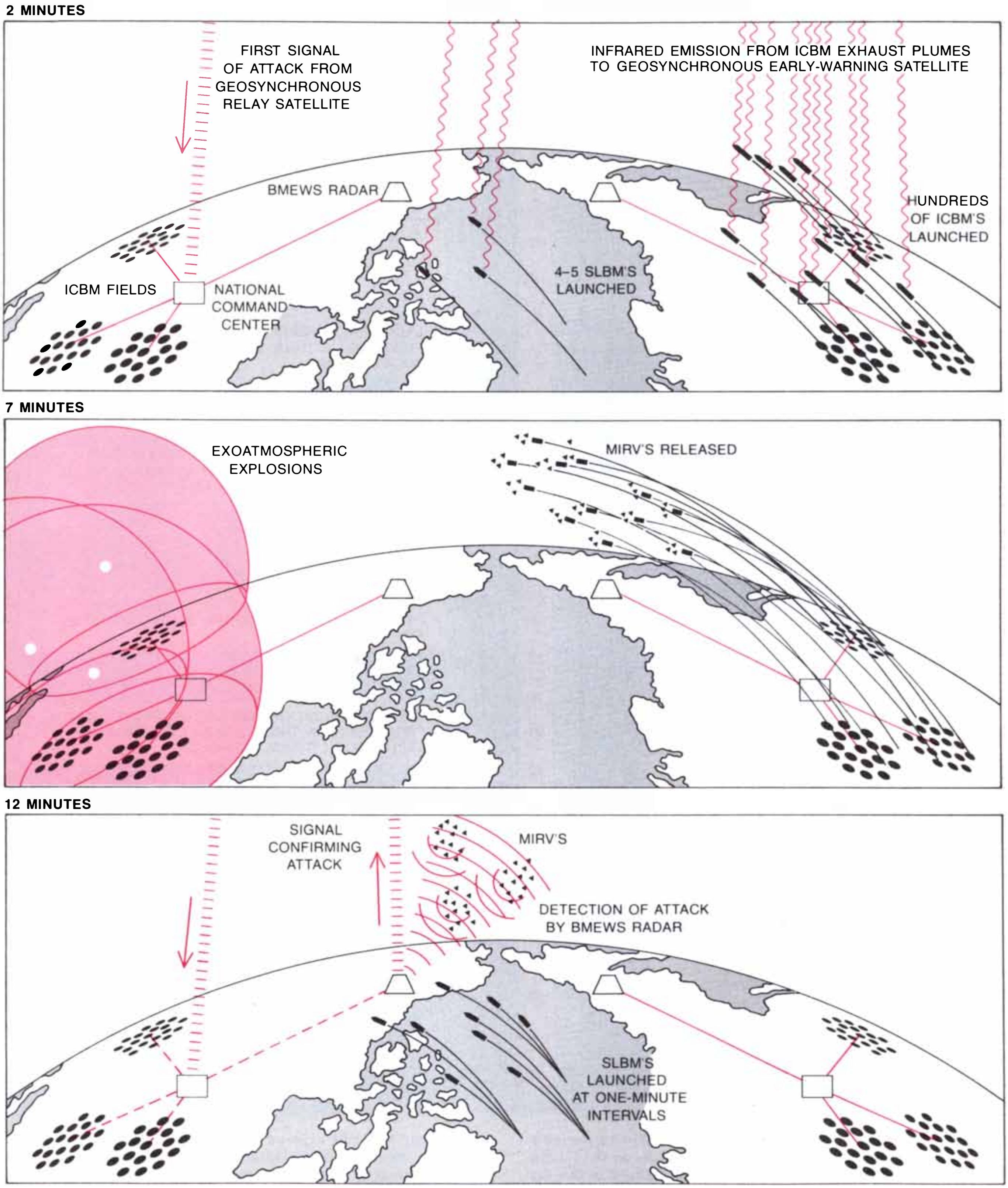
Any potential aggressor able to combine highly refined technical calculations with basic common sense will appreciate that the uncertainties of this situation actually remove much of the theoretical incentive for a preemptive strike on ICBM silos [see "The Uncertainties of a Preemptive Nuclear Attack," by Matthew Bunn and Kosta Tsipis; SCIENTIFIC AMERICAN, November, 1983]. Nevertheless, it is possible to imagine a completely cold-blooded operation exploiting the advantage of initiative to override all the negative-control procedures on the attacking side (without giving any external indication) and to automate the attack sequence completely, thereby solving the timing problem. Because of this theoretical possibility the practical problems of the attacker have generally been discounted in the U.S., and vigorous efforts have been made to devise a solution to the silo-vulnerability problem within the limits of its technical definition, in particular a solution that does not depend on timing errors in the attack sequence. Judging from the various projected costs of the MX program, expenditures on the order of $30 billion to $50 billion have been considered acceptable.

Tightly controlled timing is also nec-



HYPOTHETICAL SEQUENCE OF EVENTS following the initiation of a preemptive nuclear attack by the U.S.S.R. on the U.S. intercontinental-ballistic-missile (ICBM) force is outlined in this diagram. Tightly controlled timing would be necessary not only to execute such an attack on ICBM silos but also to execute a successful launch-under-attack policy in response. The current U.S. Minuteman system would be exposed to disruption during the process of confirming the attack and executing a counterattack. A highly automated variant of the launch-under-attack idea, advanced by Richard L. Garwin, would seek to avoid disruption by arranging for the missiles to be fired in retaliation within five minutes of the first indication of the attack. (SLBM stands for sea-launched ballistic missile.)

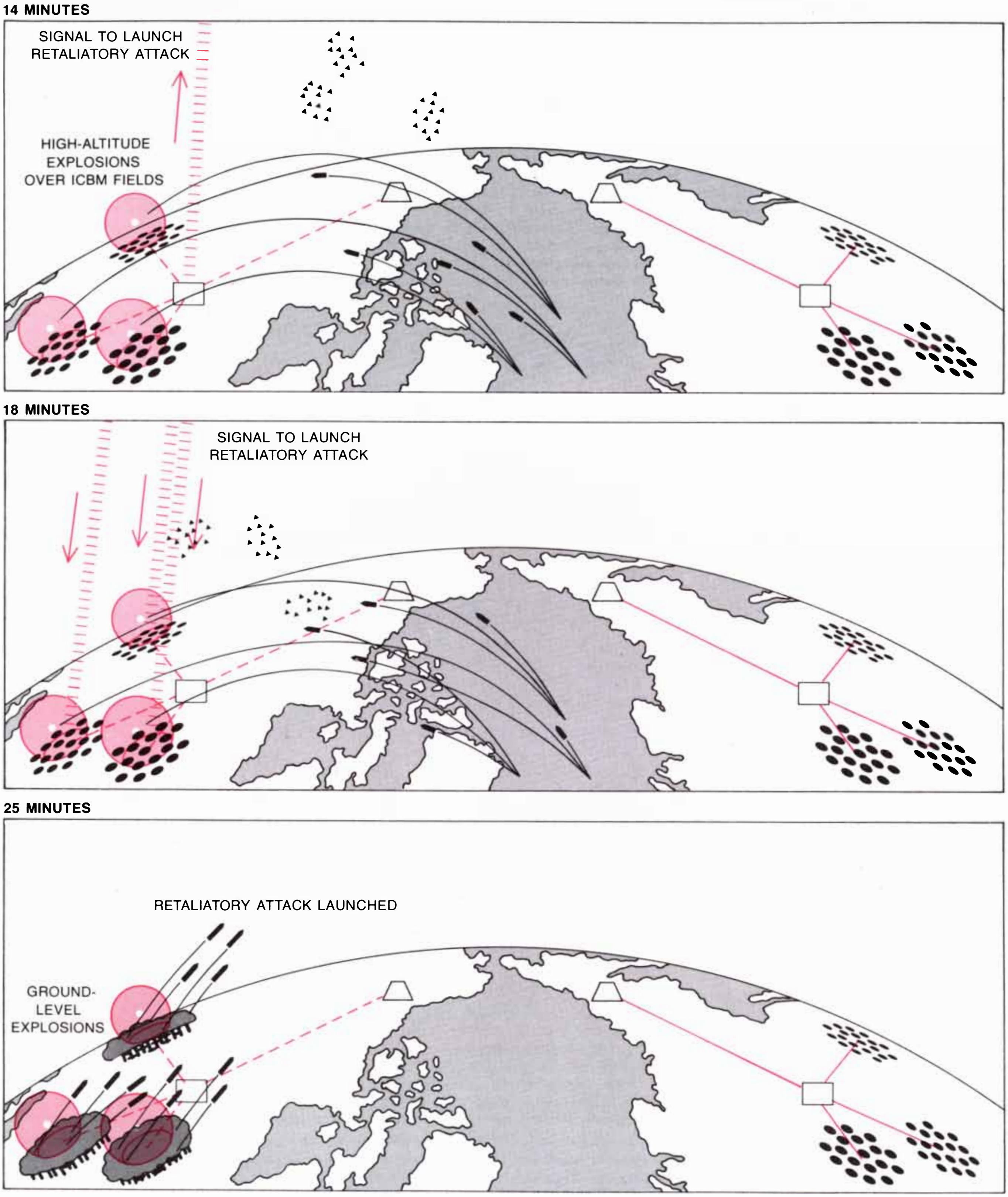
essary to execute a successful launch-under-attack policy. For the first three or four minutes of flight U.S. ICBM's would be substantially more vulnerable to nuclear explosions over their base areas than they would be in their silos. Hence small timing errors in the execution of a launch under attack could facilitate the destruction of U.S. missiles rather than their survival. At lower altitudes early in their boost phase outgoing U.S. missiles would be subjected to debris and high winds generated by ground-level explosions intended to destroy silos; beyond the atmosphere the missiles would be exposed to damage by X rays from high-altitude nuclear explosions. Since the flight time of Russian ICBM's is estimated to be between 25 and 30 minutes and since it takes about two minutes to collect, process and



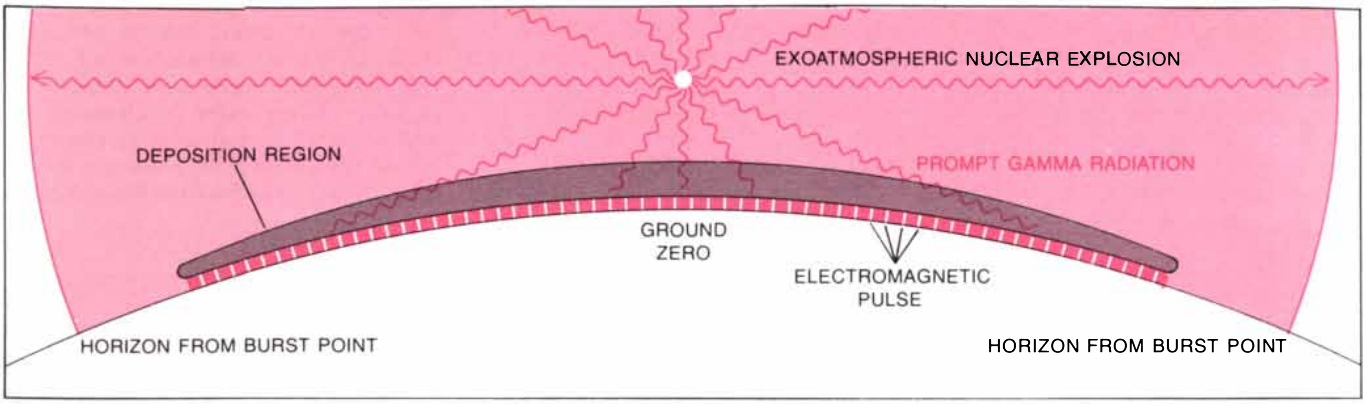
SIX STAGES in the hypothetical attack sequence diagrammed on the preceding page are represented graphically. In this case the decision to retaliate would not be made until after the attack has been confirmed by the land-based Ballistic Missile Early Warning System

transmit the initial indication of the attack (the exhaust plumes of the attacking missiles in their boost phase), slightly less than half an hour would be available for assessing the attack, deciding to counterattack and executing a retaliatory launch in response to an attack solely involving ICBM's. The available time could be further diminished if missiles from Russian submarines were used to generate X rays from nuclear explosions at an altitude of 100 miles or so over the ICBM fields.

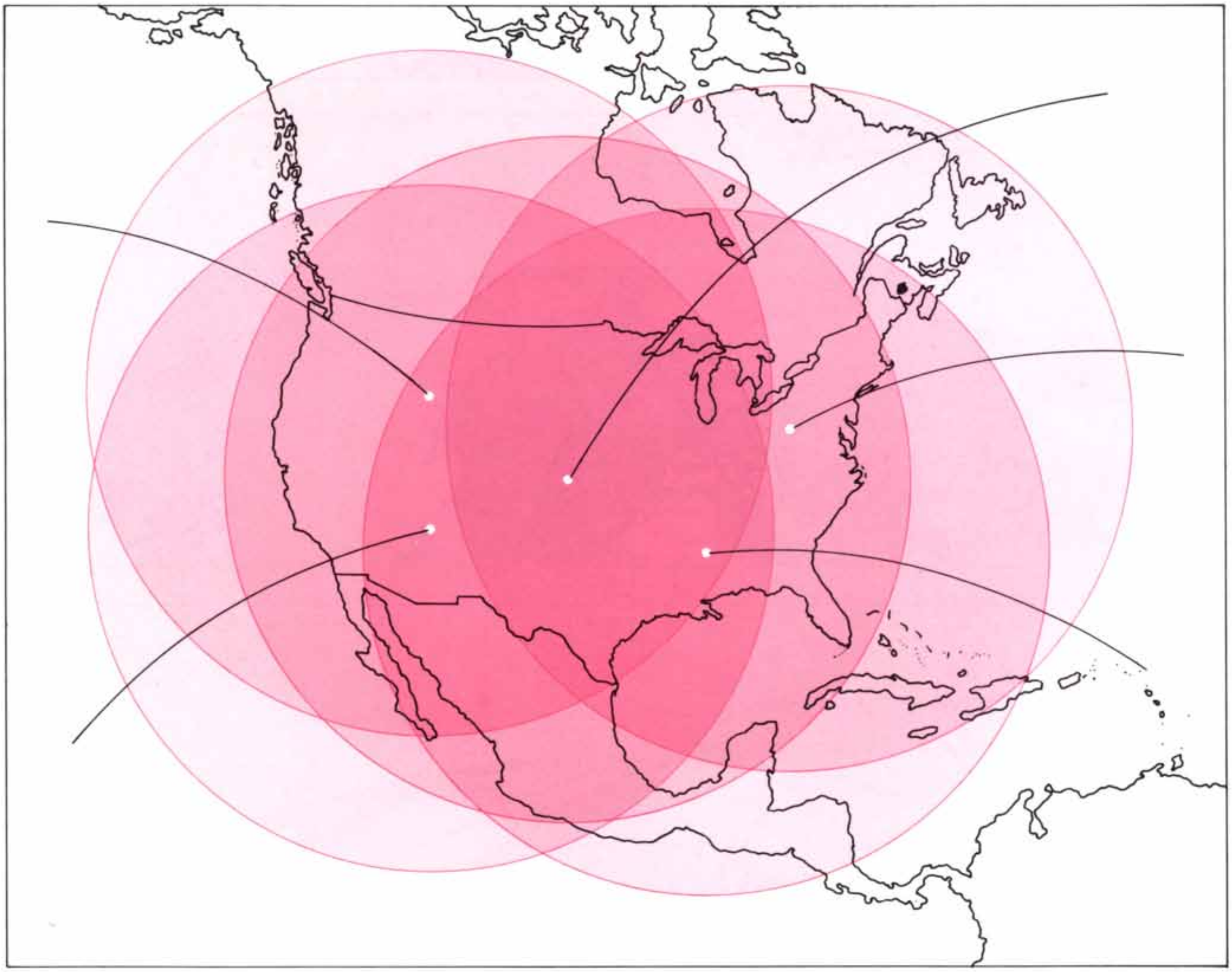
Moreover, even without expanding the scope of targets directly attacked the U.S.S.R. could seriously damage the U.S. command system as it attempts to perform the necessary functions of attack assessment, retaliation decision and launch execution. Within seven minutes of launching, missiles from Russian sub-



(BMEWS). The U.S. ICBM's would be launched early enough to avoid damage from ground-level explosions aimed at silos, but they would pass through a barrage of high-altitude explosions detonated over each U.S. ICBM base in the tactic known as X-ray pin-down.



ELECTROMAGNETIC PULSE generated by a high-yield nuclear explosion 300 miles above the ground could have a devastating effect on communications systems over a large area. The prompt, or instantaneous, gamma rays emitted by such an explosion would strip electrons from air molecules in the upper atmosphere in a roughly circular, pancake-shaped zone known as the deposition region. The free electrons would be accelerated radially by the earth's magnetic field, separating them from the heavier, positively charged ions and creating by a complex mechanism the conditions for a downward-directed high-voltage electromagnetic pulse that would manifest itself in the form of simultaneous electrical surges in exposed conductors on the ground. For a sufficiently large explosion the destructive effects of this high-voltage electromagnetic pulse would extend to the horizon, or tangent point on the earth, as viewed from the burst point (in this case to a radius of about 1,500 miles from ground zero). In addition the disruption of the earth's magnetic field caused by the explosion would result in a second damaging electromagnetic pulse of much lower magnitude but longer duration (hundreds of seconds).



COMPLETE COVERAGE OF THE U.S. by the destructive electromagnetic effects of an exoatmospheric nuclear explosion could in principle be achieved with the detonation of a single one-megaton nuclear warhead at a height of 300 miles over the center of the country. Because the damage caused by the high-voltage electromagnetic pulse would depend in part on the orientation of the exposed conductors to the direction of the pulse, however, the hypothetical attacker would presumably choose to create an overlapping pattern of effects by committing four or five nuclear warheads to this phase of the attack; if the warheads were launched from submerged submarines at their normal patrol range some 1,000 miles off the coasts of the U.S., they would reach their burst points in approximately seven minutes.

marines could produce nuclear explosions 300 miles above the continental U.S. that would not cause blast, thermal or prompt-radiation effects on the ground but would create electromagnetic pulses of two kinds. A pulse on the order of 50,000 volts per meter would arise and dissipate within a microsecond as gamma rays from the explosions stripped electrons from air molecules in the upper atmosphere and the free electrons gyrated around the earth's magnetic lines of force. To a degree dependent on the exact angle of incidence, the resulting electromagnetic pulse would be collected along any exposed conductor and delivered with disruptive or permanently destructive effect to electrical and electronic equipment connected to the conductor.

A second pulse of much lower magnitude (tens of volts per kilometer) but of longer duration (hundreds of seconds) would be induced by temporary disruptions of the earth's magnetic field created by the explosions. The result in this case would be a major disruption of communications lines on land extending over tens of kilometers. The complexities of these phenomena and the inability to test them at full scale make accurate assessments impossible, but it is clear that current land-line communications systems and domestic power grids would be at serious risk. It is not known exactly what would happen if the entire U.S. were to be subjected to the equivalent of thousands or perhaps tens of thousands of simultaneous, highly accelerated bolts of lightning, but any practical assessment must anticipate massive failures of communications systems, power supplies and electronic equipment.

Existing U.S. intelligence sensors, information-processing centers and communications systems could conceivably meet launch-under-attack requirements against an attack by ICBM's alone if there were no electromagnetic pulse or other deliberate disruption of the U.S. command system. Infrared sensors on geosynchronous satellites could detect a Russian missile launch within seconds and processed warning signals could be available in the U.S. within minutes, giving the approximate scale of the attack and the fact that it was directed toward the U.S. Independent confirmation and generalized indications of the areas of the U.S. under attack would be provided within 15 minutes by the Ballistic Missile Early Warning System (BMEWS) radars in Alaska, Greenland and Britain, and this information would be supplemented at 20 minutes by the single U.S. anti-ballistic-missile (ABM) radar at Grand Forks, N.D. In combination these indicators could reliably distinguish actual attacks from other phenomena. A telecommunications conference among the president and the nation's top military commanders could begin within minutes of the first warning signals. If the authorization to respond were communicated with the first confirmation of attack, the missile-launching procedures that embody negative control could in principle be completed in time for the U.S. ICBM's to escape from their silos and travel a safe distance from their base areas.

Electromagnetic disruption, however, could make it quite unlikely that the full sequence could be completed. The probable failure of land lines would interrupt the transmission of warning information and high-data-rate military communications either at the outset of the attack or after the initial warning indicators had been received but had not yet been acted on. Residual emergency communications depend heavily on ultrahigh-frequency radio transmissions relayed between aircraft that must themselves be launched in response to a warning of attack; at the moment these aircraft are thought to be insufficiently protected against the effects of the high-intensity electromagnetic pulse. It would require approximately six minutes to relay messages across the country through this network. With disruption of the ultrahigh-frequency relay network and supplemental jamming, the surviving communications system would operate at low data rates on the low-frequency and very-low-frequency bands and could require almost 10 minutes simply to transmit the coded message authorizing retaliation.

With only partially disseminated warning signals and disrupted communications, normal command procedures encumbered by negative-control requirements would almost certainly delay the protective firing of U.S. missiles, and improvised attempts to expedite the process would risk perverse success: slightly mistimed launches that would expose U.S. missiles to destruction in their boost phase. Given current military capabilities, the prevailing judgment that launch under attack is impractical to the point of impossibility seems to be warranted.

Under the provisions of the current five-year defense plan some of the most critical vulnerabilities of the U.S. command system to electromagnetic disruption will be alleviated. The most important provisions will protect major land-based and airborne command posts, communications-relay aircraft and, to a lesser extent, critical communications lines on land against the effects of electromagnetic pulses from exoatmospheric explosions. A network of ground-wave transmitters operating in the low-frequency band will provide a link for warning information that does not depend on inherently vulnerable cables. Small, mobile stations for receiving information from early-warning satellites will be introduced and bombers will be fitted with low-frequency and very-low-frequency receivers to allow fairly robust radio communications beyond the line of sight of the relay-aircraft network.

These improvements will make the disruption of the U.S. command system more difficult to accomplish without significantly and provocatively expanding the targets that are directly attacked with the standard blast effects of nuclear explosions. The program was not designed, however, to provide the coherent, reliable launch-under-attack capability theoretically outlined above, and as a practical matter it will leave the U.S. far short of such a capability, a discrepancy that might be measured in decades of effort. Before the consensus necessary to undertake such an effort could be marshaled serious consideration would undoubtedly have to be given to what the state of the command system and the issues of control mean for the entire conception of the problem.

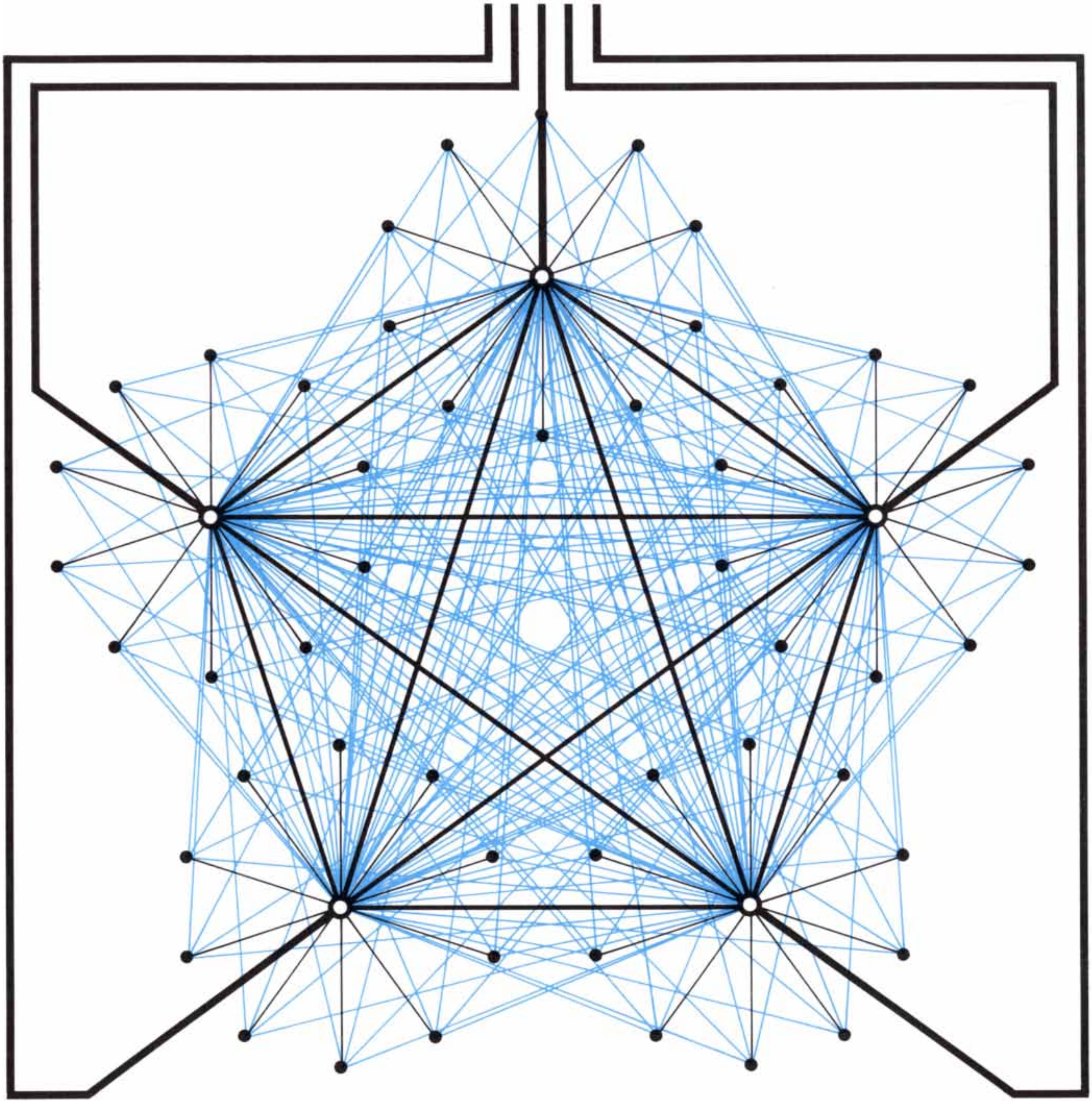
From this latter, more general perspective the widely accepted impracticality of a launch-under-attack policy raises security issues much more fundamental than the problem of silo vulnerability. If the U.S. command system can be significantly disrupted by the induced electromagnetic effects of a few exoatmospheric explosions, it can be effectively destroyed by the direct blast effects of a few hundred weapons and very severely degraded by as few as 50 weapons. These attacks would not be operationally demanding in terms of the technical performance of the warheads and the degree of potential interference among them. If the protective launch of missiles during an attack cannot be reliably accomplished with a disrupted command system, the feasibility of launching the surviving weapons after attack can be questioned and the execution of coherently planned retaliatory missions can be seriously doubted.

Moreover, although the Russian command system is thought to be more protected, the destructive effects of nuclear weapons appear to be so inherently damaging to any command network that differences in exposure between the U.S. and the U.S.S.R. are not likely to be significant given the scale of offensive firepower available. The implication, therefore, is that for both countries the vulnerability of their overall strategic-command systems is a far more important problem than the vulnerability of individual weapons. Furthermore, this has been true since the first mutual large-scale deployment of offensive ballistic missiles began two decades ago.

Compared with silo vulnerability, command-system vulnerability presents a much more powerful incentive to initiate attack before damage has actually been suffered, an incentive that is driven not by theoretically defined opportunity emerging from refined technical calculations but rather by practical fears of decisive defeat in a war that cannot be avoided. Whatever the declared national-security policy might be in peacetime, the incentive appears to impose on responsible military commanders in both the U.S. and the U.S.S.R. potentially overwhelming pressures for outright preemption under intense crisis circumstances when the prospect of an unavoidable war would be facing them.

In order to assess these implications it would be desirable to have calculations of the effects of nuclear attack on an entire command system that are commensurate with those used to estimate the effects on missile silos. Such calculations have not been done in public, however, undoubtedly in part because of the special security sensitivities involved but also in part because the problem entails complex interactions of technology and human organization, and rigorous assessment of such problems is not well established. Nevertheless, without attempting to make a statistically valid estimate, it is possible to conceptually model the interactions of command-system vulnerabilities, procedures for ensuring both negative and positive control, and the time sequencing of both the attack and the response.

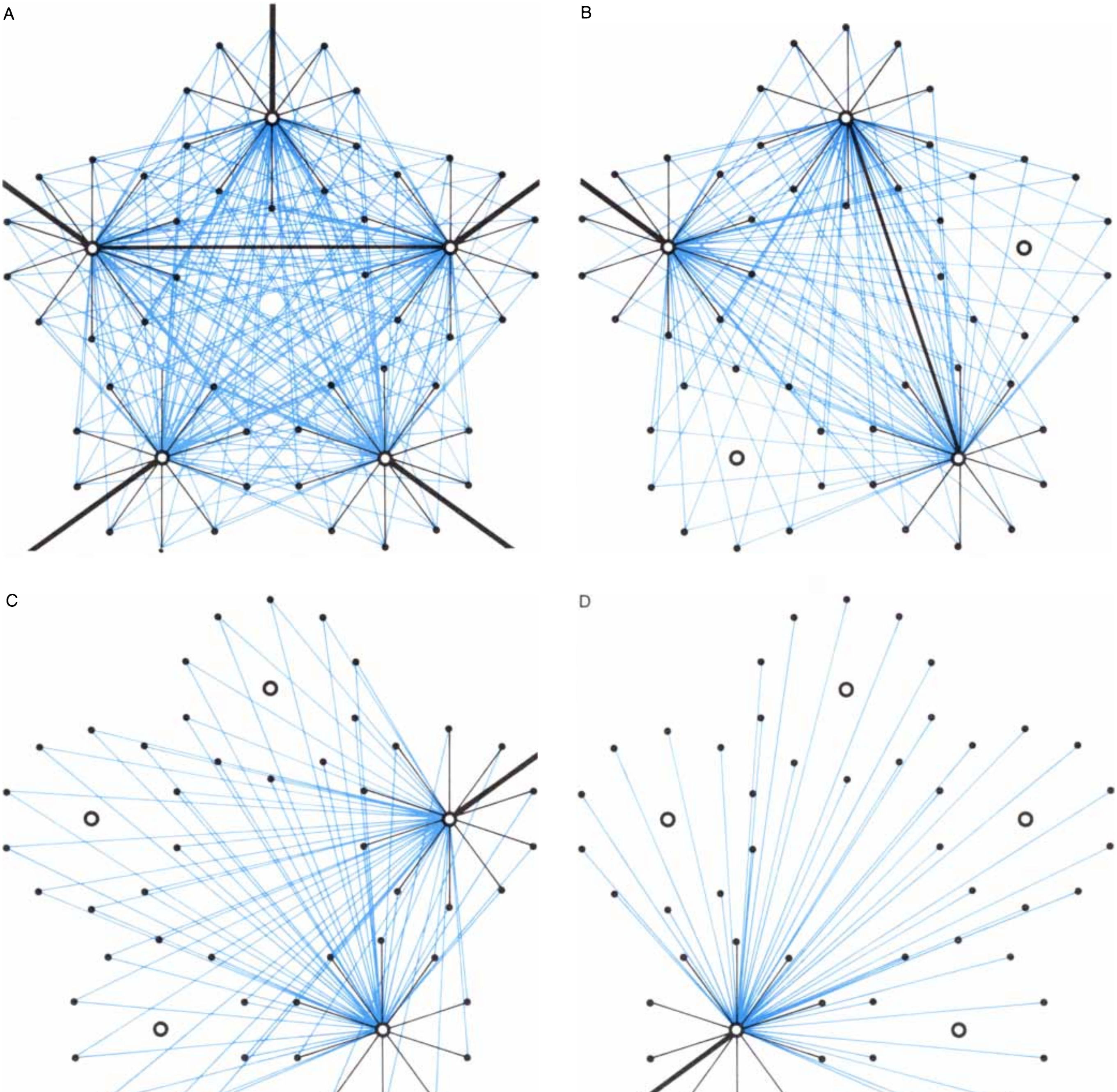
A simple model for this purpose has been developed by Paul Morawski and



COMMAND SYSTEM for the 50 missiles of a Minuteman II squadron is represented schematically in this illustration, based on a simplified computer model developed by Paul Morawski and Bruce G. Blair at the Brookings Institution. Communications lines link the five control centers to the outside world (*thick black lines*), to one another (*medium black lines*) and to each of the 50 missiles. Each center normally has primary responsibility for 10 missiles (*thin black lines*) and monitoring responsibilities for the other 40 (*color lines*). The computer program incorporates complex rules that embody negative and positive control procedures governing the operation of the missiles.

Bruce G. Blair at the Brookings Institution in the form of a computer program that incorporates the publicly documented command arrangements and procedural rules by which the 50 missiles of a Minuteman II squadron were said to be launched as of the mid-1970's. This system is a very restricted part of the U.S. strategic-command network and conveys only part of its complexity and little of its hierarchical depth. Moreover, the model reflects only part of an actual squadron's operations. Within this highly simplified representation, however, some results appear to suggest some of the internal pressures that can be expected to confront actual strategic forces.

The model depicts five separate control centers linked to one another and to each of the 50 missiles [*see illustration on opposite page*]. The programmed rules that determine each center's actions reflect the negative and positive control procedures needed to operate Minuteman II missiles. Each center normally has primary responsibility for 10 missiles (an instance of positive control) and monitoring responsibilities for the other 40 (negative control). Each center is a designated substitute for one other center; if a given center is destroyed, its substitute will assume responsibility for that center's missiles (positive control). A center will execute instructions to set an overall target plan for all 50 missiles, to assign targets to individual missiles under its jurisdiction and to launch or withhold its primary missiles (positive



FOUR CONDITIONS OF DAMAGE to the command system of a Minuteman II squadron are ordered in this illustration by degree of severity. In condition *A* two control centers are assigned to the operational state labeled normal (with all communications links intact) and three are rated vertical (with links to other centers destroyed). In condition *B* one center is vertical, two centers are horizontal (with the link to external commanders destroyed) and two are gone (all links destroyed). In condition *C* one center is vertical, one is silent (with links to only the missiles intact) and three centers are gone. In condition *D* one center is vertical and the other four centers are gone.

control), but only after completing a procedure to determine that the instructions are valid (negative control).

For each center the validation procedure requires confirmation from one other control center in the system (negative control). A missile will be launched toward its assigned target promptly only if the instruction it receives to do so is confirmed by a second control center. Each center will veto the launch instructions of another center if it considers them to be improper, either because it has not received and validated an authoritative instruction (negative control) or because its target assignments to individual missiles have not been completed (positive control). A veto will cancel the launch instruction of any single center (negative control) but will have no consequence if two centers concur in the launch instruction (positive control). A single center's launch order that is neither confirmed nor vetoed will ultimately be executed (positive control), but only after an extended delay (negative control). These rules embody some of the procedural complexity of launching ballistic missiles under dispersed command arrangements.

The model executes these procedures with constant and efficient timing and allows none of the errors that might ensue if confused, frightened human beings were performing them at the brink of nuclear war. It does account, however, for physical damage to communications links that might result from the opponent's attack. It does so by assigning each center to one of five operational states: normal (with all communications links intact), vertical (with links to other centers destroyed), horizontal (with the link to external commanders destroyed), silent (with links to only the missiles intact) and gone (all links destroyed and unable to operate). From this array of possibilities 525 significantly different conditions of partial damage can arise within the five-center network. Even with all procedural errors excluded and no possibility allowed for initiating actions other than those mandated by given instructions, interactions between the timing of damage in the network and the execution of the established control procedures can produce results that deviate from the intentions represented in the instructions. Failures of both negative and positive control are possible.

Some representative results of the model were obtained under four conditions of damage to the command-center network, ordered roughly by degree of severity [*see illustration below*]. In all cases the instructions simply require the execution of a single targeting plan for all the missiles. Three different targeting plans were tested, representing crudely an unrestricted retaliation, one focused exclusively on military targets and one that is deliberately limited in character. In each case the performance of the system when instructions are received shortly before damage occurs is compared with its operation when the damage precedes the instructions. In spite of the severe damage depicted the modeled system performs well in general, but there are some notable problems. The intermediate damage conditions (labeled *B* and *C*) create a catastrophic positive-control failure when instructions to retaliate arrive after the damage occurs. In these circumstances the procedural rules create internal contradictions and no missiles are launched. The system does manage to launch the instructed retaliation under the most severe damage condition (*D*) but with a substantial delay—90 minutes—that reduces the presumed value of attacking "time urgent" targets such as airfields, submarine bases, mobile missiles and command centers. Negative control generally succeeds, but under the intermediate damage conditions missiles intended to be withheld in the limited retaliation plan are in fact launched.

The performance of the modeled command system deteriorates as it is asked to execute more complicated instructions representing conceptually an attempt to manage the process of re-

	PLAN I: STANDARD RETALIATION		PLAN II: TIME-URGENT RETALIATION		PLAN III: LIMITED RETALIATION	
COMMAND-CENTER CONDITIONS	ORDERS ARRIVE BEFORE DAMAGE	ORDERS ARRIVE AFTER DAMAGE	ORDERS ARRIVE BEFORE DAMAGE	ORDERS ARRIVE AFTER DAMAGE	ORDERS ARRIVE BEFORE DAMAGE	ORDERS ARRIVE AFTER DAMAGE
A	P = 100 N = 0	P = 94 N = 1	P = 100 N = 0	P = 93 N = 0	P = 100 N = 0	P = 95 N = 3
B	P = 99 N = 0	P = 0 N = 0	P = 95 N = 3	P = 0 N = 0	P = 96 N = 8	P = 0 N = 0
C	P = 98 N = 2	P = 0 N = 0	P = 91 N = 5	P = 0 N = 0	P = 90 N = 14	P = 0 N = 0
D	P = 42 N = 1	P = 44 N = 2	P = 19 N = 1	P = 23 N = 0	P = 36 N = 0	P = 40 N = 0

RESULTS OBTAINED FROM THE MODEL of the operation of a Minuteman II squadron are summarized in this table by separate scores for positive and negative control, averaged over 10 iterations of the model. The positive-control score (*P*) gives the percentage of successfully executed retaliatory strikes defined by assigning weights to four categories of targets in the ratio of 1 : 2 : 3 : 4. Categories 2 and 4 represent "time urgent" military targets, and their value is reduced respectively by 50 percent over 90 minutes and by 80 percent over 60 minutes. Perfect execution (*P* = 100) is scored when all missiles are sent to their designated targets on time. The negative-control score (*N*) gives the proportion of missiles sent to unintended targets as a percentage of the total number of missiles launched. Plan I assigns missiles more or less equally to the four categories of targets. Plan II assigns 15 missiles to category 3 and 35 to category 4. Plan III assigns 10 to category 3 and 10 to category 4, withholding 30 missiles. One problem that arises is that under the intermediate damage conditions (*B*, *C*) a catastrophic positive-control failure is created when the instructions to retaliate arrive after the damage has occurred. The system manages to launch the retaliation under the severest damage condition (*D*) but only after a delay of 90 minutes. Under conditions *B* and *C* there is also a statistically smaller failure of negative control: some of the missiles intended to be withheld are in fact launched.

taliation under the time pressures that launch-under-attack operations would impose [*see illustration at right*]. The system is first given instructions to execute one of the limited plans and then to switch to full retaliation. Damage is inflicted either in the interval between the receipt of the two sets of instructions or after the arrival of the new instructions. The average results over 10 trials shown in the table indicate failures of both positive and negative control.

A review of the disaggregated results by individual trials revealed two different sources of failure. In switching between the two standard targeting plans the original plan prevailed on two of the 10 trials in spite of the subsequent instruction, as determined by the random factor of which of the centers happened to initiate its operations first. When greater requirements were imposed (modifying the standard plan by giving supplementary target assignments to specific missiles), the entire process consumed more time. In that circumstance the implementation of the instructed switch in plans was achieved on all 10 trials, but some of the retargeting operations were not completed by the time the missiles were launched. This process of switching plans is similar in its demands on the system to that of terminating a launch plan short of full execution. Under some damage conditions termination is not possible once execution of the instructions is under way.

	PLAN II FOLLOWED BY PLAN I		PLAN III FOLLOWED BY PLAN I	
TARGETING PLANS	NEW ORDERS ARRIVE BEFORE DAMAGE	NEW ORDERS ARRIVE AFTER DAMAGE	NEW ORDERS ARRIVE BEFORE DAMAGE	NEW ORDERS ARRIVE AFTER DAMAGE
STANDARD PLANS	P = 98 N = 2	P = 51 N = 13	P = 98 N = 2	P = 69 N = 3
RETARGETING APPLIED TO PLAN I	P = 87 N = 7	P = 49 N = 15	P = 87 N = 7	P = 68 N = 13

PERFORMANCE OF THE MODEL deteriorates as it is asked to execute more complicated instructions. In this case the command system is first given instructions to execute either Plan II or Plan III and then to switch to Plan I. Damage is inflicted either during the interval between receipt of the two sets of instructions or after the new instructions arrive. In general the results reflect failures of both positive and negative control. When the supplementary target assignments of the standard plans are modified (*bottom row*), the entire process takes more time and some of the retargeting operations are not completed by the time the missiles are launched.

The Brookings model is an exercise in logic rather than an empirical assessment, and its specific results are not as important as its general implications. In actual practice Minuteman missiles can be controlled through an airborne command post as well as through the ground network of fixed centers; if the aircraft were itself able to avoid serious disruption, it could overcome some of the specific failures reported in the model's results. Even with greater depth and procedural adroitness, however, the real strategic-command system cannot remove the inherent problems the model illustrates. Partial disruption is a major threat to the coherence of dispersed control arrangements. It creates a variety of possible circumstances too substantial to allow clear prediction and significantly sharpens the trade-off between negative and positive control.

Military commanders aware of these problems will not have confidence in the net performance of their forces once serious disruption of the command system occurs, nor will they know how to strike an optimum balance between negative and positive control. This factor translates into a very strong incentive to initiate offensive operations before damage occurs in order to ensure both the timing and the coverage of targets necessary to achieve military objectives. It also identifies the opponent's command system as the target of greatest opportunity and the most likely means of achieving victory. Thus the two opposing strategic forces appear to impose on each other powerful incentives for preemption as the most promising means of conducting nuclear war.

This conclusion reveals a threat of war very different from those commonly imagined. For decades deterrence has seemed strong enough to guarantee that political leaders will avoid nuclear war as long as events remain subject to their conscious decisions. Existing command systems provide reasonable assurance that a technical failure or a human error will not be the prime cause of catastrophe. Under normal circumstances, therefore, nuclear war has been considered a sufficiently remote possibility. Both nuclear establishments, however, are subject to potentially fatal stress under crisis conditions. Both are committed to conducting a nuclear war by means of elaborately programmed, carefully timed offensive operations that require coordination among the weapons involved. Although both aspire in principle to the ability to adjust these operations to the exact circumstances of war and although both have made a considerable investment in the protection of their forces in pursuit of that principle, actually neither can guarantee such coordination after the first 50 or 100 nuclear explosions have been absorbed. If war should ever appear unavoidable, military commanders on both sides charged with executing their assigned missions would inevitably seek authority to initiate an attack, whatever prior national-security policy may have been. They would do so with a forcefulness that would depend directly on the intensity of the crisis. The pressures on political leaders at that point would be severe. Although there is no reason to doubt their continuing desire to avoid war, there are strong reasons to doubt their ability to contain their respective strategic organizations.

Since there has not yet been a crisis severe enough to test these pressures and since prevailing strategic ideas do not readily identify the issue, this problem has not been widely appreciated. The Cuban missile crisis of 1962 happened before either of the two strategic forces had fully matured, and the Middle East crisis of 1973 did not trigger a full alert of nuclear weapons on both sides. Moreover, the weapons systems that most exacerbate the problem are not those with advanced accuracy whose offensive capabilities are analytically best understood but rather those that affect the timing and coordination of strategic operations. The American Pershing II intermediate-range missile in Europe, for instance, significantly reduces the reaction time available to the Russian command system. Equally troubling is an often threatened but as yet undefined Russian capability to "put the U.S. in an analogous position." In addition antisatellite weapons capable of attacking critical warning and communications systems in space could severely degrade the coherence of an opponent's command system.

The importance of those weapons developments has not been recognized in the public debate over national-security policy. In the final analysis, therefore, the idea of launch under attack is best seen neither as a theoretical answer to the silo-vulnerability problem nor as an impractical policy that can be avoided, but rather as a symptom of perilous security conditions that demand more thought than they have yet received.

The Atmospheric Effects of El Chichón

This relatively small but sulfur-rich eruption in 1982 injected into the stratosphere a fine mist of sulfuric acid droplets denser than any volcanic cloud since the great 1883 eruption of Krakatau

by Michael R. Rampino and Stephen Self

The eruption of the Mexican volcano El Chichón in late March and early April of 1982 was not particularly large but it injected an unusually large amount of volcanic material into the stratosphere. The eruption sent ash and volcanic gases more than 25 kilometers aloft. Satellite observations made immediately afterward showed that the material in the stratosphere moved rapidly westward; within a few weeks a veil of fine material extended around the globe. In less than a year the stratospheric cloud had blanketed the entire Northern Hemisphere and a good part of the Southern Hemisphere.

It has been known for some time that volcanic clouds in the stratosphere can affect world climate, notably by causing a decrease in the mean global or hemispheric temperature. Until recently it was thought the volume of fine ash and dust ejected by an explosive eruption was a good measure of the density of the cloud it would produce and therefore of its effect on climate. For example, the dust-veil index introduced in 1970 by the British climatologist Hubert H. Lamb ranks the clouds created by historical eruptions on the basis of estimates of the amount of ash erupted, reports of atmospheric optical effects and the decrease in surface temperatures in the years following the eruption. Over the past 10 years, however, it has become clear that most of the dust settles out of the stratosphere in a few months and that long-lived volcanic clouds are composed not of dust but of an aerosol of sulfuric acid droplets. The amount of sulfur-rich gases released by an explosive volcanic eruption may therefore be a better indicator of its atmospheric effects than the volume of ash erupted.

The eruption of El Chichón demonstrated for the first time that a relatively small but sulfur-rich volcanic eruption can indeed produce a dense, widespread stratospheric cloud. Although El Chichón ejected a volume of ash comparable to that erupted by Mount St. Helens in May, 1980, the stratospheric cloud it created was on the average some 100 times denser. Indeed, it was the densest cloud seen in the Northern Hemisphere since the eruption of Krakatau in Indonesia in 1883. The cloud thrown up by Mount St. Helens was composed primarily of fine ash that rapidly aggregated into larger particles and settled out of the atmosphere, whereas El Chichón left a dense haze of sulfuric acid droplets that will take several years to settle out completely.

The size of the aerosol cloud can be related to the volcano's geologic characteristics, such as its general compositional type and its eruptive style. El Chichón emitted enough sulfur gases to manufacture substantial amounts of sulfuric acid and was powerful enough to inject the gases into the stratosphere partly because it was a volcano of intermediate silica content. Volcanoes with a low content of silica tend to have a high content of sulfur; they also tend to erupt less explosively than those with a high content of silica. El Chichón, however, had a much higher sulfur content than volcanoes of the same general compositional type, an anomaly geologists are still seeking to explain. The sulfur may have come either from sedimentary deposits underlying the volcano or from sulfide deposits on one plate of the earth's crust that is plunging under another and melting.

The sulfuric acid aerosol is formed by the photochemical reaction of sulfur gases released by a volcano with water vapor in the stratosphere. The aerosol causes a decrease in the mean global temperature because the droplets both absorb solar radiation and scatter it back into space. The reactions by which the sulfuric acid is synthesized are complex, however, and the reaction pathways have not been completely worked out. It is also not completely understood how the climatic effects of a volcanic aerosol cloud depend on variables such as the rate at which the sulfuric acid forms, the rate at which the aerosol droplets grow and the rate at which the cloud spreads. Part of the problem has been a lack of data on volcanic clouds in the stratosphere.

The atmospheric effects of El Chichón are better documented than those of any other volcano that generated a large stratospheric cloud. The eruption is thus serving as a natural experiment in which monitoring instruments and theoretical models of the chemistry and climatic effects of volcanic aerosols in the stratosphere can be tested. Although the wealth of information collected during and after the eruption is still being analyzed, it is already apparent that there will be some surprises. For example, one analysis of temperature data suggests that the aerosol cloud may have had its maximum effect on Northern Hemisphere temperatures two months after the eruption, far sooner than most climate models had predicted.

El Chichón is at 17.33 degrees north latitude and 93.2 west longitude in Chiapas, the southernmost state of Mexico. Three plates of the earth's crust meet near southern Mexico: the North American plate, the Caribbean plate and the Cocos plate. The volcanic activity is related to the subduction of the Cocos plate under the North American plate. El Chichón and several associated volcanic centers, however, lie in a gap between the trans-Mexican volcanic belt to the north and the Guatemalan volcanic belt to the south. This isolated position may be related to a break in the Cocos plate caused by the subduction of the Tehuantepec Ridge off the southwestern coast of Mexico.

El Chichón, which was first described in 1928, was a volcanic peak rising only 1,260 meters (4,130 feet) above sea level. The conduits within it were capped with domes or plugs of solidified lava. Although fumaroles in the crater had been steaming for many years, the volcano had long been dormant. Recent carbon-14 dating indicates that the last eruption, which may have been 10 times the size of the one in 1982, occurred sometime between A.D. 1350 and 1400.

The 1982 eruption was preceded by a month of earthquake activity that was

ERUPTION OF EL CHICHÓN in 1982 blasted away the upper 200 meters of the old volcanic cone. The volcano, which is in a remote region of southern Mexico, had been dormant for roughly 600 years. One of the few photographs of the volcano before the 1982 eruption (*top*) shows the rim of the old crater, the dome that capped the central conduit and a flank dome (*lower left*) capping a subsidiary conduit. The crater left by the 1982 eruption (*bottom*) lies within the old crater, which is masked by new volcanic deposits. Water collected in the crater during the rainy season, forming a hot acidic lake; the temperature of the water is about 50 degrees Celsius and the *p*H is less than 1.

SULFUR CONTENT of volcanic deposits just outside the crater is indicated by their yellow color. Analysis of the volcanic ash El Chichón erupted showed anomalously high values of sulfate: up to 2 percent by weight. The sulfur is present in several forms, including crystals of anhydrite ($CaSO_4$), which is rare in volcanic rocks. The yellow crust on these deposits is elemental sulfur that is probably being released by the breakdown of anhydrite. Estimates of the mass of the sulfur gases injected into the atmosphere by the eruption run as high as 20 million tons.

recorded by a network of seismographs set up in the area in 1980 to monitor possible seismic activity caused by water impounded behind dams. The foci of the earthquakes were initially shallow (less than five kilometers deep), and before the eruption they became even shallower (less than two kilometers deep). The seismic activity was probably caused by the motion of magma up toward the surface or by the interaction of the magma with groundwater.

The eruption sequence has been determined from the deposits left by the volcano and their correlation with eyewitness accounts. The deposits were mapped soon after the eruption by Haraldur Sigurdsson and Steven N. Carey of the University of Rhode Island and by J. M. Espindola of the National Autonomous University of Mexico. The divisions between the three major eruptions on March 28, April 3 and April 4 are marked by changes in the grain size of the ash deposits. The three eruptions were of the type called plinian after Pliny the Younger, who described the eruption of Vesuvius in A.D. 79. Plinian eruptions are marked by great columns of gases, dust, ash and pumice that are vigorously driven by convection to heights of tens of kilometers. These eruptions, which lofted the gases and dust that became the stratospheric cloud, account for most of the deposits. The second eruption column, however, collapsed before it dissipated and produced the ground-hugging avalanches of hot gas, ash and pumice called pyroclastic flows. The pyroclastic flows left distinctive deposits of their own.

The first plinian eruption began at 11:32 P.M. on March 28 and lasted for five or six hours. The second, which began at 7:35 P.M. on April 3, differed from the first in that a large amount of old volcanic rock was erupted. The presence of the older rock in the ash layer from this eruption indicates that the conduit of the volcano was being abraded and widened by the passage of magma. It was the widening of the vent and an accompanying decrease in the eruption velocity that may have led to the collapse of the eruption column.

Plinian eruptions occur when the rate of eruption is high and the vent is relatively small (less than 200 meters in diameter). The lower part of the column is powered by the kinetic energy the material acquires in shooting through the conduit and the upper part is powered by convection currents set up in the atmosphere by the hot material. If the eruption velocity decreases, the upper part of the column may become denser than the air and the column may collapse. Material that is still hot then falls back to the ground around the vent and surges outward, powered by the kinetic energy it acquires in falling and fluidized by the hot gases and air. As the pyroclastic material moves outward it segregates into denser, slower-moving pyroclastic flows and lighter, faster pyroclastic surges. The pyroclastic surges spread in all directions from the volcano ahead of the flows, which are generally channeled by the topography.

The surges generated by El Chichón destroyed the dense forest on its flanks and the adjacent hills up to a radius of eight or nine kilometers from the crater. Most of the trees within this zone were blown down and charred, blackened or sandblasted on the side facing the volcano. Temperatures were high enough to ignite dead wood and the wood of houses but generally not high enough to ignite living trees. Although the surges were destructive, they left only a thin deposit that thickens gradually toward the vent. Pyroclastic flows immediately followed the surges, coursing down the narrow river valleys on the slopes of the

TECTONIC MAP of southern Mexico shows the relation of El Chichón to the Mexican and Guatemalan volcanic belts and to local tectonic features. Three major plates of the earth's crust meet near this region. The Caribbean plate is sliding past the North American plate along a series of faults in Guatemala, and the Cocos plate is being subducted under the North American and Caribbean plates at the Middle American trench off the coast of Mexico. The volcanism is related to the subduction of the plate but has features that indicate the tectonics are not straightforward, such as the offset between the trans-Mexican and Guatemalan volcanic belts. The isolated position of El Chichón may be the result of a break in the downgoing plate where the Tehuantepec Ridge, an inactive fracture zone with which it is aligned, is being subducted.

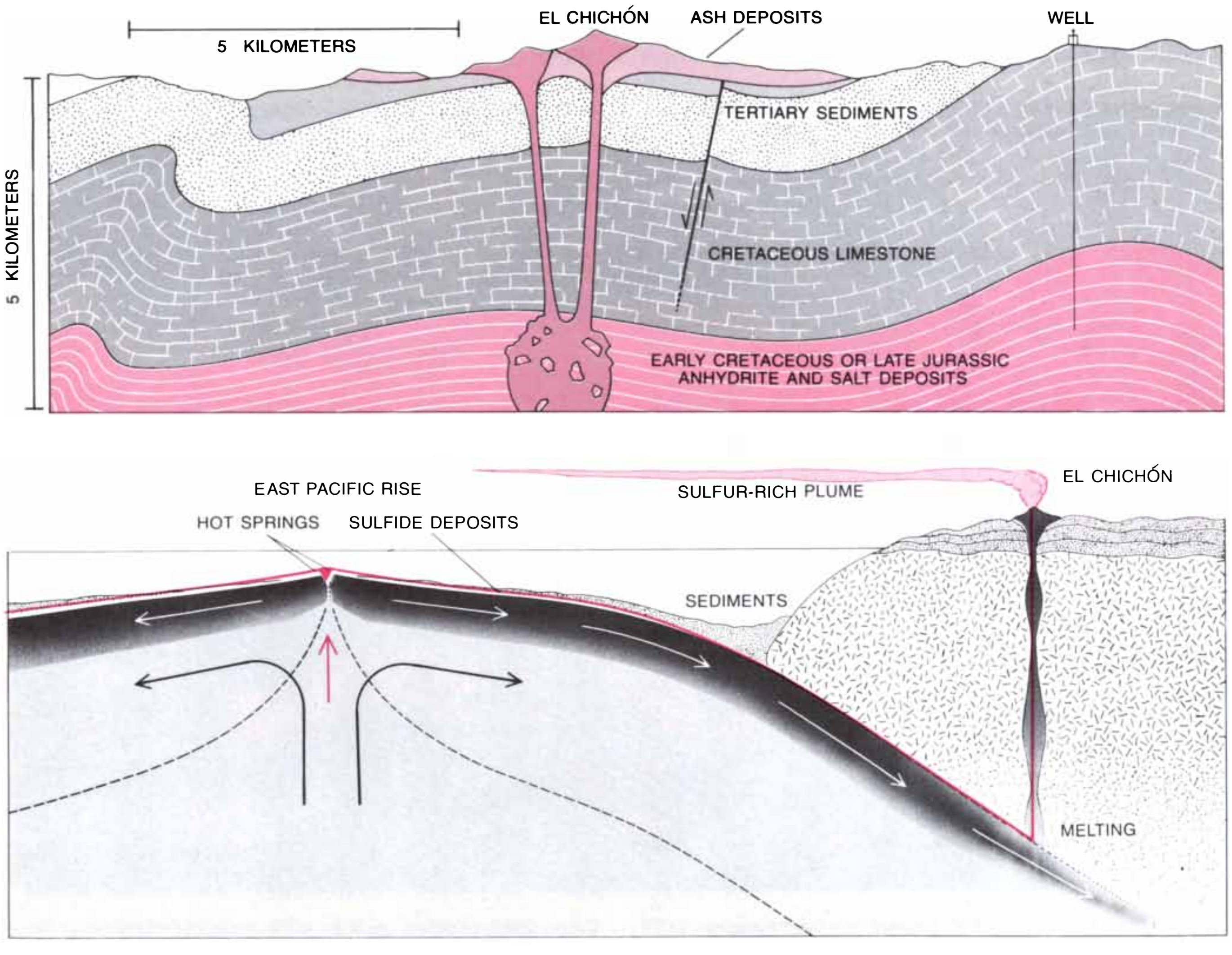
volcano. The flows left comparatively thick radial deposits of ash and pumice that include blocks of the dense volcanic rock up to a meter in diameter and carbonized logs.

The third plinian eruption began at 5:22 A.M. on April 4 and like the second produced ash and pumice deposits including large amounts of old volcanic rock. This time, however, the eruption column appears to have dissipated without collapsing. Close to the volcano the deposits from the third eruption are covered with fine-grained ash. This thin ash layer was probably created when hot pyroclastic-flow material in the valleys came in contact with seasonal floodwaters and caused steam explosions.

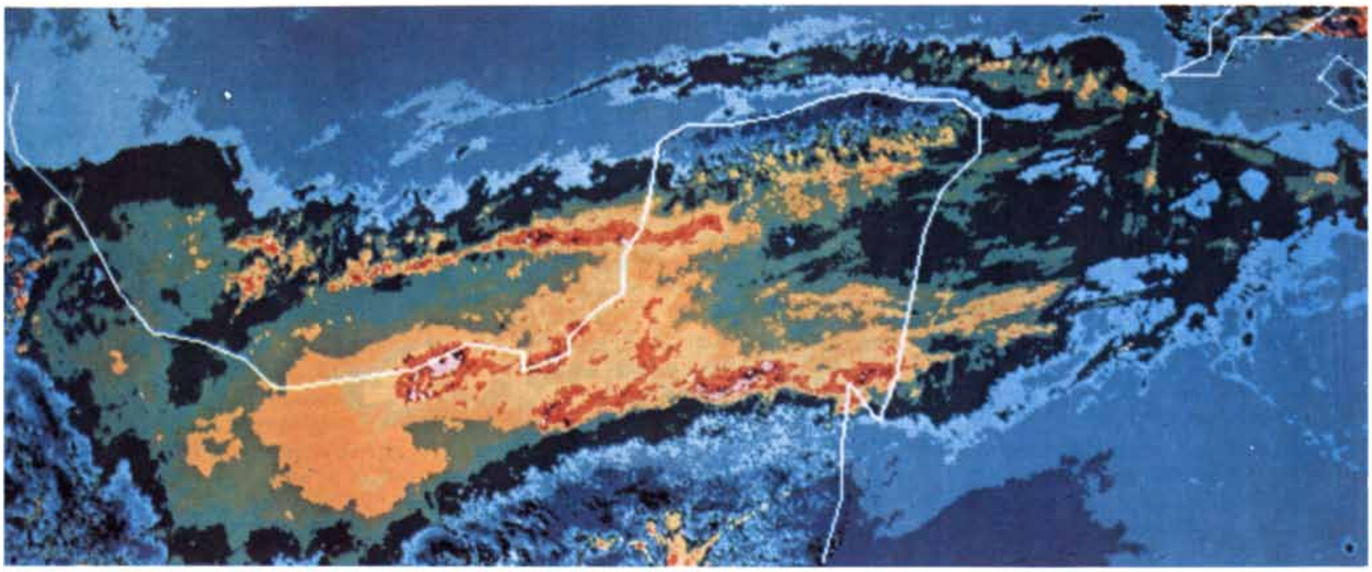
The volume of ash and pumice discharged during the eruption of El Chichón was not particularly large: between .5 and .6 cubic kilometer. For comparison the 1902 eruption of the Guatemalan volcano Santa Maria discharged 10 cubic kilometers of material, the 1883 eruption of Krakatau 20 cubic kilometers and the 1815 eruption of Tambora in Indonesia more than 175 cubic kilometers.

The volcanic ash ejected by El Chichón was of the compositional type called andesite, commonly erupted by subduction-zone volcanoes. (The El Chichón magma is a potassium-rich one called trachyandesite.) The chemical composition of magmas varies from silica-poor, iron-rich basalt through intermediate-composition andesite to silica-rich, iron-poor dacite and rhyolite. Sulfur is more soluble in iron-rich magmas, and so basalt tends to have higher concentrations of sulfur than rhyolite does. To generate significant amounts of stratospheric aerosol, however, an eruption must be sufficiently explosive to loft the sulfur gases high into the atmosphere, and silica-poor magmas tend to erupt less explosively than silica-rich ones. Eruptions of intermediate-composition material are well suited to injecting large quantities of sulfur gases into the stratosphere. The historical record suggests that the atmospheric effects of volcanoes and their general compositional type can be roughly correlated. For example, Mount Agung on Bali, which produced a large stratospheric cloud when it erupted in 1963, was a volcano of andesitic composition. The eruption of Krakatau, the first eruption whose atmospheric effects were intensively studied, was a very large one of dacitic composition.

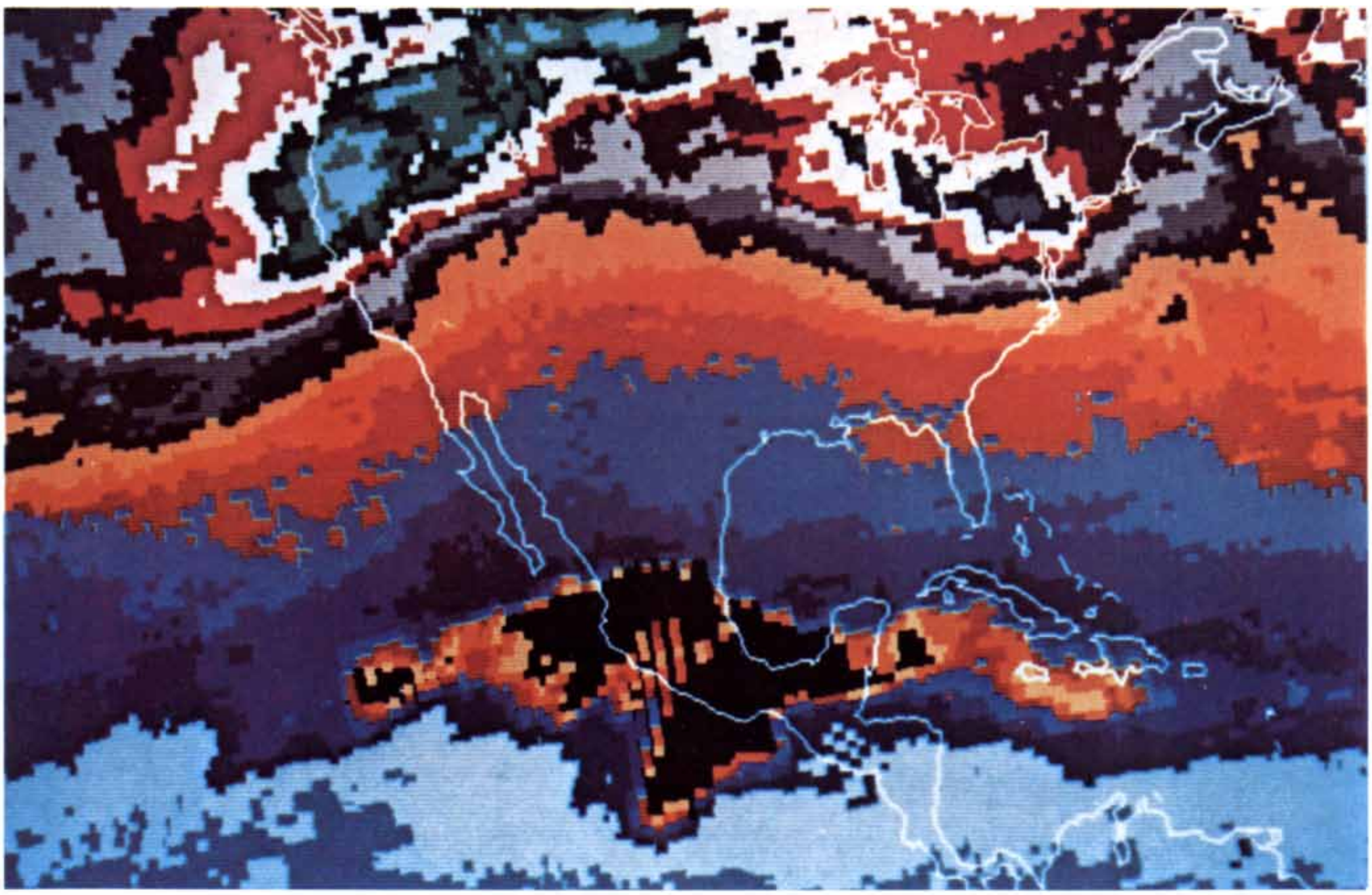
The ash deposited by El Chichón has a sulfur content that is anomalously high for any eruption. Johan C. Varekamp of Wesleyan University and James F. Luhr of the University of California at Berkeley, who were among the first geologists to collect samples and analyze the chemical composition of



TWO SOURCES for the sulfur erupted by El Chichón have been proposed. A deep drill hole near the volcano tapped sedimentary deposits of anhydrite and salt formed by the evaporation of shallow seas some 100 million years ago (*top*). The magma may have assimilated large amounts of sulfur in passing through these layers on its way to the surface. The geologic cross section is based on the work of Robert I. Tilling and Wendell A. Duffield of the U.S. Geological Survey. It is also possible the sulfur was in the rock that melted to form the magma (*bottom*). In some sections of the midocean rifts where new oceanic crust is manufactured vents release sulfur-rich solutions. When the solutions precipitate, they coat the newly formed crust with sulfide deposits. A plate carrying such deposits may have produced a sulfur-rich magma when it was subducted and melted. The sulfur in the magma El Chichón erupted may have come from both sources.



MAP OF TEMPERATURE of the top of volcanic cloud on March 29, 1982, one day after the first major eruption of El Chichón, was generated from data collected by the Advanced Very High Resolution Radiometer carried by the satellite *NOAA 7*. One channel of the radiometer measures the amount of radiation the atmosphere and the earth are emitting at infrared wavelengths. Before the eruption sea-surface temperatures, which might serve as a basis for long-range weather forecasts, were being calculated from the radiometer data. When the volcanic cloud passed under the radiometer, however, the instrument detected primarily the emission from the cloud top. The color in this image varies from green to yellow to red with increasing temperature. The center of the cloud is colder than the periphery because it is higher. Height of cloud can be determined by comparing infrared data with temperature profiles from balloon radiosondes.



GASEOUS SULFUR DIOXIDE injected into the atmosphere by El Chichón shows up clearly in an image generated by the Total Ozone Mapping Spectrometer carried by the satellite *Nimbus 7* on April 5, one day after the last major eruption. The cloud of volcanic sulfur dioxide is the black area extending from Haiti across the Yucatán Peninsula into the Pacific. In this area of the image the sulfur dioxide in the cloud is causing the spectrometer to register falsely high ozone levels. The spectrometer measures the amount of radiation reflected by the atmosphere in the ultraviolet region of the spectrum, where the principal absorbing gas is normally ozone but where sulfur dioxide also has strong absorption lines. The spectrometer also calculates the ozone level by comparing the measured amount with a value computed by an atmospheric model into which information on general atmospheric radiance is fed. The colors in the upper half of the image correspond to ozone levels, varying from blue to brown to green with increasing concentration of ozone. The wave pattern in the ozone distribution (a ridge is centered over the U.S.) is characteristic of the upper tropospheric circulation over the northern U.S. and Canada.

the ash deposits, found remarkably high values of sulfate: up to 2 percent by weight. Under the microscope one could see free crystals of anhydrite ($CaSO_4$), a calcium sulfate mineral rare in volcanic rocks. When the ash was washed, it was found that sulfur was adsorbed on the surface of the ash particles; in other words, some of the sulfur gases released in the course of the eruption had been removed from the volcanic cloud by the ash.

The anomalously high levels of sulfur in the ash might be explained by the presence of sedimentary deposits of sulfur-rich minerals under the volcano. A deep drill hole made near the volcano in the course of exploration for oil had penetrated thick layers of sedimentary anhydrite and salt formed by the evaporation of shallow seas some 100 million years ago. Several investigators have suggested that the magma rising toward the volcano may have assimilated large amounts of sulfur in passing through these layers. The question of the origin of the sulfur remains, however, unsettled. William I. Rose, Jr., of Michigan Technical University argues that the sulfur could have come up from great depths along with the magma. Large quantities of sulfur in the newly formed magma would seem to require an unusual source of sulfur in the subducting plate. At some of the midocean rifts where new crustal plate is manufactured hydrothermal vents emit hot, sulfur-rich solutions that coat the newly formed crust with sulfide deposits. When such crust is subducted and melted, magmas unusually rich in sulfur may be formed.

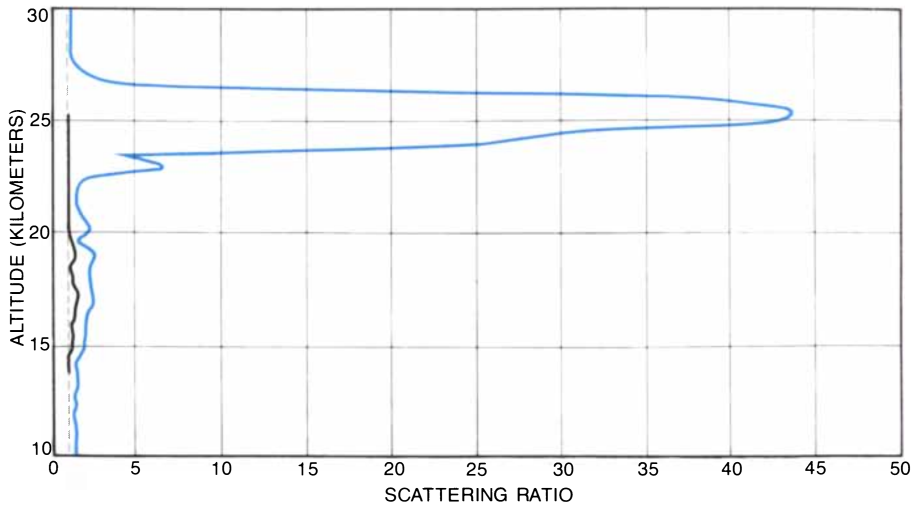
Studies of the isotopic composition of the sulfur in the anhydrite crystals and other sulfur-bearing products of the eruption of El Chichón may provide a way to determine their origin. The sulfur in sediments deposited in the sea should be richer in the heavier of two sulfur isotopes than the sulfur in magma. Initial studies suggest that the sulfur erupted by El Chichón was of mixed origin, but exactly how much sulfur was contributed by each source has not yet been determined.

Given a magma rich in sulfur, how much sulfur is actually released from the molten rock in the eruption and how can the fraction of sulfur released as sulfur gases be determined? Joseph D. Devine of the University of Rhode Island and Sigurdsson have developed a method of measuring the amount of sulfur in magma before an eruption from the composition of inclusions in crystals. Crystals that form in the magma just before an eruption sometimes trap small amounts of the melt within them as they grow. The trapped magma may form beads of glass instead of crystallizing. The composition of the glass reflects the preeruption content of volatile constituents such as sulfur-rich gases. Devine and Sigurdsson measured the sulfur content of the melt inclusions found in the ash at a number of volcanic eruption sites by means of an electron microprobe. They then determined the sulfur content of samples of volcanic glass formed in each eruption. The difference between the two is a measure of the extent of outgassing in the course of the eruption. The total amount of sulfur gases released in an eruption can be estimated from this value and from measurements of the thickness of the ash deposits and the area they cover, which give the total volume of magma erupted.

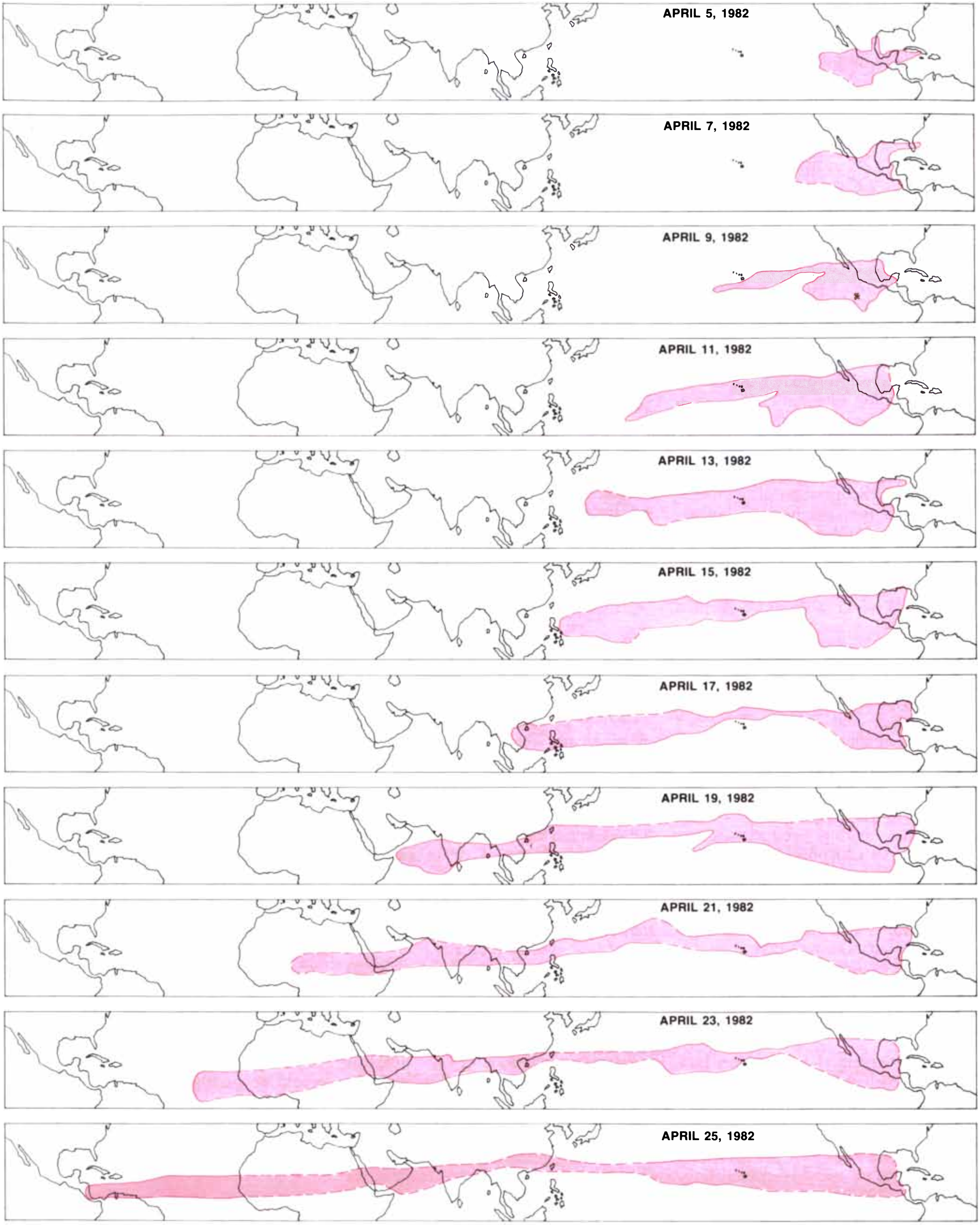
Devine and Sigurdsson have found a good correlation between their estimates of the amount of sulfur gases volcanoes have released and decreases in mean hemispheric temperature; such a temperature decrease is one indicator of the size of the aerosol cloud the volcano produced. Their crystal-inclusion estimate of the amount of sulfur released as gas by El Chichón, however, is much lower than direct estimates of the amount of sulfur in the stratosphere after the eruption. In the case of El Chichón some of the sulfur in the magma was in a form that did not show up in the measurements of the gases in the crystal inclusions. Sulfur may have been sequestered before the eruption in tiny crystals of anhydrite like those found on the ash particles. During the eruption much of this sulfur would have been released as sulfur-rich gases.

Analyses of volcanic emissions show that sulfur is emitted as sulfur dioxide (SO_2) and to a lesser extent as hydrogen sulfide (H_2S) that is soon oxidized to sulfur dioxide. In the stratosphere the sulfur dioxide reacts with hydroxide (OH^-) radicals created by the photodissociation of water vapor to yield intermediate compounds such as hydrogen sulfite (HSO_3), which eventually condense into droplets of sulfuric acid (H_2SO_4) and water. Exactly how the droplets form is not known. Gaseous sulfuric acid will condense on minute seed particles such as volcanic dust or even on ions or small clusters of molecules. Molecules of hydrogen sulfite and water can form droplets simply by clustering, and molecules of sulfuric acid vapor and water may do the same. The photochemical reactions are protracted: the complete conversion of the emitted sulfur gases into aerosol may take weeks or months. New droplets continue to form as others grow and settle out of the stratosphere. The cloud therefore replenishes itself for some time.

The spread of the cloud is determined by stratospheric circulation patterns. The stratosphere is the region between the lower atmosphere (the troposphere) and the upper atmosphere (the mesosphere) where the temperature of the air does not change greatly with altitude. Because of its weak temperature gradient there is little vertical mixing in it, and the volcanic gases tend to stay at the altitude to which they were injected by the eruption, thereby forming layered



SCATTERING PROFILES of the atmosphere show that the eruption of El Chichón produced a stratospheric cloud far denser than the one produced by the 1980 eruption of Mount St. Helens, which ejected a comparable amount of ash but was not sulfur-rich. The profiles were made from data obtained with the lidar (light detection and ranging) system at the National Aeronautics and Space Administration's Langley Research Center in Virginia on July 18, 1980 (*black*), and July 1, 1982 (*color*). The amount of material in the atmosphere is determined from measurements of the amount of light in a laser pulse scattered back to the instrument; the altitude from which the light is scattered is determined from the travel time of the scattered pulse. The scattering varies with the density of the atmosphere as well as with the amount of dust or aerosol in the air. The ratio shown here is the ratio of the amount of light received from a given altitude to the amount the atmospheric molecules normally scatter. The scattering ratio of 45 (at an altitude of 25 kilometers) is the highest ever recorded at the Langley Research Center.



GLOBAL MAPS of the location of the volcanic cloud in the first three weeks after the eruption were made by combining information from the geostationary *GOES East* and *GOES West* satellites and the polar-orbiting *NOAA 7* satellite. The extent of the cloud was determined by examining the imagery the satellites produced from measurements of reflected visible radiation. Just before the eruption the wind in the stratosphere had begun to blow from the east as the stratospheric circulation shifted into its summer pattern. The time it took for the cloud to circle the globe corresponds to a wind speed of about 20 meters per second (about 45 miles per hour). The cloud was confined, however, to a narrow range of latitudes by the circulation pattern. The boundary of the cloud is a broken line where its edge could not be determined exactly. Volcanic clouds generated by other eruptions in the past two years (including Mount St. Helens, Alaid in the U.S.S.R. and Galunggung in Indonesia) dispersed to such an extent in only a few days that they could not be detected in satellite images.

clouds. The stratosphere has its own seasonal weather patterns; in the late spring and summer of the Northern Hemisphere the stratospheric winds in the subtropical location of El Chichón blow generally westward. Stratospheric winds quickly carry aerosols around the world, but it takes longer for the aerosols to spread north and south, a phenomenon that has been thought to be important in delaying the effect of a volcanic eruption on climate.

All three of the plinian eruptions of El Chichón apparently injected sulfur-rich volcanic gases and ash into the base of the subtropical stratosphere. Satellite observations suggest that the eruption on March 28 created a cloud at an altitude of about 20 kilometers, the eruption on April 3 a somewhat lower cloud and the April 4 eruption a massive cloud centered at 26 kilometers. The stratospheric clouds could be seen to spread westward, whereas the short-lived cloud of ash in the upper troposphere spread eastward before it dispersed.

The stratospheric cloud was tracked by ground-based laser range-finding systems and by a variety of instruments carried by satellites. The purpose of the laser ranging systems is specifically to monitor atmospheric aerosols, but some of the satellite instruments had been designed to monitor the wavelengths at which particular gaseous constituents of the air absorb or emit radiation, and it was something of a surprise that the position of the cloud could be determined from the data they collected.

The laser system is known as lidar: light detection and ranging. It measures the amount of light in a laser pulse that is scattered back by the atmosphere; the amount of aerosol can be determined by comparing the amount backscattered with that normally backscattered by the air molecules from a given altitude. As the El Chichón aerosol cloud spread around the globe, high backscattering values were successively recorded by ground lidar stations in the U.S., Japan and Europe.

The most dramatic lidar measurements were those made at the Mauna Loa Observatory in Hawaii right after the eruption. Hawaii is immediately downwind of El Chichón, and the stratospheric cloud had not dispersed appreciably when it first passed overhead. The backscattering values registered by lidar at the observatory were the highest ever recorded since the system went into operation in 1973. The cloud was 140 times denser than the Mount St. Helens cloud.

Initially the cloud was dense enough to be detected by the solar radiation in the visible region of the spectrum it reflected back into space. As the cloud moved around the globe from east to west it was tracked by means of imagery generated from measurements of reflection made by many satellites, including *NOAA 6, NOAA 7, GOES East, GOES West, GMS* (Japan) and *Meteosat* (Western Europe). The tongue of densest aerosol, which was centered at about 20 degrees north latitude, reached Hawaii by April 9, Japan by April 16, the Red Sea by April 20 and had circled the globe by April 26. The movement of the cloud corresponded to an easterly stratospheric wind speed of about 20 meters per second, or about 40 miles per hour. On its first circuit of the globe the cloud spread out somewhat until it covered a 25-degree belt extending from about five degrees north latitude to about 30 degrees.

As the aerosol began to disperse it became increasingly difficult to distinguish it from normal clouds, water vapor, the glint of sunlight on ocean waves and other sources of reflected light. It turned out, however, that satellite instruments that had not been designed to detect the aerosol were able to track it for much longer. For example, the infrared radiometer carried by *Solar Mesosphere Explorer,* which normally measures the radiation emitted by stratospheric water vapor at a wavelength of 6.3 micrometers, was able to detect infrared emission from the volcanic aerosol. One of the most interesting features of the maps generated from data gathered by this satellite is that they show that the bulk of the cloud remained south of 30 degrees north latitude for more than six months after the eruption. A blocking pattern such as this in the stratospheric circulation was unexpected and is all the more intriguing because one analysis of temperature data suggests that El Chichón had its maximum effect on Northern Hemisphere temperatures when most of the aerosol was still confined to latitudes only a third of the way from the Equator to the pole.

Most of the satellite measurements of the cloud could be taken only as indicators of the amount of material in the stratosphere, including dust and water vapor as well as sulfuric acid aerosol. It was discovered soon after the eruption, however, that the amount of sulfur dioxide gas in the volcanic cloud could be determined from measurements taken by the Total Ozone Mapping Spectrometer carried by the satellite *Nimbus 7.* The spectrometer normally determines the level of ozone absorption in the ultraviolet region of the spectrum by comparing the amount of ultraviolet the stratosphere is reflecting with a value obtained from a theoretical model of atmospheric absorption, scattering and radiative transfer.

Sulfur dioxide gas decreases the reflectance of the stratosphere at ultraviolet wavelengths, indicating higher apparent ozone levels, but the contribution of sulfur dioxide can be separated from that of ozone and other gases if the reflectance at different wavelengths is examined. Sulfur dioxide absorbs at two wavelengths near .3 micrometer, whereas the other gases do not. On the basis of spectrometer data and estimates of the extent of the El Chichón cloud Arlin J. Krueger of the Goddard Space Flight Center of NASA estimates that the volcano injected 3.3 million tons of gaseous sulfur dioxide into the stratosphere. The data also suggest that all the sulfur dioxide had been converted into sulfuric acid by July, 1982, about three months after the eruption.

That estimate of the initial amount of sulfur dioxide gas is less than estimates by other groups of the amount of sulfuric acid aerosol in the stratosphere. This raises the question of whether the sulfur was injected in the form of other gases. One possibility is that some of the hydrogen sulfide emitted by the volcano was not oxidized before it was injected into the stratosphere.

Samples of the cloud taken by aircraft and balloons provided a more intimate look at its composition and the dynamics of its growth and dispersal. The aerosol was collected in flights made as part of the Lyndon B. Johnson Space Center's program for collecting cosmic dust and by a group at the Ames Research Center. David S. McKay and Ian D. R. MacKinnon of the Johnson Center found that the cloud still held a significant amount of ash in May and early June of 1982. About 85 percent of the ash consisted of angular glass fragments. The glass shards were coated with droplets of sulfuric acid, and some of them bore small sulfur-containing crystals, probably anhydrite or another sulfate mineral that had sublimated onto the ash during or after the eruption.

In May the ash particles had an average diameter of between three and six micrometers. By July the larger ash particles had settled out of the stratosphere and the average diameter had decreased to between one and two micrometers. The samples also showed that aggregates of ash and sulfuric acid aerosol were forming that were low enough in density to reach much larger sizes before settling out. Aggregate particles as large as 80 micrometers in diameter with densities as low as .1 gram per cubic centimeter were found.

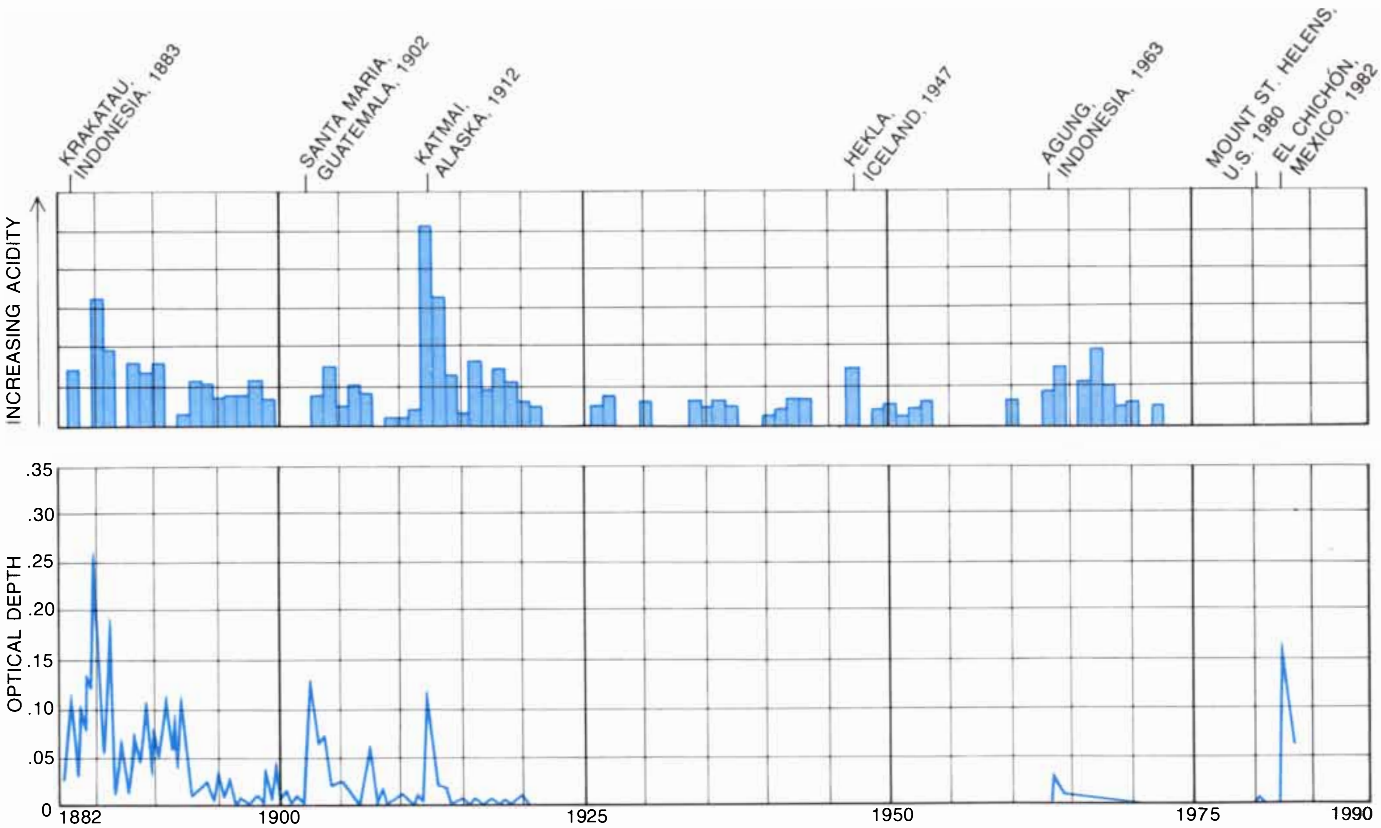
The best information about the dynamics of aerosol growth and removal in the cloud was gathered by balloon-borne instruments. David J. Hofmann and James M. Rosen of the University of Wyoming have been sampling stratospheric aerosols by balloon since 1971; they sent aloft photoelectric particle counters that could detect condensation nuclei (minute particles on which vapors can condense) as small as .02 micrometer in diameter and could distinguish between liquid droplets and solid particles by heating them. Plots of the size of the aerosol droplets in the cloud in May made by Hofmann and Rosen show two peaks, one near .04 micrometer in diameter that is due to newly forming droplets and one near 1.4 micrometers that is due to droplets that had grown by accumulating sulfuric acid vapor. Data from later balloon flights in August and October of 1982 show a decrease in the number of very small droplets; the production of condensation nuclei had apparently stopped by that time. Even as late as October, however, there was little decrease in the number of large droplets; growing droplets were apparently still replacing those that had settled out of the stratosphere.

The balloon-borne measurements of the average density and size of the droplets and estimates of the extent of the cloud have been utilized to estimate the total amount of aerosol in the stratosphere. According to these data, a month after the eruption there may have been as much as 20 million tons of sulfuric acid aerosol in the stratosphere. By April, 1983, a year after the eruption, the amount of aerosol had decreased to less than eight million tons.

Stratospheric aerosols affect the global radiation budget mainly because they absorb and backscatter incoming solar radiation, although they also absorb infrared radiation emitted by the lower atmosphere. Absorption and backscattering give rise to atmospheric optical effects and decreases in surface temperatures. The absorption of the infrared in the incoming solar radiation shows up mainly as an increase in stratospheric temperatures.

After the eruption of El Chichón the temperature of the equatorial stratosphere rose by about four degrees Celsius; the temperature was the warmest ever recorded since continuous measurements of stratospheric temperature were begun in 1958. Soon after the eruption atmospheric optical phenomena like those that followed the eruption of Krakatau in 1883 and of Mount Agung in 1963 began to be reported, and some were still being seen in late 1983. For example, Aden B. Meinel and Marjorie P. Meinel of the University of Arizona reported unusually brilliant and prolonged sunrises and sunsets in Tucson beginning in late April. Long windrows of aerosol were visible in the upper atmosphere on May 7, when the main body of the cloud apparently passed northwest of the city. For some time the sky over Arizona was a whitish blue instead of the normal bright blue characteristic of the U.S. Southwest.

The sunsets that indicate the presence of a stratospheric aerosol typically begin with a lavender glow that appears high over the horizon and changes gradually to yellow and orange. After the sun sets there is often a strong red afterglow from sunlight reflected by the aerosol cloud that higher above the horizon shades into purple. The stratification of the cloud may become apparent as horizontal streaks. Several observers also reported seeing Bishop's rings, another optical effect typical of volcanic aerosols. Such rings, named for the Reverend Sereno E. Bishop, who first described them after the Krakatau eruption, are halos around the sun in which the normal order of the colors is reversed, with the red on the outside.



PEAKS IN ACIDITY of ice in the Greenland ice sheet correlate with periods when the atmosphere was reported to be less clear than normal. The acidity of annual layers in an ice core taken from the ice sheet in the early 1970's (*top*) is a good measure of the amount of sulfuric acid aerosol in the stratosphere at the time. The record of the optical depth (opacity) of the stratosphere in the Northern Hemisphere (*bottom*) was compiled from the reports made by observatories on the intensity of sunlight or starlight. Most of the peaks can be matched with contemporaneous eruptions, some of which are indicated. The period between 1920 and 1960 was one of comparatively few large explosive eruptions that produced stratospheric clouds. The latitude of an eruption must be taken into account in interpreting the records. For example, Katmai, which erupted in 1912, and Hekla, which erupted in 1947, may have had a large effect on the acidity of Greenland ice because of their high latitude. The acidity of the ice core was measured by Claus U. Hammer of the University of Copenhagen. The record of optical depth up to 1970 was compiled by Owen B. Toon and James B. Pollack of NASA's Ames Research Center.

The most important effect of the El Chichón aerosol cloud on climate will probably be a decrease in mean global temperatures. Volcanic aerosols typically cause a drop of a fraction of a degree C. Although such an apparently small decrease can be significant for climate, separating it from natural fluctuations can be difficult. This is particularly true in the case of the eruption of El Chichón, because 1981 was the warmest year on record in the Northern Hemisphere (about .5 degree C. above the long-term average) and temperatures had begun to decline in early 1982 even before the eruption.

Investigators who have examined the statistical connection between volcanic eruptions and decreases in mean temperature find that temperatures averaged over a hemisphere drop by about .3 degree C. for one to three years after major volcanic eruptions. More specific predictions of the effect the eruption of El Chichón will have on temperatures were made from the change produced by the eruption in the optical depth (opacity) of the stratosphere and from several different mathematical models of climate. Measurements of optical depth can be made by comparing with normal values the amount of light received by a photometer or a pyrheliometer. On the basis of the drop in temperature caused by previous eruptions that gave rise to similar changes in the stratospheric optical depth it was predicted that El Chichón would cause Northern Hemisphere temperatures to drop by about .4 degree by late 1983. The climate model developed by Alan Robock of the University of Maryland predicts a cooling of .5 degree by next winter (the winter of 1984–85). Other climate models predict a similar decrease in temperature, although the delay in the drop varies from model to model with the assumptions about the spread and dissipation of the volcanic cloud.

One recent analysis of the temperature decreases following volcanic eruptions (including the eruption of El Chichón) arrived at some startling conclusions. P. M. Kelly and Chris B. Sear of the University of East Anglia focused their analysis on mean monthly temperatures in the Northern Hemisphere after the major eruptions of the past 100 years. It shows that the maximum temperature decrease occurs as soon as two months after a Northern Hemisphere eruption and is followed by a second decrease about a year later, after which temperatures slowly recover. The response of Northern Hemisphere temperatures to eruptions in the Southern Hemisphere may be delayed by seven or eight months.

Kelly and Sear's analysis suggests that the eruption of El Chichón, at 17 degrees north latitude, had its maximum effect on monthly temperatures (a drop of .2 degree C.) as early as June, 1982, two months after the eruption. Moreover, since this maximum came before the cloud had blanketed the globe and while it still held significant amounts of dust, the inference is that a volcanic aerosol cloud can affect temperature in a wider belt of latitudes than it covers, and that both dust and sulfuric acid contribute to the cooling.

A drop in surface temperature may not be the only effect the eruption of El Chichón has had on global climate. By altering the vertical profile of atmospheric temperatures it may have altered atmospheric circulation patterns, perhaps contributing to the intense El Niño event of last winter (the winter of 1982–83). An El Niño event is an abrupt change in the patterns of both atmospheric and oceanic circulation in the equatorial Pacific. Although the El Niño pattern is recognizable around the Pacific, it is commonly associated with the sudden change in ocean currents off the west coast of South America, with which it typically begins. There the southeast trade winds periodically weaken. As the winds die down, the water in the western Pacific surges eastward, cutting off the Humboldt Current that normally flows northward along the coast of South America, a disruption that in turn produces anomalous weather over the continent.

The aerosol cloud could conceivably have caused the El Niño event by warming the upper atmosphere in the Tropics. The warming of the stratosphere would have reduced the temperature difference between it and the surface of the earth, which could have led to a weakening of the atmospheric circulation and of the oceanic circulation it drives. The 1982–83 event is particularly suspect because it did not follow the normal pattern of development. It was unusual in that it began to develop in May instead of October or November and in that the change in circulation patterns occurred across the entire Pacific at about the same time instead of moving westward from the coast of South America.

Of the nine El Niño events since 1950 only two have not followed the normal pattern: the 1982–83 event and the 1963 event that coincided with the eruption of Mount Agung, the last previous eruption to produce a dense aerosol cloud affecting tropical latitudes. On the other hand, some of the phenomena that precede an El Niño event, such as the drop in atmospheric pressure at Tahiti and Easter Island in relation to pressures at Australia and Indonesia, seem to have begun before the eruption of El Chichón. For the present any causal connection between volcanic eruptions and atypical El Niño events remains unproved, and the question may not be settled until future events are examined.

Now that El Chichón has established the importance of the sulfur content of a volcanic eruption to its atmospheric impact, will it be possible to go back and establish the connection between the sulfur content of past volcanic eruptions and climate changes? Unfortunately andesitic eruptions such as El Chichón leave little in the way of a geologic record. Andesitic volcanoes tend to erupt a small volume of magma and thus leave thin, easily eroded ash deposits. It will therefore be difficult if not impossible for geologists to assemble a complete record of past sulfur-rich eruptions from studies of ash deposits. Nevertheless, frequent small eruptions may have as important an effect on climate as one large eruption. Research with climate models suggests that periods of increased volcanic activity could potentially have accounted for some of the relatively long-term decreases in global temperatures on record.

One promising method of detecting past sulfur-rich eruptions is measuring the acidity of annual layers of ice in the earth's polar ice sheets (in Greenland and Antarctica). The total amount of sulfuric acid aerosol in the atmosphere can be calculated from the acidity of an ice layer if the latitude of the eruption is known, although volcanoes close to the ice sheet may have an inordinately large effect on acidity because their sulfurous gases arrive through the troposphere. Comparison of climate data with the acidity of layers in ice cores has brought to light interesting correlations. For example, relatively acidic ice was deposited throughout the period from about A.D. 1350 to 1700 known as the Little Ice Age. The correlations suggest that sulfur-rich volcanism may indeed be an important modulator of long-term trends in global climate.

The ash deposited by El Chichón is still being analyzed; the stratospheric aerosol is still being sampled by balloon, aircraft and satellite, and ground-based instruments are still detecting it. The information has already deepened understanding of the link between volcanoes and global climate, and further analyses of the data and their application to modeling and predicting the effect of volcanic eruptions on climate will occupy atmospheric scientists and volcanologists for years to come. The eruption provided a much needed test case for the hypothesis that the sulfur content of a volcanic eruption has a more important effect than the volume of ash it ejects on the size of the aerosol cloud it creates. What conclusions further analyses will yield remains to be seen, but the eruption of El Chichón has provided a collection of data unrivaled by anything previously available on the effects of volcanic eruptions on world climate.

High-Energy Collisions between Atomic Nuclei

Heavy nuclei traveling at 95 percent of the speed of light are now made to collide with target nuclei. The debris includes anomalons, mysterious fragments that collide again much sooner than expected

by Wm. C. McHarris and John O. Rasmussen

On the first manned flight to the moon, in 1969, the astronauts aboard *Apollo 11* reported a curious phenomenon: when they closed their eyes to go to sleep, they saw an occasional pinpoint flash of light. It was soon surmised on the earth that the flashes were caused by heavy atomic nuclei among the cosmic rays impinging on the spacecraft. The amount of energy deposited by each such nucleus inside an astronaut's retina was apparently greater than the minimum required to stimulate the retina's light-sensitive cells.

Soon after the flight this hypothesis was tested. At the Lawrence Berkeley Laboratory of the University of California the Bevatron, a high-energy proton accelerator built in the early 1950's, had recently been modified so that it could accelerate nuclei heavier than individual protons. Edwin M. McMillan, who was then the director of the laboratory, volunteered to put his head in the path of a low-intensity beam of high-energy heavy nuclei. He saw the same kind of flashes the astronauts had seen.

High-energy nucleus-nucleus collisions since then have exhibited properties more difficult to explain. For example, the collisions give rise to what are called anomalons: nuclear fragments that collide again within a distance that is puzzlingly short. Some of the properties of the collisions are said to manifest a "leakage" of the force that binds nucleons (protons and neutrons) into a nucleus. Indeed, it is said the collisions may reveal the activity of quarks, the unseen theoretical constituents of protons and neutrons. It seems certain that the extremes of temperature and pressure generated by high-energy nucleus-nucleus collisions will create new states of nuclear matter, such as the state the universe was in when its age could be measured in millionths of a second.

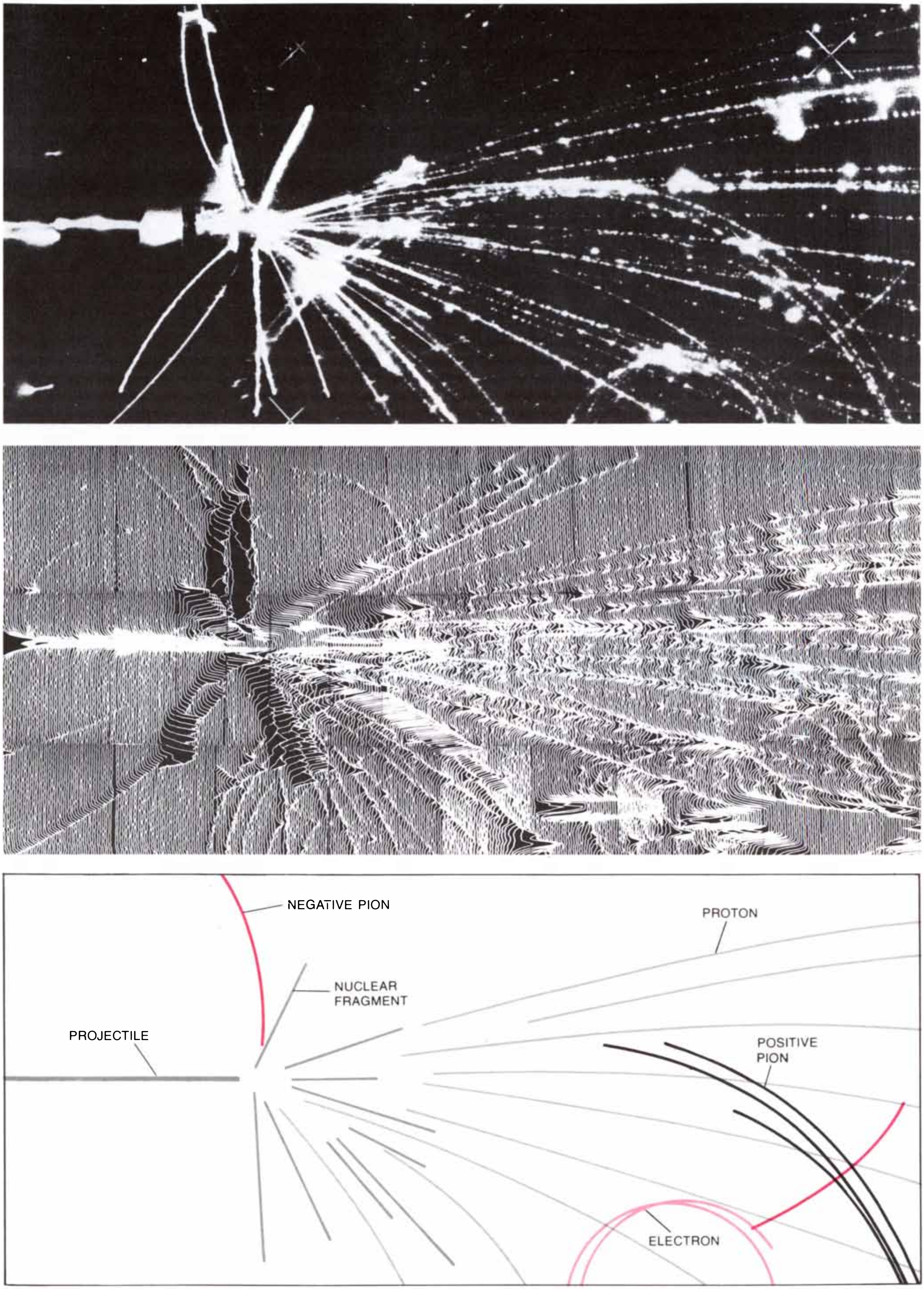
What distinguishes high-energy nucleus-nucleus collisions so that the flashes they can cause in the eye were discovered in space and not, say, by someone who was trying to sleep on an airplane? The most important difference lies in the rate at which a collision deposits energy in matter. The rate is proportional to the square of the electric charge of the "projectile" nucleus participating in the collision. Thus a cosmic-ray iron nucleus, which when it is stripped of the electrons that would otherwise be part of an iron atom becomes an ion with a positive charge of 26 units, deposits its energy at 26^2, or 676, times the rate of a cosmic-ray proton with the same velocity. In addition the heavy nuclei in cosmic rays are likely to break up when they collide with nuclei of the gases in the upper atmosphere. This ensures that no heavy nuclei from cosmic rays penetrate into the lower atmosphere.

Well before the flight of *Apollo 11* the rate at which heavy ions deposit energy had led medical scientists to advocate high-energy heavy-ion accelerators. They reasoned that a single hit by such an ion might kill a cancer cell instead of merely disrupting its activity. (The biological damage done by most forms of radiation is caused by the production of peroxides and other chemical radicals in cells.) By the early 1970's several heavy-ion accelerators were in place, but their beams had ranges in tissue of a few millimeters at the most. Boosting the ions to greater energies would require accelerator rings like the ones serving particle physicists.

At that time physicists were reluctantly preparing to close down old accelerators such as the Bevatron. Instead a new use arose: a vacuum transfer line was constructed so that the Heavy-Ion Linear Accelerator, or HILAC, also at Berkeley, could send its beam into the Bevatron. In this way the ions emerging from HILAC with an energy of 8.5 MeV per nucleon (that is, 8.5 million electron volts for each proton and neutron in the nucleus) could be further accelerated to energies as great as 2 GeV (two billion electron volts) per nucleon. The combined machines, named Bevalac, could for the first time accelerate heavy nuclei in the laboratory to velocities as great as 95 percent of the speed of light and thereby give them energies approaching those of cosmic rays. These velocities and energies are said to be "relativistic" because bodies moving at appreciable fractions of the speed of light behave in ways predicted by the special theory of relativity. Meanwhile at the Joint Institute for Nuclear Research at Dubna in the U.S.S.R. an accelerator named the Synchrophasotron began to accelerate carbon nuclei to twice the energy the Bevalac could give them. The beam, however, was less intense.

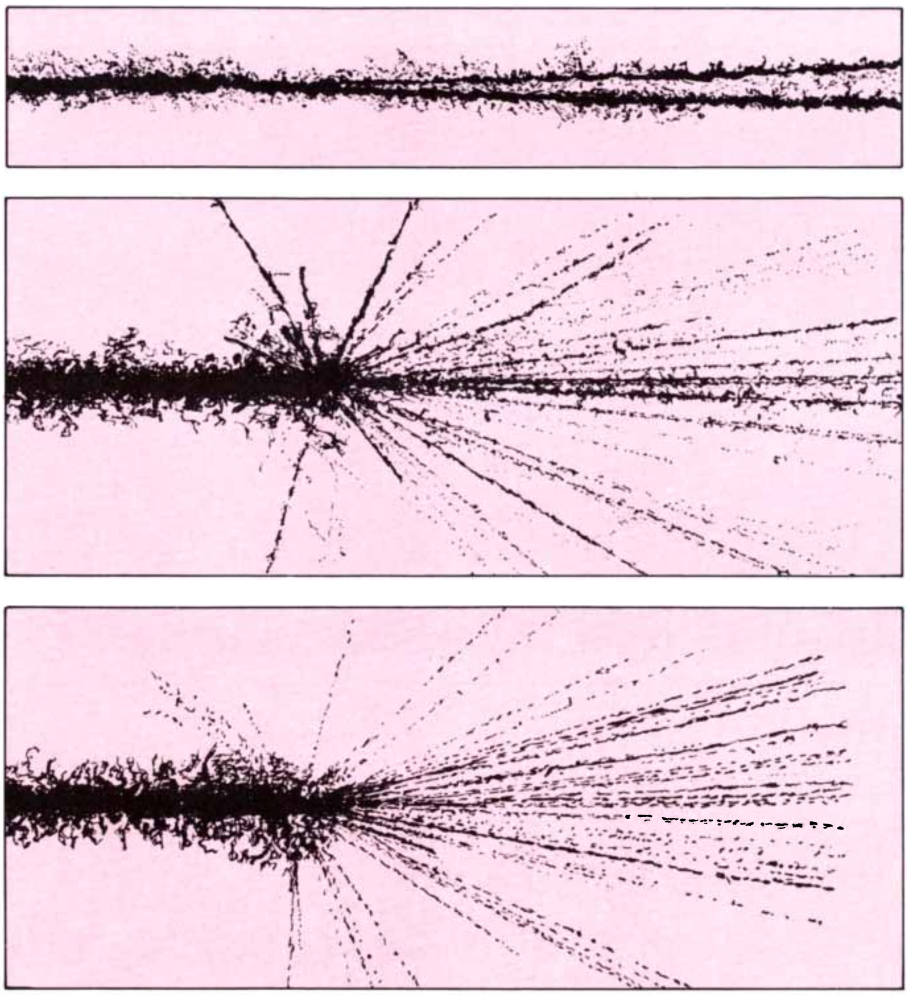
Concurrent with these developments

NUCLEUS-NUCLEUS COLLISION was photographed (*top*) in a streamer chamber at the Lawrence Berkeley Laboratory of the University of California. In such a chamber the passage of a charged particle leaves behind a track of ionized gas. A charge-coupled device (in effect a computer-controlled television camera) reconstructed the collision (*middle*). Peaks indicate the intensity of light, which represents both the ionization density of the tracks and the distance of each track from the camera. A map (*bottom*) identifies some of the particles. The nucleus of an argon-40 atom entered at the left. Its kinetic energy was 72 billion electron volts (72 GeV). About .3 meter into the chamber it struck the nucleus of a lead atom in a lead oxide target. Most of the detectable debris flung forward by the collision consisted of protons; their paths were bent slightly downward by a magnetic field in the chamber. A few of the particles in the debris were pions, which were created from the energy of the collision. Some positive pions crossed the tracks of the protons; a negative one traveled back toward the left. The heaviest fragments were aggregations of protons and neutrons. They left short, highly ionized tracks.


NEGATIVE PION
PROTON
NUCLEAR
FRAGMENT
PROJECTILE
POSITIVE
PION
ELECTRON

an innovation in theoretical physics was opening a deeper level of understanding of the composition of matter: the quark theory of elementary particles was taking shape as an account of the internal structure of protons and neutrons. It was already understood that the atomic nucleus consists of protons and neutrons in a set of orbital "shells." Now it was affirmed that protons and neutrons each consist of three quarks. The proton consists of two "up" quarks and a "down" quark; the neutron consists of two "down" quarks and an "up" quark. The quarks have fractional electric charge: in units of the charge of the proton the charge of the down quark is $-1/3$ and the charge of the up quark is $+2/3$. Already the quark theory has answered some basic questions about the admixture of states entailed when a proton and a neutron combine to form a deuteron (a hydrogen-2 nucleus). Moreover, as we shall see, the quark theory has yielded insight into some otherwise mysterious occurrences in high-energy nucleus-nucleus collisions.

It is gratifying to note how the new capabilities for producing high-energy heavy-nucleus beams fill a gap in knowledge. Consider a chart in which the energy one can give particles in the laboratory is plotted against their mass. High-energy physicists have progressed along the energy axis, employing beams of individual particles: generally beams of electrons, protons and also mesons, the class of particles intermediate in mass between the electron and the proton. Conversely, nuclear scientists have progressed along the mass axis. This leaves the area distant from both of the axes as a terra incognita waiting to be explored. Let us examine what is already known to be there.



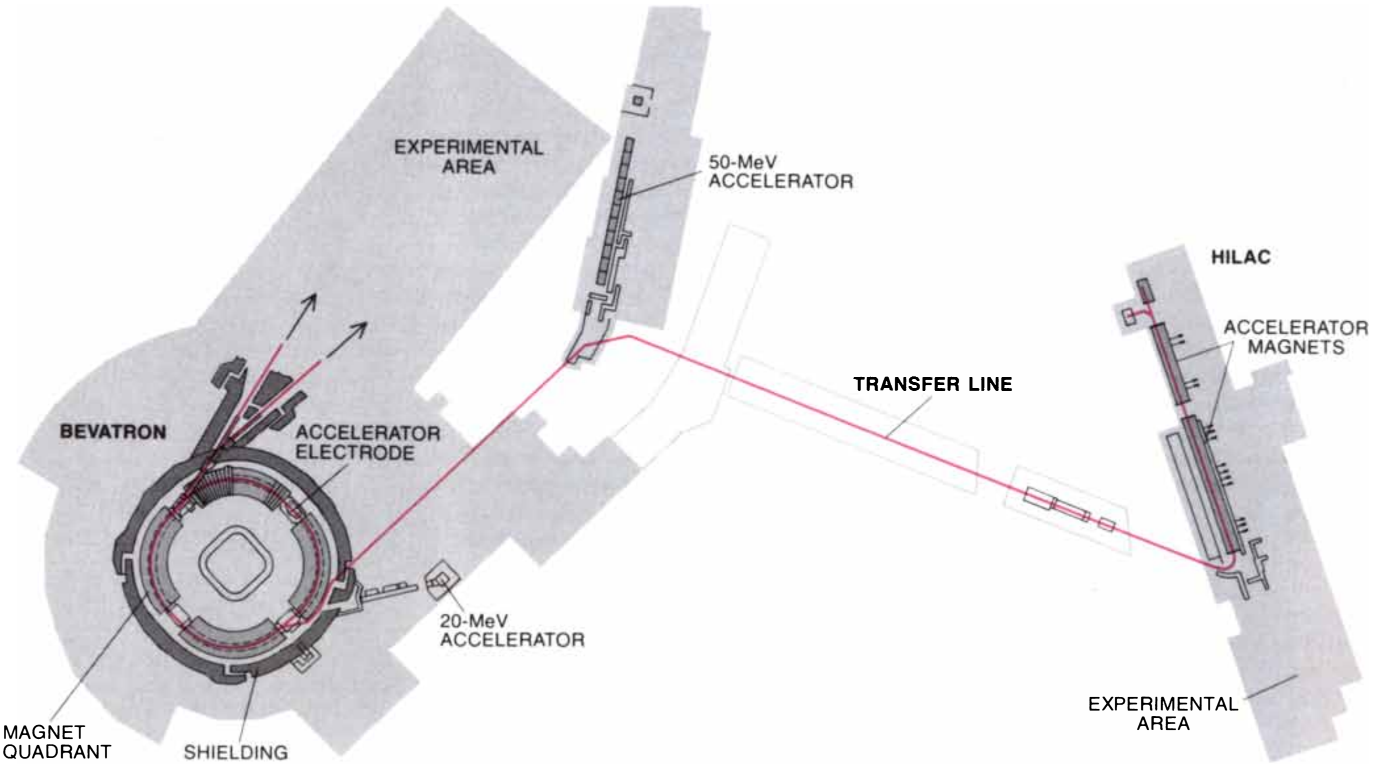
THREE TYPES of high-energy nucleus-nucleus collisions were recorded, also at Berkeley, by placing photographic emulsions in the path of a nuclear beam of uranium 238 after each nucleus had been accelerated to an energy of 228 GeV. In a peripheral collision (*top*) a uranium nucleus passed close enough to a nucleus in the emulsion so that the two interacted, if only electromagnetically. The interaction caused the uranium to fission into two nuclear fragments. The target nucleus remained intact. (The squiggles emanating from the tracks are caused by energetic electrons stripped off numerous atoms in the emulsion.) In a hybrid collision (*middle*) a uranium nucleus struck a nucleus in the emulsion but not head on. Both nuclei disintegrated. Some of the darker tracks were made by fragments of the target nucleus flying off at large and even backward angles. In a central collision (*bottom*) a uranium nucleus engulfed a nucleus in the emulsion. The agglomerate then disintegrated and its momentum threw most of the fragments forward. The tracks in each collision have been magnified by some 400 diameters.

Nucleus-nucleus collisions at relativistic energies can be classified in three categories: peripheral, hybrid and central. Although the division is somewhat arbitrary, any given collision can generally be assigned to one category or another. In peripheral collisions one nucleus, a "projectile," grazes a second nucleus, a "target," or passes close enough to the target so that the two interact at least electrically. A large piece of the projectile continues forward. In hybrid collisions the projectile strikes the target, but not head on. For example, the projectile may shear off half of the target. Large parts of both nuclei disintegrate. In central collisions the projectile hits the target more or less head on. The amalgam of the two then disintegrates, and the momentum of the projectile tends to throw the fragments forward.

The peripheral collisions reward a close examination. The results of the collisions can be recorded by a variety of detectors: photographic emulsions, in which the passage of a charged particle through a sheet of silver bromide lays down a track of metallic silver; bubble chambers, in which the passage of a charged particle through a superheated liquid under pressure lays down a track of bubbles (in essence, the charged particle causes local boiling), and streamer chambers, in which the passage of a charged particle through a gas across which a high voltage is applied lays down a track of streamers, or electric discharges. All such detectors confirm that the fragments of a peripheral collision have little change in direction from that of the projectile.

What these detectors alone cannot show is the momentum or velocity of the fragments. For that a complementary technique is brought into play: the collision is made to take place in a magnetic field, which bends the trajectories of the fragments if they carry electric charge. The higher the momentum of a fragment, the less its trajectory is bent; this bending can be measured from the visual record of such a track or by an array of particle detectors positioned around the collision. The measurements reveal that the fragments of a peripheral collision have almost the same velocity as the projectile has. The small variations that do occur can be related to the size of the fragment: the maximum variation is shown by fragments with half the mass of the projectile. Moreover, the variation in the velocity is on the order of the variation in velocity that a nucleon has as it jiggles about in a nucleus. (This jiggling is called Fermi motion and is much the same in all nuclei.)

The fragments of peripheral collisions at lower energies do not show such



HEAVY-NUCLEUS ACCELERATOR at Berkeley is actually two accelerators connected by a beam-transfer line. The Bevatron (*left*), a ring-shaped accelerator 100 feet in diameter, was built in the early 1950's. It was designed to accelerate protons to 6 GeV. By the 1970's it was obsolescent; then it proved to be capable of accelerating heavy nuclei. Today the Heavy-Ion Linear Accelerator, or HILAC (*right*), accelerates nuclei to an energy of 8.5 million electron volts (8.5 MeV) per nucleon (neutron or proton). The Bevatron boosts them to 2 GeV per nucleon and velocities as great as 95 percent of the speed of light. The combination of Bevatron and HILAC is referred to as Bevalac.

simple behavior. At lower energies the colliding nuclei can be in contact for a period of time many orders of magnitude longer than the time it takes a nucleus to vibrate or the time it takes a relativistic particle to travel the diameter of a nucleus. Hence in even a grazing encounter two low-energy nuclei have opportunity to fuse or partially fuse. In effect they form a liquid drop of nuclear matter whose energy is sufficient to boil off nucleons in a number of directions at a number of velocities.

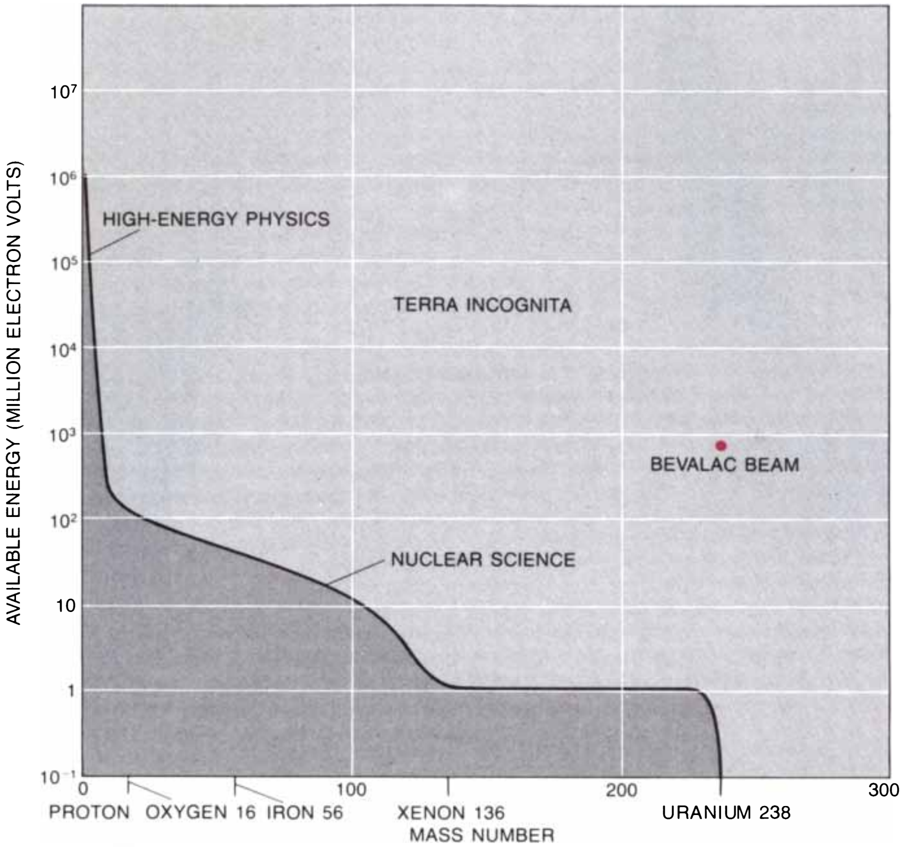
Much of the energy of a high-energy nucleus-nucleus collision is typically converted into mass by the copious production of exotic, unstable subatomic particles, notably pions, which are the lightest of the mesons. They have an eventful history. In the 1930's Hideki Yukawa predicted their existence as "virtual particles" inside the nucleus. He argued that the strong force (the force that binds nucleons into a nucleus) is carried by particles the nucleons exchange. The exchange would be undetectable, and so a pion appearing spontaneously as a fluctuation in the state of a nucleus would be "virtual," not real. Nevertheless, the mass of a pion, expressed as the equivalent in energy, would be 139 MeV.

A remarkable property of the pions generated by high-energy nucleus-nucleus collisions is that in peripheral collisions many of the negative pions are electrically focused: the positive charge of the forward-flying fragment attracts the negative charge of each pion, so that it tends to acquire a velocity and a direction much like those of the fragment. This could have practical applications. In Los Alamos, Vancouver and Zurich particle accelerators derive a beam of pions from the collisions of intense beams of high-energy protons; with a heavy-ion accelerator one might produce pion beams more economically. The heavy ions of choice would be the nuclei of silicon 28. It happens that for pions and for silicon-28 nuclei passing through matter the rate of energy loss as a proportion of the total energy is identical. Hence it makes no difference at what point a collision gives rise to pions inside matter: a silicon-28 nucleus traversing a target can collide with a nucleus anywhere inside the target, and the resulting negative pions will emerge well collimated in velocity with the pions produced by other collisions.

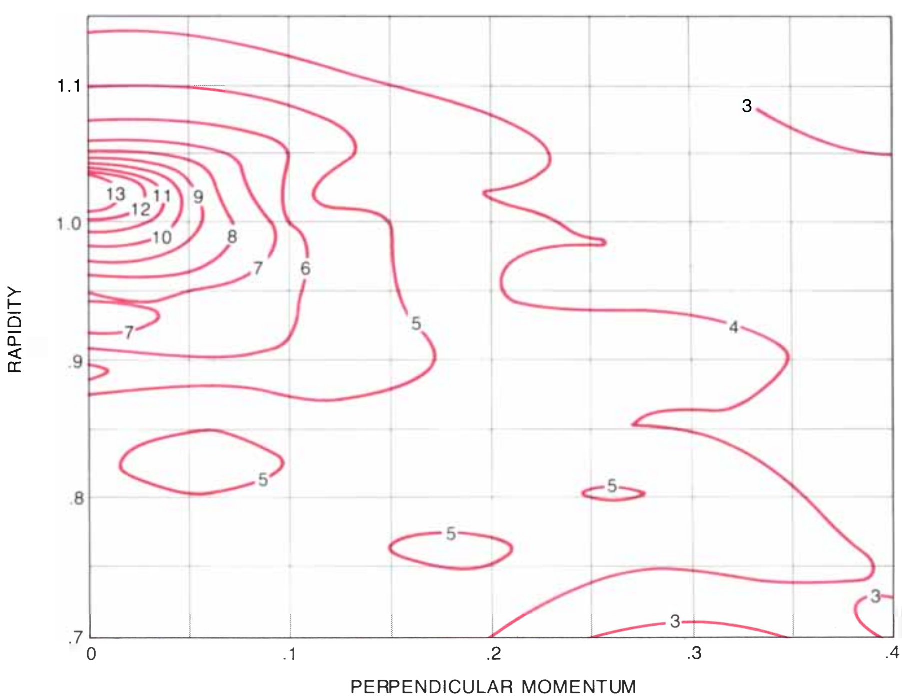
A second product of peripheral high-energy nucleus-nucleus collisions (and the occasional hybrid collision) is more mysterious than the pion. It is the anomalon: a projectile fragment whose mean free path (the distance it travels before colliding again) is anomalously short. Perhaps the anomalon can best be described as a nucleus that interacts before its time. Over the past two years experimental evidence has been accumulating that anomalons indeed exist. The experiments are tedious, however: they require that sequences of collisions be traced, most often by the examination of photographic emulsions under the microscope. Moreover, the very nature of the evidence, consisting as it does of a small number of seemingly premature interactions, invites contention. Today the nuclear-science community is split into camps of believers and disbelievers. Why have anomalons generated so much attention and controversy? If they do exist, they could well have far-reaching implications not only for nuclear science but also for elementary-particle physics and thus for theories concerning the basic structure of matter.

Anomalons were first noted in the early 1950's in photographic records of cosmic-ray collisions. Every once in a while a heavy-nucleus cosmic ray would collide with a nucleus in a photographic emulsion, creating a "star" (a burst of photographic tracks), and then a fragment of the collision would create a second star within a much shorter distance than one would predict: perhaps only one or two centimeters rather than the normal 10 or more.

One can readily imagine the quandary of investigators trying to explain this curious finding. Suppose we repeatedly conduct the experiment of driving



NEW SCIENCE explored with high-energy nucleus-nucleus collisions is evoked when the energy imparted to particles in the laboratory (*vertical axis*) is plotted against their mass (*horizontal axis*). The highest energies have been imparted to the least massive particles (mostly subatomic), leaving a terra incognita that occupies most of the chart. The energy of the Bevalac beam of uranium-238 nuclei (*colored dot*) places it well within the unknown territory.



PION FOCUSING is an attribute of high-energy nucleus-nucleus collisions; thus an argon-40 beam at 21 GeV striking a potassium chloride target at the Bevalac produces negative pions "focused" so that they have nearly the velocity of the beam. Basically the heavier nuclear fragments of the collisions, which have a positive charge, attract the negative pions. The chart plots "rapidity," a measure appropriate for speeds approaching the speed of light, against momentum perpendicular to beam direction. Contour lines show relative numbers of negative pions.

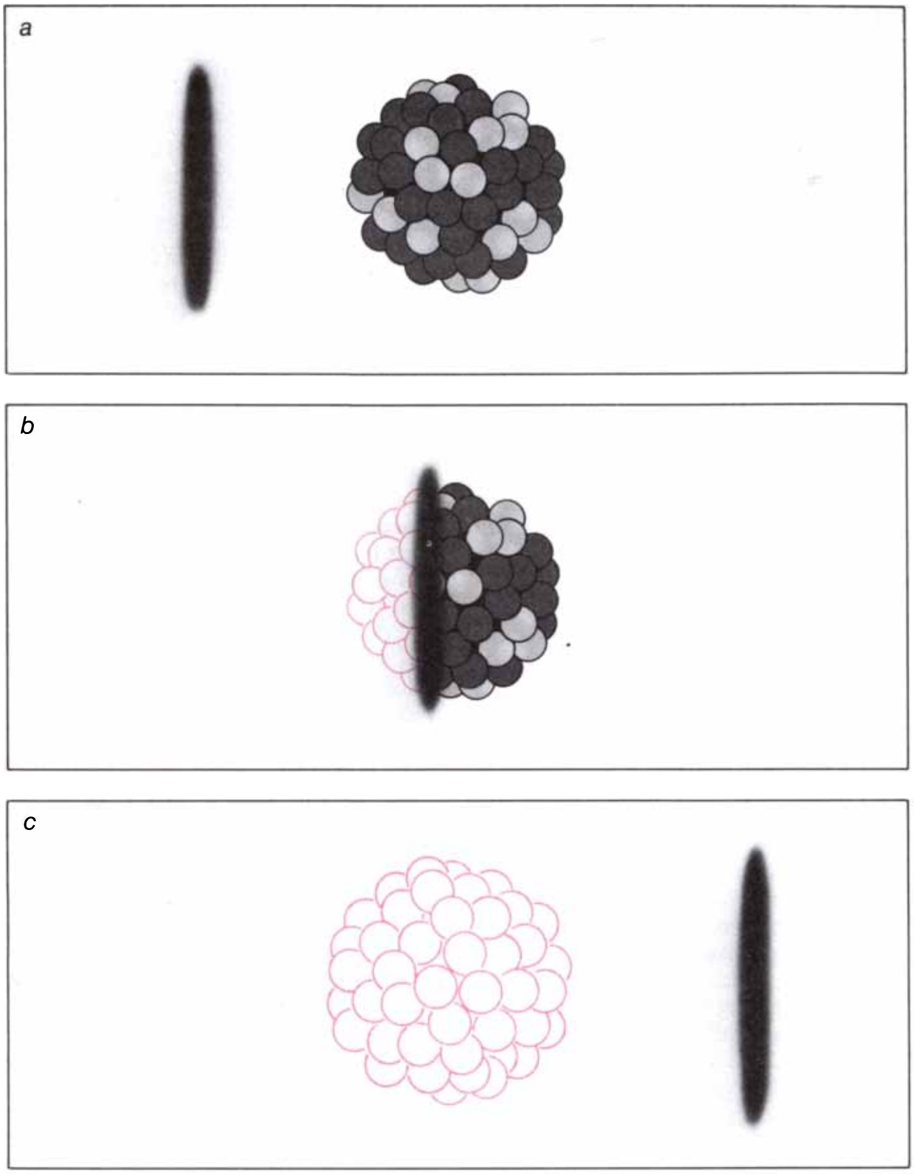
an automobile down a highway at high speed and then running it off the road into a forest of randomly spaced large trees. Sometimes we crash almost immediately, sometimes we make it between many trees and therefore travel a longer distance. After many trials an average distance is established, depending on the width of the car and the average spacing between the trees. In a subsequent set of trials the automobiles are outwardly the same. The only difference we know of is that they have been repaired after one or more of the earlier crashes. We are astonished to find they crash into trees after an average of one-tenth the distance the new cars traveled. Clearly something unexpected has happened. After one crash the cars act as if they were much wider. Such cars would qualify as anomalons.

Observations of anomalons in photographic records of cosmic-ray collisions were reported sporadically throughout the 1950's and 1960's. Then the advent of beams from the Bevalac made records of controlled high-energy nucleus-nucleus collisions available in sufficient quantity so that the properties of anomalons could be gauged. Three groups, one at the Lawrence Berkeley Laboratory, one at the National Research Council of Canada in Ottawa and one at the State University of New York at Buffalo, examined photographic emulsions that had been exposed to Bevalac beams of various nuclei at various energies, ranging up to iron-56 nuclei at 1.88 GeV per nucleon. A group at the University of Minnesota reexamined the cosmic-ray records. As of two years ago, the time of the First International Workshop on Anomalons, held at Berkeley, the combined efforts had led to the following results.

First, anomalons did indeed have mean free paths that were anomalously short. Among the projectile fragments traced by the various groups a small percentage (roughly 6 percent, although the reported figures varied from 3 to 10 percent) had mean free paths no greater than 2.5 centimeters, or less than a tenth the length of the paths exhibited by the projectiles. Second, anomalons were observed only when the energy of the projectiles was greater than 1 GeV per nucleon. (Little work had been done below 1.8 GeV per nucleon, and so the precise energy threshold was uncertain.) Third, anomalons had to have a lifetime on the order of at least 10^{-10} second; otherwise they would not have traveled 2.5 centimeters. That lifetime is roughly 13 orders of magnitude longer than the time required for events such as the vibration of a nucleus. Fourth, if anomalons decayed spontaneously in flight, the decay would have to be accompanied by the emission of only electrically neutral particles, namely neutrons or gamma rays (high-energy quanta of electromagnetic radiation), because no charged particles were observed emanating from the paths of anomalons across photographic emulsions. Fifth, the properties of anomalons appeared to be independent of the amount of charge of the fragment. Some tenuous evidence did suggest, however, that anomalons with only one or two units of charge never form.

Sixth, and perhaps strangest of all, anomalons exhibited a memory effect that can be expressed as "once an anomalon, always an anomalon." More precisely, it seems a fragment emerging from a star produced by an anomalon has a greater chance of being an anomalon itself than it would have if it emerged from a star produced by something other than an anomalon. Either anomalons have a good chance of surviving a collision or something about anomalons makes it relatively easy to remake one in a collision.

It is quite a set of properties. Over the past two years groups in laboratories around the world (Egypt, Germany, India, Italy, Sweden, Switzerland and the U.S.S.R.) have been examining photographic records from both the Bevalac and the Synchrophasotron; thus the statistics on anomalons have been improved. Moreover, some groups have begun to employ new detectors, notably plastic-track detectors. Certain plastics are susceptible to damage by heavy ions; hence stacked sheets of such a plastic are exposed to ions and then chemically etched. The etching enlarges the damage sites into tiny conical pits that trace the path of the ion. The great advantage of the technique is that for ions of a given velocity the diameter of the pit is a measure of the charge, and the pit diameters can be measured by computerized scanners. At the Second International Workshop on Anomalons, held in Berkeley last summer, the participants agreed that the list of six properties was still viable, although still not fully proved. In addition several groups reported that the electric charge of anomalons seems always to be a multiple of the unit electric charge. Thus a seventh property can be added: anomalons do not have fractional electric charges. In other words, anomalons appear not to include stray quarks.



RELATIVISTIC CONTRACTION is a further attribute of the collisions. Here a hypothetical ultrahigh-energy collision is shown. A uranium nucleus with an energy of one trillion electron volts per nucleon approaches a uranium nucleus at rest (*a*). The projectile, which travels at more than 99.999 percent of the speed of light, appears to be a disk; its contraction is predicted by the special theory of relativity. Its encounter with the target lasts for only 10^{-22} second, far too short a time for the two to come to an equilibrium; thus the projectile passes through (*b*). Nevertheless, it raises the target to a temperature of more than one trillion degrees (*c*). Such temperatures may approach those of the universe immediately after the big bang.

Devising a satisfactory theoretical explanation of anomalons is at least as difficult and confusing as the experimental work has been. Many theories have been proposed, but none so far has covered the facts so well that it is accepted as probably being correct. The most spectacular theories are those in which anomalons are manifestations of quarks. Such theories rest on an analogy. Consider the chemical bond. It is a short-range, saturable interaction. That is, it binds only particular numbers of neighboring atoms. Nevertheless, it arises from a long-range force: electromagnetism. The bond is short-range only because the electric charge of the nucleus and the charge of the electrons in an atom are equal and opposite, so that they cancel when the atom is at a distance.

Now consider the bond that holds nucleons together. It too is short-range and saturable. Its normal range is less than the size of a nucleus. Yet it too may arise from a long-range force. The bond holding nucleons together would then be short-range only because each of the three quarks in a nucleon has a strong-force charge, called "color," and the colors normally cancel.

Anomalons furnish one of the first real grounds for speculation that the strong force may have a long-range leakage. Therefore attempts have been made to visualize anomalons as assemblages of quarks in which the colors do not cancel, so that a long-range strong interaction facilitates collisions and shortens a nuclear fragment's mean free path. The attempts have ranged from the "demon deuteron" (a deuteron taken to consist of three quark pairs rather than two quark triplets) to great doughnut-shaped masses of quarks. No attempt has been compatible with all the items in the list of anomalons' properties. Not the least of the difficulties is the finding that anomalons are created at energies less than the ones at which one would expect an extensive rearrangement of quarks or much production of quarks.

The explanations of anomalons that do not rely on quarks have been similarly diverse. They range from bubble nuclei (hollow balls consisting of nucleons) to quasi-molecular nuclei (dumbbell-shaped aggregates arising when two nuclei join but do not amalgamate). The idea is to make the anomalon a swollen or distended object, so that its chance of being struck in a collision is increased. The explanations also include a proposal that we ourselves have devised, in which anomalons consist of negative pions strongly bound to neutron-rich nuclear fragments. The pions could draw neutrons outward in the nuclear matter by convection and thereby extend the surface, giving the fragment protuberances. We argue that pions are produced in copious quantity by nucleus-nucleus collisions and that the negative ones are focused around the larger fragments. In addition we note that pions are attracted to protons and neutrons by the strong interaction. The attraction of negative pions is known to be much greater for neutrons than it is for protons.

A problem with our model is that any contact of negative pions with protons makes the matter disintegrate in times that are short compared with the lifetime of anomalons. We are encouraged all the same by recent theoretical calculations suggesting that a body consisting of a few negative pions and a few neutrons might be sufficiently long-lived to be detectable. No such "pineuts" have ever been seen, but searches at various heavy-ion accelerators may have a chance of finding them. Indeed, it is beginning to appear that negative pions bound in nuclei may be a phenomenon not peculiar to a theory of anomalons. Certain data gathered by particle detectors imply that pions have orbited nuclear fragments and then have been flung away. Accordingly we speculate that the anomalon is a pineut halo surrounding a nuclear core, with centrifugal force on the pineuts keeping them away from the protons at the core. Perhaps the requisite neutron-rich nucleus (or neutron-rich regions in a nucleus) explains the memory effect observed in anomalons. A neutron-rich nuclear fragment would tend in subsequent collisions to make further neutron-rich fragments.

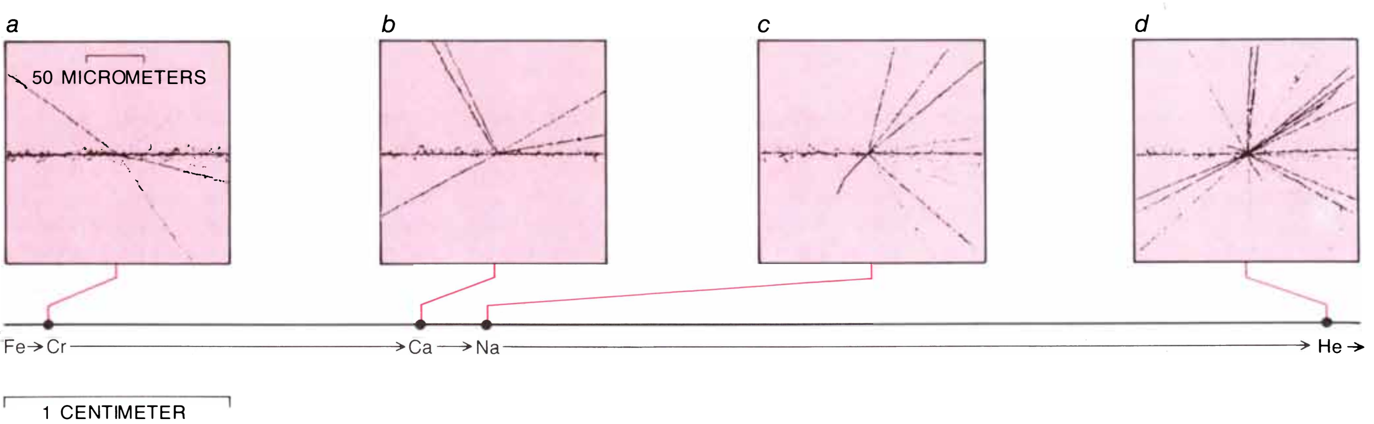
Can a connection be seen between negative pions bound to neutrons and a long-range strong-force leakage? So far the efforts to treat simple nuclei as clusters of quarks have centered on multiples of quark triplets. Such multiples virtually guarantee that the color charges will cancel. What about calculations based, say, on a negative pion and two neutrons? That is the simplest pineut. It would consist of eight quarks: five down quarks, two up quarks and one anti-up quark. (A negative pion is taken to be a pairing of a down quark and an anti-up quark.) Such a cluster seems less likely to guarantee canceling color charges; it ought to be pursued.

One aspect of high-energy nucleus-nucleus collisions makes them particularly promising for cosmological research: the impressive rate at which a high-energy heavy nucleus can deposit energy in a target nucleus suggests that the target nucleus could be heated to a temperature approaching that of the universe a short time after the big bang. Such temperatures could not be reached by bombarding heavy nuclei with beams of electrons or protons; the lightness of these projectiles means they would dissipate only a small fraction of their total kinetic energy as they passed through even the heaviest nucleus. In contrast, a central collision between two uranium atoms could stop the projectile, which thus would dump all its kinetic energy.

A German-U.S. collaboration at the Bevalac has looked into the extent to which nuclei at top Bevalac energies really stop in central collisions. They have constructed an elaborate spherical shell of multiple detectors that determine both the energy and the rate of energy deposition of particles emerging from a target. A further wall of detectors is positioned a few meters "downstream." Together the "ball and wall" incorporate more than 1,000 detectors. The data show that calcium nuclei are too small to stop each other completely but that niobium nuclei are large enough.

The concept of a nucleus having temperature calls for an explanation. In general temperature is simply a measure of the random motions in a group of particles. The energy of a nuclear collision is transformed, at least in part, into the random motions of the nucleons in a nucleus; hence it is appropriate to speak of a nuclear temperature. Somewhat more specifically, it is characteristic of matter in thermal equilibrium, whether it consists of nucleons, of molecules or of macroscopic particles in random motion, that the distribution of velocities falls off exponentially. In other words, high-energy constituents are relatively scarce. The average kinetic energy is proportional to the temperature, T, of the matter, expressed in degrees Kelvin; the constant of proportionality is Boltzmann's constant, k. Nuclear scientists prefer not to deal with T, which can assume unwieldy values (billions of degrees). They employ as a surrogate kT, the average kinetic energy per nucleon.

Let us give some examples. The temperature of the surface of the sun (T) is about 6,000 degrees K. The corresponding kinetic energy kT is .5 electron volt. This means the atoms at the surface of the sun have less kinetic energy than electrons get when they are accelerated by the voltage of a flashlight battery. (One realizes that the kinetic energy a



ANOMALONS are a mysterious aspect of high-energy nucleus-nucleus collisions; they are nuclear fragments that make subsequent collisions within a distance that is anomalously short. This illustration gives an example. An iron nucleus with an energy of 1.88 GeV per nucleon collided with a nucleus in a photographic emulsion, producing a chromium nucleus (*a*). After a flight of 1.62 centimeters the chromium nucleus collided, producing a calcium nucleus (*b*); one would have expected an average flight of 10 centimeters or more. After .3 centimeter the calcium nucleus collided, producing a sodium nucleus (*c*); again the flight was anomalously short. After 3.75 centimeters the sodium nucleus collided, producing most notably a helium nucleus (*d*). The sequence of collisions was recorded at the Bevalac.

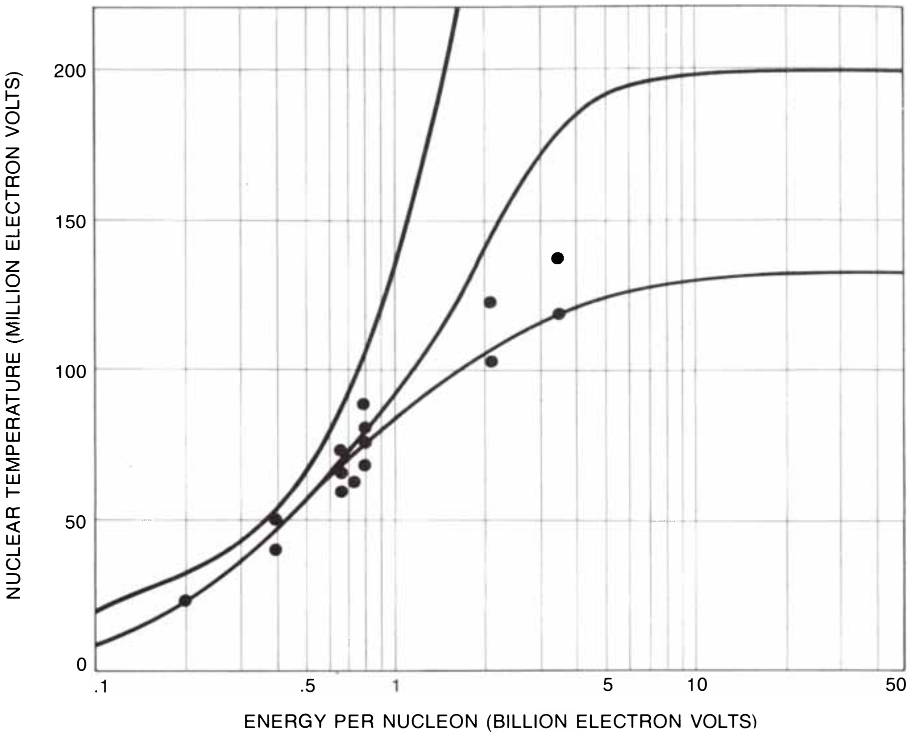
particle gets when it is heated is typically not impressive.) The temperature at the center of the sun is thought to be 15 million degrees. The corresponding kinetic energy is 1,300 electron volts, or about the energy of electrons striking the screen of a picture tube in a television set. The temperature at the center of a star as it becomes a supernova is thought to be a few billion degrees. The corresponding kinetic energy approaches 1 MeV, or about the energy imparted to electrons in the most powerful electron microscopes. That is the upper limit of the temperatures generated by processes active in the universe today. In the early universe much greater temperatures are thought to have prevailed.

Can anything special happen to nuclear matter when it is raised to extremely high nuclear temperatures? Consider what happens to ordinary matter when it is heated. Initially most such matter is solid. As it is heated it melts, becoming a liquid; then it boils, becoming a gas. Next its molecules are torn apart into atoms. Finally electrons are torn from the atoms, so that electrons and ionized nuclei, some of them completely stripped of electrons, move independently in a gas of charged particles: a plasma.

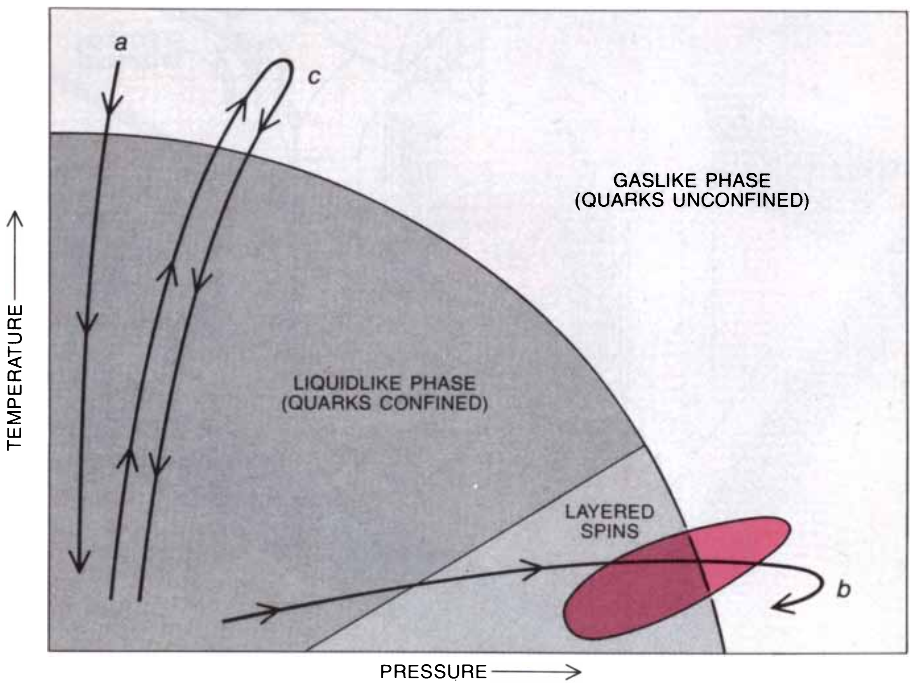
Nuclear matter is thought not to have a solid phase; it more nearly resembles a liquid. Theorists predict, however, that high temperature (or great pressure) would bring on a phase transition, in which nucleons break up into quarks and gluons: the exchange particles that bind quarks together. The result would be analogous to a plasma: individual quarks would range throughout the nucleus instead of moving only inside quark-triplet "bags."

These phases of nuclear matter are of course quite speculative. Yet one can argue that in addition to the normal, liquid phase and the predicted plasma phase, a "pion condensate" phase can be achieved at high nuclear densities without high nuclear temperatures. In such a phase the nucleons in a nucleus would array themselves in layers so that the spin, or intrinsic angular momentum, of the nucleons in one layer would be aligned in one direction and the spin in the next layer would be aligned the opposite way. ("Pion condensation" is a bit of a misnomer and has to do with the calculated stability of pion binding in dense nuclear matter.) There has been speculation that the center of neutron stars might be nuclear matter in a pion-condensate phase. Unfortunately the efforts to produce such a phase on the earth are hampered by the problem of giving nuclear matter a high density without giving it high temperature.

How hot and dense does nuclear matter have to get to become a quark-gluon



NUCLEAR TEMPERATURES attained in nucleus-nucleus collisions at Bevalac energies serve as a probe of various models of nuclear matter under extreme conditions. The curves are theoretical calculations that embody different assumptions about nuclear matter. One assumption (*bottom curve*) is that the heating of nuclei converts all the available energy into matter: namely hyperons, excited states of nucleons. The hyperons can then decay into nucleons and pions. The opposite assumption (*top curve*) is that the heating takes place without changing the nucleons into other particles at all. A third assumption (*middle curve*) is intermediate. The temperatures in the chart are all expressed as their kinetic-energy equivalent in MeV.

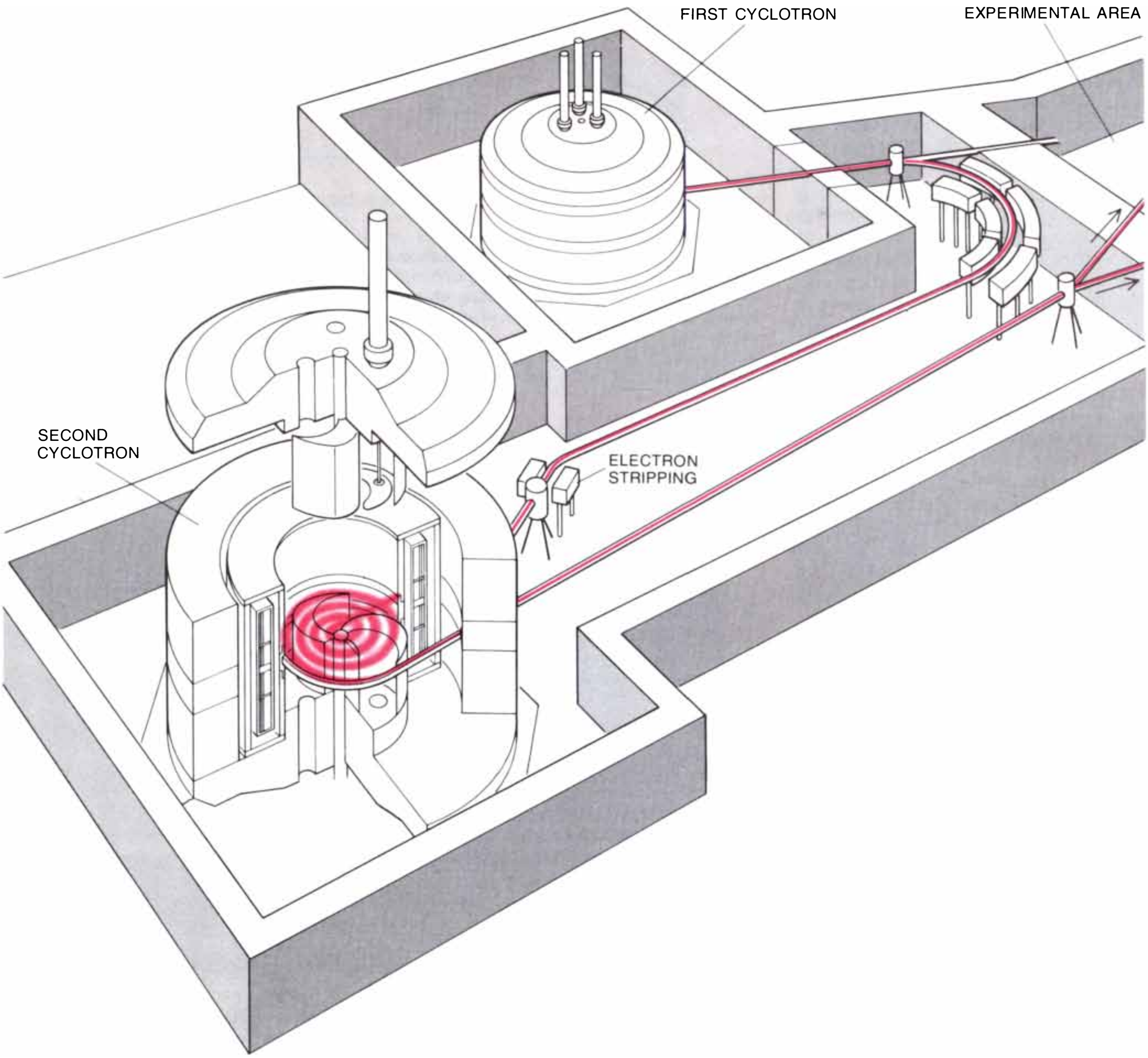


HYPOTHETICAL PHASE DIAGRAM shows the forms of nuclear matter that may emerge under widely varying conditions of temperature and pressure. Normal nuclei are in a liquid-like phase (*dark gray*): they are a "fluid" of neutrons and protons. At extremely high temperatures and pressures these nucleons may dissociate into quarks and gluons, their supposed constituents (*light gray*). At combinations of high pressure and intermediate temperature a third phase may exist, in which the spins of the nucleons are organized into layers (*medium gray*). One trajectory in the chart follows the evolution of nuclear matter in the early universe (*a*). A second trajectory follows the evolution of nuclear matter in a supernova explosion (*b*). A third trajectory follows the evolution of nuclear matter in a central nucleus-nucleus collision (*c*). A different realm might be accessible to peripheral nucleus-nucleus collisions (*color*).

plasma? To calculate an answer with any certainty is a task now occupying powerful computers. The results are not yet in, but the best estimates so far suggest that the temperature at normal nuclear densities would have to approach 200 MeV. Such conditions are probably just beyond the reach of existing heavy-ion accelerators. Still, experiments are being mounted just in case to look for indications that a quark-gluon plasma might already have been created.

A number of still newer accelerators may provide better chances yet. Some have just been completed and others are being built in laboratories around the world. Thanks to superconducting magnets and dependable microelectronic controls they can be more compact and less costly than their advocates had feared. Nevertheless, the extremely high-energy nuclear beams needed for the exploration of nuclear phases and their cosmological implications remain prospective. Both the Lawrence Berkeley Laboratory and the Brookhaven National Laboratory have made preliminary proposals for heavy-ion accelerators (and the associated ion-storage rings) in the TeV (trillion electron volt) range.

For the time being the available accelerators yield a glimpse of the science that deals with relativistic nucleus-nucleus collisions. It is a hybrid science, relying on knowledge of two fields: high-energy subatomic-particle physics on the one hand and nuclear science on the other. In recent decades the two have been veering apart, and today they have little interaction. Yet the juncture of two climates can generate not only great turbulence but also great fertility. The juncture of two sciences, particle physics and nuclear science, for the study of nucleus-nucleus collisions has a similar potential.



SUPERCONDUCTING CYCLOTRONS at Michigan State University are the first such devices; a pair of them will accelerate heavy nuclei to energies as great as 200 MeV per nucleon. One cyclotron, already completed, will accelerate slightly ionized atoms along a spiraling path between the propeller-shaped poles of a superconducting magnet only 10 feet in diameter. (The acceleration region itself is only 52 inches in diameter.) The resulting beam of ions will be directed through a foil, which strips the atoms of many remaining electrons, and into the second cyclotron, which will accelerate the ions further. The pair of machines will be modest in size and in cost compared with recent accelerators but will generate beams whose energy per nucleon is greater than that of most heavy-ion accelerators.

Strong medicine for feverish health-care costs.

A new blood analysis system which embodies Kodak's expertise in chemistry, optics, physics, and electronics is helping clinical labs control costs. Its flexible software design is bringing new ease and reliability to the process of blood analysis. It can improve laboratory efficiency, increase productivity, and help to put health-care costs on the road to recovery.

The new system incorporates numerous technological inventions with such intricate design and engineering that it can perform a full range of routine tests—including kinetic enzyme, as well as potentiometric and colorimetric analysis—in operator-preferred sequence. If this sounds like quite an accomplishment, it is.

The Kodak Ektachem 700 analyzer incorporates 20 megabytes of hard-disk memory, relies upon Kodak's dry layer-coated Ektachem clinical chemistry slides like the one shown,* and it produces hard-copy results for physicians' evaluation. In just 5 minutes!

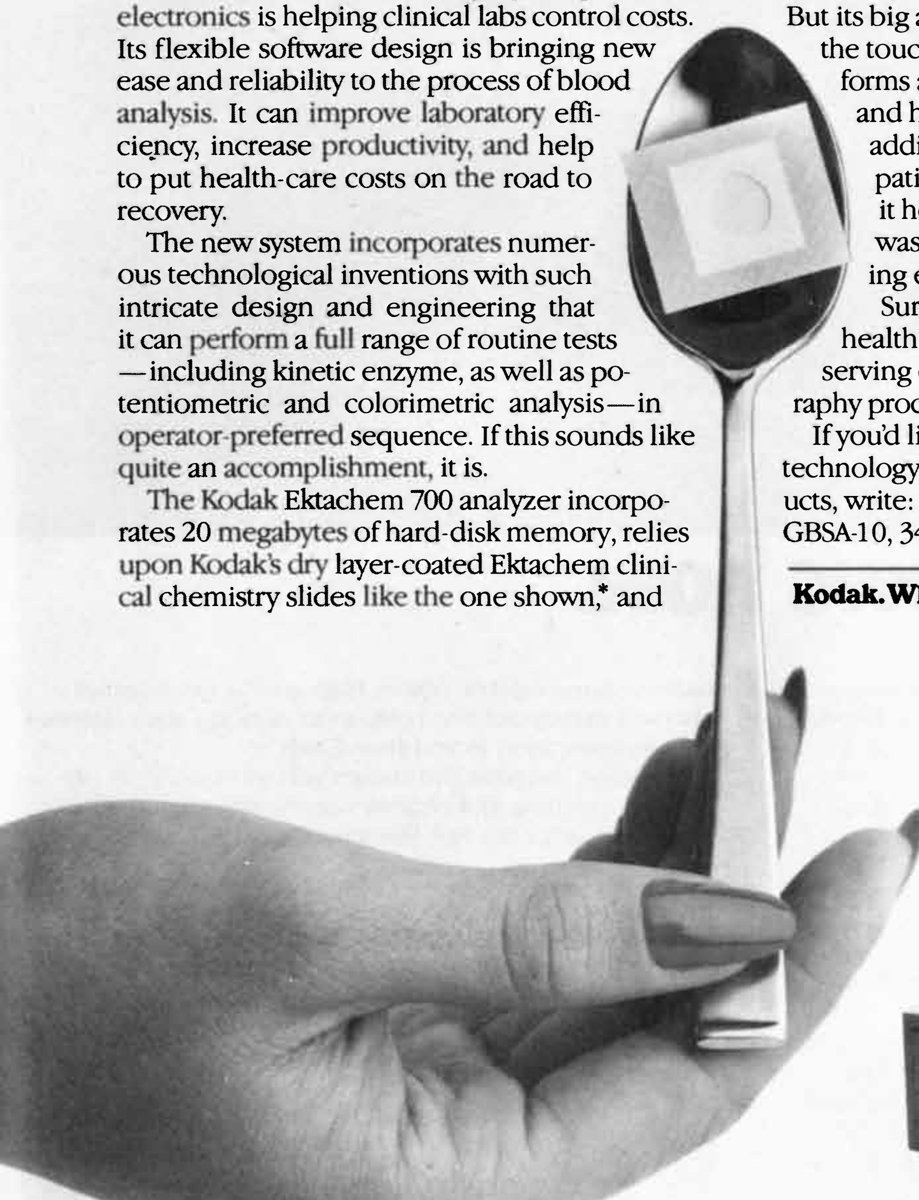
But its big advantage is selective testing. At the touch of a CRT screen, this analyzer performs any combination of one to 26 assays and has the potential to report up to 7 additional calculated results on a single patient sample. That's a big plus, because it helps to eliminate wasted tests and wasted time, and contributes to operating economy.

Surprised that we're so committed to health care? You shouldn't be. We've been serving diagnostic medicine with radiography products for more than 80 years.

If you'd like technical papers dealing with the technology involved in Kodak Ektachem products, write: Eastman Kodak Company, Dept. GBSA-10, 343 State Street, Rochester, NY 14650.

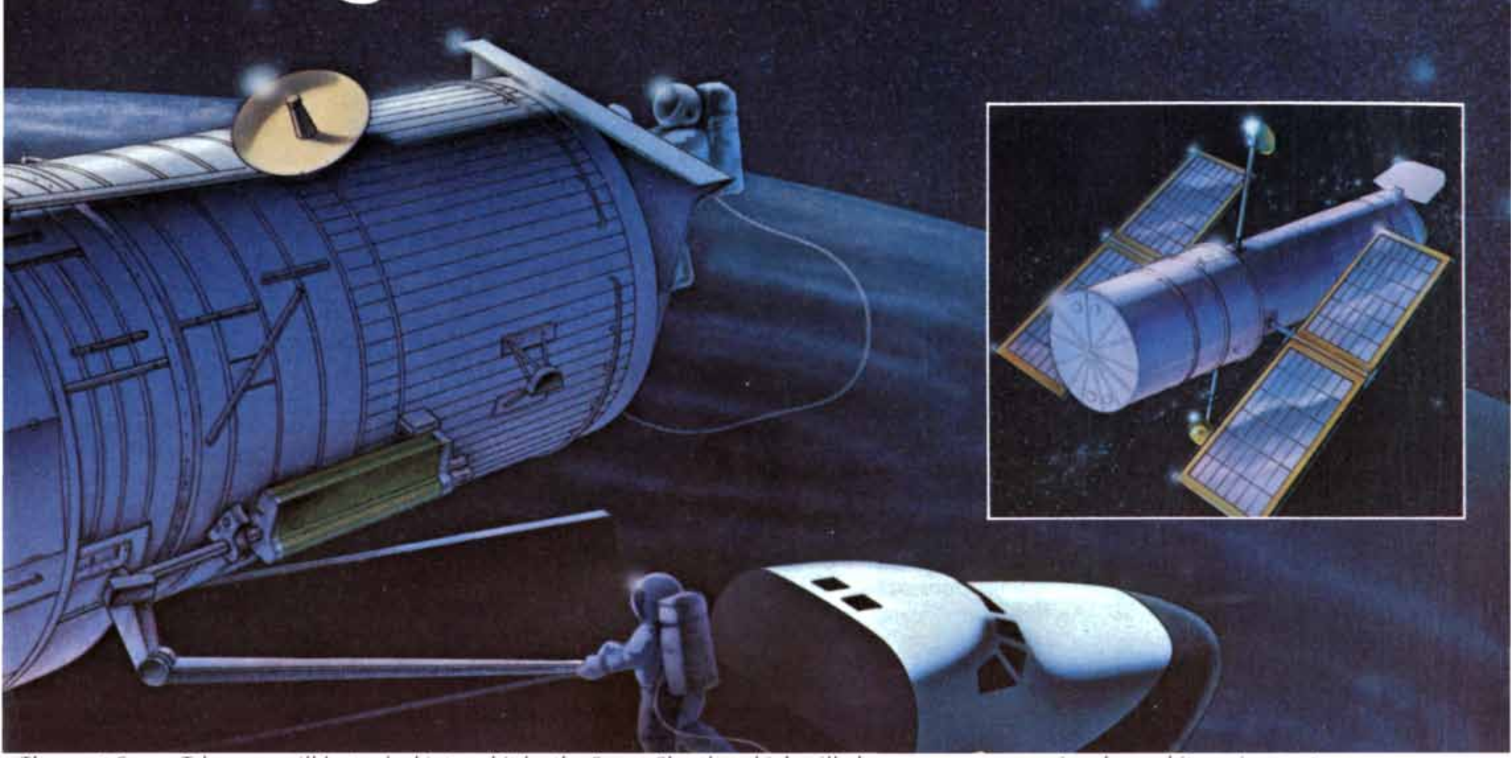
Kodak. Where technology anticipates need.

Kodak



*With Ektachem slides, complex sequential reactions can be carried out in ways not possible in solution chemistry. Multilayer coatings offer domains for multiple reactions within the single-use slide. In some layers, interfering substances can be trapped or altered; in other layers, reactions can be run which produce measurable signals.

the edges of the universe.



The great Space Telescope will be tucked into orbit by the Space Shuttle, which will also serve as an occasional on-orbit service center. INSET: The telescope deployed.

and a photometer (being built by several institutions).

In their interference-free environment, the mirrors and instruments will help produce and transmit images with resolution 10 times greater than now possible with ground-based telescopes.

Fixing mankind's steadiest gaze.

Those mirrors and instruments, however, couldn't do their best without a unique Lockheed pointing control system (PCS) that positions and holds them precisely on target. And that is incorporated in the Lockheed Support Systems Module, containing the interfacing, communications, thermal control, data management, and electric power systems.

Using sensors and gyroscopic devices, the PCS will move and adjust the huge telescope into a programmed viewing attitude. The Perkin-Elmer fine-guidance sensors will fix upon a pair of known, bright 'guide stars,' references for finding the target.

When the target object is dead in the telescope's sight line, designated instruments will be turned on and other functions shut down. That's because the tiniest torque or motion in a small motor or electric relay could jar the 12-ton vehicle off its lock-on point.

And how precise is that lock-on? In angular measurement, there are 1,296,000 arc-seconds in a circle. For perhaps as long as 20 hours at a time, the telescope will hold onto a target with a stability accuracy of just *seven thousandths of one arc-second* in all directions.

That stability is 100 times greater than any ever achieved so far. It about equals standing in Boston and holding rifle sights steady on a dime suspended over Washington, D.C.!

In a neutral buoyancy tank, an astronaut goes through special training simulating the service tasks to be performed in space.

Keeping a beauty in shape.

Lockheed's innovative design will also let the telescope have periodic 'house calls' for checkups and maintenance during its planned 15-year life. Armed with nothing more than a 7/16″ ratchet wrench and a Lockheed-devised foot-anchoring tool, a Space Shuttle astronaut will be able to open access doors, make inspections, replace modular units, and see to the telescope's well-being. When necessary, the Shuttle will retrieve the telescope and return it temporarily to Earth for general servicing.

The world of science is waiting expectantly. The wonders yet to be discovered are waiting, too. The Space Telescope will arrange a meeting.

When it comes to bringing all this together—testing, refining, integrating all the critical elements of a miraculous scientific breakthrough—Lockheed knows how.

Lockheed

SCIENCE AND THE CITIZEN

Depressed Area

When the necessity of improving the quality of education in the U.S. is discussed, it is often observed that teachers are inadequately paid. Insofar as college teachers are concerned the situation is a lot worse than has been generally realized. Data assembled by the American Association of University Professors show that in terms of purchasing power the pay of college teachers declined 18.9 percent between 1967–68 and 1981–82, whereas in the race with inflation people in nonacademic jobs came out even.

The data are presented in "The Annual Report on the Economic Status of the Profession, 1982–83," written by W. Lee Hansen of the University of Wisconsin at Madison, who is chairman of the A.A.U.P. committee on the economic status of the profession. The report is published in *Academe,* the bulletin of the A.A.U.P.

The committee found that average academic salaries in 1981–82 were $33,480 for professors, $25,210 for associate professors, $20,630 for assistant professors and $16,310 for instructors. For full professors at major universities the range was from $29,400 at the University of South Dakota at Vermillion to $53,600 at Harvard University. For all ranks combined in all colleges and universities the median salary was $25,750.

Comparing the average academic salaries in 1981–82 with the compensation of similar groups of workers in nonacademic jobs, the committee found that the teachers fared poorly. "College professors," Hansen writes, "earn substantially less than self-employed physicians, dentists and related practitioners ($46,530); salaried physicians, dentists and related practitioners ($38,500), and even engineers ($31,070), who for the most part hold only baccalaureate degrees. Faculty members earn only slightly less than other self-employed professionals ($27,880)." Hansen notes that the average assistant professor receives a salary ($20,630) that is below the full-year earnings of the average male sales worker and only slightly above those of the average craftsman.

The committee chose 1967–68 as the base year for its study of real salaries because inflation then was minimal, Government controls and guidelines did not affect the determination of faculty salaries, the economy was reasonably prosperous and funding for higher education was not unduly constrained. Against that background the committee found that real faculty salaries rose 2.1 percent from 1967–68 to 1971–72, fell 10 percent during the next five years and fell another 11.7 percent in the five years up to 1981–82. For 1982–83, Hansen reports, real salaries rose 2 percent, mainly because of the slowing in the rate of inflation.

The prospects for improvement in the pay of teachers are not encouraging, according to Hansen. "The dramatic drop in the inflation rate and the virtual stability of consumer prices since the late summer of 1982 have weakened faculty arguments for higher nominal salary increases. At the same time pressures to hold down salary increases in the public sector have been exceedingly strong because of the economic difficulties faced by many states whose revenues have fallen sharply as a consequence of the protracted recession. Even without these pressures there is little evidence of any significant change in public attitudes that would support higher salary increases to make up for past declines in real incomes."

Insult on Insult

Evidence from epidemiological studies of the incidence of cancer and from the examination of tumor cells has long suggested that carcinogenesis is a multistep process. Now that impression has been confirmed at the molecular level. In a number of experiments the transformation of normal cells into cancer cells has been shown to depend on at least two distinct genetic changes, each of them presumably directed by a different gene. The evidence comes from experiments involving not only oncogenes, the cancer-initiating genes found in human tumor cells, but also some environmental carcinogens and the genes of two DNA viruses that can cause cancer in laboratory animals. It appears more and more likely that a limited number of transforming genes will turn out to be responsible for a wide range of cancers arising in various tissues in response to seemingly disparate triggering events.

Oncogenes are altered versions of ordinarily benign genes ("proto-oncogenes") in normal animal cells. In one way or another they are activated to play a role in the conversion of a normal cell into the cancerous state. Their role has been demonstrated in the laboratory by the ability of a single oncogene to transform a line of cultured mouse cells called *NIH* 3*T*3; the transformed cells form colonies displaying the varied characteristics of tumor cells and, when they are injected into laboratory animals, give rise to tumors. The ability of an "established" cell line such as 3*T*3 to grow indefinitely in culture suggests, however, that the cells have already been partially transformed: they have been immortalized, and they may have undergone other changes characteristic of cancer cells. The question arose: What would it take to transform more normal cells? Given the epidemiological and pathological evidence for multistep carcinogenesis and the finding that some human tumors harbor two different oncogenes activated in different ways, investigators in three laboratories undertook to transform nonimmortal cells, newly cultured from normal rat or hamster tissues, with varying combinations of cancer genes or of cancer genes and carcinogens. Their independent findings were reported in *Nature.*

In each case one of the oncogenes was a version of the well-studied *ras* family of genes isolated from human bladder carcinomas, which had been shown to transform 3*T*3 cells. At the Institute of Cancer Research in Buckinghamshire in England, Robert F. Newbold and Robert W. Overell found that *ras* alone would not transform connective-tissue cells from the skin of hamsters. They had shown previously that treating such cells with either of two chemical carcinogens or with X rays would give rise to colonies of immortal but nonmalignant cells. When colonies of these immortal cells were exposed to *ras,* the cells were fully transformed. Newbold and Overell suggest that the activation of a *ras* gene may be a late step in carcinogenesis, that a preceding step is required and that "carcinogen-induced immortality [may be] a strong candidate for this critical early event."

H. Earl Ruley of the Cold Spring Harbor Laboratory found a different way to complement the activity of a *ras* oncogene. He worked with genes isolated from the polyoma virus and the adenovirus, two DNA viruses that infect human beings but cause cancer only in certain laboratory animals. In the case of both viruses it had been shown that the transforming activity depends on two different regions of the viral DNA. The *E*1*A* region of the adenovirus DNA and the polyoma gene coding for what is called the large-*T* antigen seem merely to immortalize cells; the *E*1*B* region of the adenovirus and the polyoma-virus gene for the middle-*T* antigen seem to account for other changes, including the transformed cells' ability in turn to give rise to tumors in animals. Ruley showed that neither the *E*1*B* region nor the middle-*T* gene alone would transform rat kidney cells. Nor would a *ras* gene. When the adenovirus *E*1*A* region was introduced into cells along with either the *E*1*B* region, the polyoma middle-*T* gene or *ras,* however, the rat cells were transformed.

Much of the earlier work on *ras* had

been done in Robert A. Weinberg's laboratory at the Massachusetts Institute of Technology (see "A Molecular Basis of Cancer," by Robert A. Weinberg; SCIENTIFIC AMERICAN, November, 1983). Hartmut Land, Luis F. Parada and Weinberg proceeded to test the ability of their *ras* oncogene to transform fibroblasts cultured from rat embryos. The oncogene alone did not convert the rat cells into tumor cells, although it did increase their ability to proliferate. The M.I.T. workers then introduced *ras* into the rat cells along with *myc,* another oncogene. The two genes together gave rise to numerous rapidly growing colonies of transformed cells, which in turn gave rise to tumors in animals. Weinberg's group went on to test *ras* and *myc* in combination with the two polyoma genes. They found that the middle-*T* gene complemented *myc,* whereas the large-*T* gene complemented *ras.*

Taken together, the three reports suggest that at least some oncogenes and genes of DNA cancer viruses can be classified into two complementation groups, as the M.I.T. workers put it. Some genes (*myc* and the large-*T* gene, for example) seem to mimic the effects of establishment; other genes (*ras* and the middle-*T* gene, for example) seem to confer on immortalized cells other traits necessary for transformation. The actual situation is probably more complicated. There is some evidence that more than two steps may be required; the clinical development of cancer may be somewhat different from the transformation of cells in a laboratory culture. Yet it begins to appear that cancer results from a succession of different insults to cells, each of them activating a specific gene. Those genes may well be identified, and their mode of action understood, before long.

Restless Quarks

Quarks emerged from theory and acquired reality. Specifically, they emerged from the attempt to derive families of subatomic particles from combinations of more fundamental entities and then acquired reality when the electrons fired at protons at the Stanford Linear Accelerator Center (SLAC) revealed pointlike objects inside the proton. The theories proposed since that time have had the task of placing limits on the quarks' freedom. The problem is that the quarks inside the proton seem to be bound to one another by only a small force, yet they cannot be driven out of the proton. Hence the idea arose that quarks are confined in the proton by a force that gets stronger with distance, so that quarks are unable to draw apart from one another to more than about one fermi (10^{-13} centimeter), the diameter of the proton. The proton thus is taken to be an inescapable "bag." The force is called the color force, and each quark is said to carry a color charge. The theory of color interactions (developed in the past decade) is called quantum chromodynamics.

Some newly reported data now challenge the limits on quark confinement: they imply that the mere agglomeration of protons into a heavy atomic nucleus (that of iron 56) gives the quarks composing those protons a surprising amount of mobility. The data come from experiments done in 1980 at the CERN synchrotron in Geneva by the European Muon Collaboration, or EMC, a consortium of physicists. In each experiment iron nuclei were bombarded by muons: subatomic particles with the negative charge of the electron but 207 times the mass. Like electrons, muons interact with nuclei only electromagnetically; hence the pattern with which they are "scattered" by the encounter is relatively uncomplicated. The experimenters recorded the energy and the angle of each scattered muon. From that they could deduce the distribution of momentum among the pointlike objects—presumably quarks—constituting the internal structure of the protons the muons had probed.

The seminal SLAC experiments (in which electrons were scattered off isolated protons) had already demonstrated that only about 35 percent of the momentum in a proton is carried by valence quarks, that is, the three quarks (two "up" quarks and one "down" quark) of which the proton is said to consist. Another 50 percent is carried by gluons, the particles whose passage from one quark to another is said to bind quarks together. (Gluons are the "exchange particles" that transmit the color force.) The remaining 15 percent is carried by "ocean" quarks: a sea of quark-antiquark pairs whose existence is required by quantum field theory, the marriage of quantum mechanics and relativity in which the dynamics of quarks are described. In effect part of the internal energy of the proton is stored in quantum fields that manifest themselves when they are probed, say by a muon, as a multitude of quarks and gluons.

The quarks in iron nuclei turn out to differ in two ways. For one thing the valence quarks carry less momentum than they would carry in an isolated proton. In addition ocean quarks are present in greater numbers: they are observed to be 60 percent more numerous. The binding of protons and neutrons into a nucleus involves energies on the order of 10 million electron volts; the internal structure of a nucleon (a proton or a neutron) is thought to be sustained by energies at least 10 times greater. Yet the binding somehow brings on a reshuffling of momentum in which new ocean quarks arise.

An attempt to account for these findings is offered by Robert L. Jaffe of the Massachusetts Institute of Technology. He proposes that the valence quarks in iron nuclei are "partially deconfined" because the three-quark bags in a heavy nucleus are large compared with the spacing between nucleons. The bags thus overlap, and valence quarks can "percolate" from one bag to another. The uncertainty principle, a basic tenet of quantum mechanics, then comes into play. It dictates that an increase in the range of positions available to a percolating quark must be accompanied by a decrease in its momentum.

Why do new ocean quarks appear? An extension of Jaffe's ideas offered by Frank Close, Richard G. Roberts and G. Graham Ross at the Rutherford Laboratory in England suggests an explanation. Close and his colleagues consider two aspects of the distance scale that characterizes the interaction between a muon and a quark inside a proton. First there is the scale at which the muon probes its target. The theory of quantum chromodynamics affirms that the color charge of a quark gives rise to a surrounding quantum field that amounts to a cloud of gluons. The abundance of these gluons decreases with increasing distance from the trajectory of the quark. In much the same way the electric charge of an electron traveling at an appreciable fraction of the speed of light gives rise to an electromagnetic field that amounts to a cloud of photons. A muon probe may send a quark off on an altered path. Among the surrounding gluons, however, the ones at least a certain minimum distance from the quark will continue their flight without change. Increasing the energy of the muon means it will shave the quark more closely and hence will free more of the gluons. The gluons can then manifest themselves as ocean-quark pairs. In much the same way the photons freed from the cloud surrounding a high-speed electron give rise to pairs of electrons and positrons.

The second aspect of the scale of the interaction between a muon and a quark involves the distances over which the quark can range through nuclear matter. In quantum field theory the flight of a quark is described by a mathematical construction (an operator) representing the removal of a quark from a certain position in a nucleus and an operator representing the replacement of that quark at a second position, with the nucleus unchanged. The multiplication of the two yields an "expectation value" representing the "mean free path" of the quark. The expectation value is a measure of the freedom the particle has.

What Close and his colleagues note is that in quantum chromodynamics a scale (that is, a parameter representing length) enters only once. Thus a change in the scale at which a muon probes a quark must act in the same way that a change in the scale of the freedom of the quark does. If the former can alter the number of ocean quarks detected, the latter must alter that number too. Close and his colleagues make a specific calculation: they show that if quarks have from 10 to 20 percent more freedom in a heavy nucleus than they have in an isolated proton, the "EMC effect" will arise in both of its aspects. Valence quarks will show a decrease in momentum; ocean quarks will show an increase in number.

The prospect that the quarks in heavy nuclei are partially deconfined has even encouraged some theorists to suggest that muons of sufficient energy might increase the scale of the freedom of quarks until it equaled the size of the nucleus overall. Quarks would range throughout the nucleus, which would then conduct color charge much as a metal conducts electricity. A less extreme position is a simple satisfaction at the apparent verification of a prediction of quantum chromodynamics. It is hypothesized that nuclear matter in the very early universe was a gas of unconfined quarks and gluons that coalesced into protons and neutrons. Any attempt to produce such a gas today would doubtless require that nuclei be subjected to extreme values of temperature or density or both. Those favoring the attempt may be heartened by the evidence that even normal nuclear densities yield a partial deconfinement.

Zip plus 4

Collecting, sorting and delivering the roughly 118 billion pieces of mail that Americans send annually is the job of some 650,000 postal workers. For several years the Postal Service has been working on a plan to cut its labor costs by automating the sorting process and by introducing a nine-digit ZIP code with which the new machines could sort mail down to city blocks and even to individual buildings. The "ZIP + 4" program began in earnest last October, when the service's Board of Governors approved rate incentives designed to encourage businesses to put the nine-digit codes on their large-volume, first-class mailings and to make their envelopes machine-readable.

Post offices began receiving two types of new equipment late in 1982. "Optical character readers" (OCR's) read the city, state and ZIP code off an envelope, check to make sure the city matches the code, translate the ZIP code into a bar code that is then stamped on the envelope, and finally channel the letter into one of 60 stacks. Less expensive "bar-code sorters" (BCS's) subsequently sort five-digit mail to the destination post office and nine-digit mail directly to the carrier. Observant people may already have noticed bar codes (sequences of long and short vertical lines) at the bottom of envelopes they receive from some businesses.

Reading machines are by now a well-established technology; machines that can read most types of printed material aloud to blind people are in service in many libraries and universities. Post offices in Japan and Europe began deploying OCR's in the 1970's. The basic technology is the same for all such machines: an electronic camera scans the printed material, bouncing a continuous beam of light off it and measuring the difference in reflectivity between print and background. The scanner then transmits the image as a digital signal to a computer that is programmed to recognize individual letters or numbers based on characteristic features that are relatively invariant across type fonts.

Although it may at first sound less complicated than reading a book, sorting mail presents special problems to a machine because of the diversity of the material. An OCR can read a wide variety of type fonts, but it will balk at handwriting, script fonts and some laser

printing. Skewed lines, poor contrast between print and background, and extraneous printing can all confuse the machine. Finally, an OCR must also deal with the fact that people often misspell or abbreviate words in an address. The OCR manufactured for the Postal Service by Pitney Bowes, for example, contains more than 20 possible spellings for "Philadelphia" in its memory, and it is even programmed to know that "Big Apple" means New York City. OCR machines are currently successfully reading from 60 to 70 percent of the high-volume business mail they are being fed, but the Postal Service reckons it will have to improve that rate by about 10 percent, chiefly by teaching businesses how to avoid thwarting the machines, if the automation program is to achieve its desired savings.

OCR's and BCS's eliminate the need for most of the clerks who sort mail; with automatically sorted ZIP + 4 mail the first person to read an envelope is the carrier. The new machines process about 10,000 pieces of mail per man-hour, compared with 1,800 pieces for nonautomatic sorting machines that require clerks to read the envelopes and about 800 pieces per man-hour for sorting by hand.

The Postal Service plans to buy about 650 OCR's and 700 BCS's by 1988, at an estimated cost of $740 million for the equipment alone, in order to automate letter sorting at the 211 major post offices that handle approximately 90 percent of all "machinable mail." When the acquisition program is complete, the service believes that it will require 15,600 fewer man-years of labor, and that it will save $600 million per year as a result. About $150 million of this would be returned to business mailers through rate incentives, and the remaining $450 million would help to stabilize postal rates.

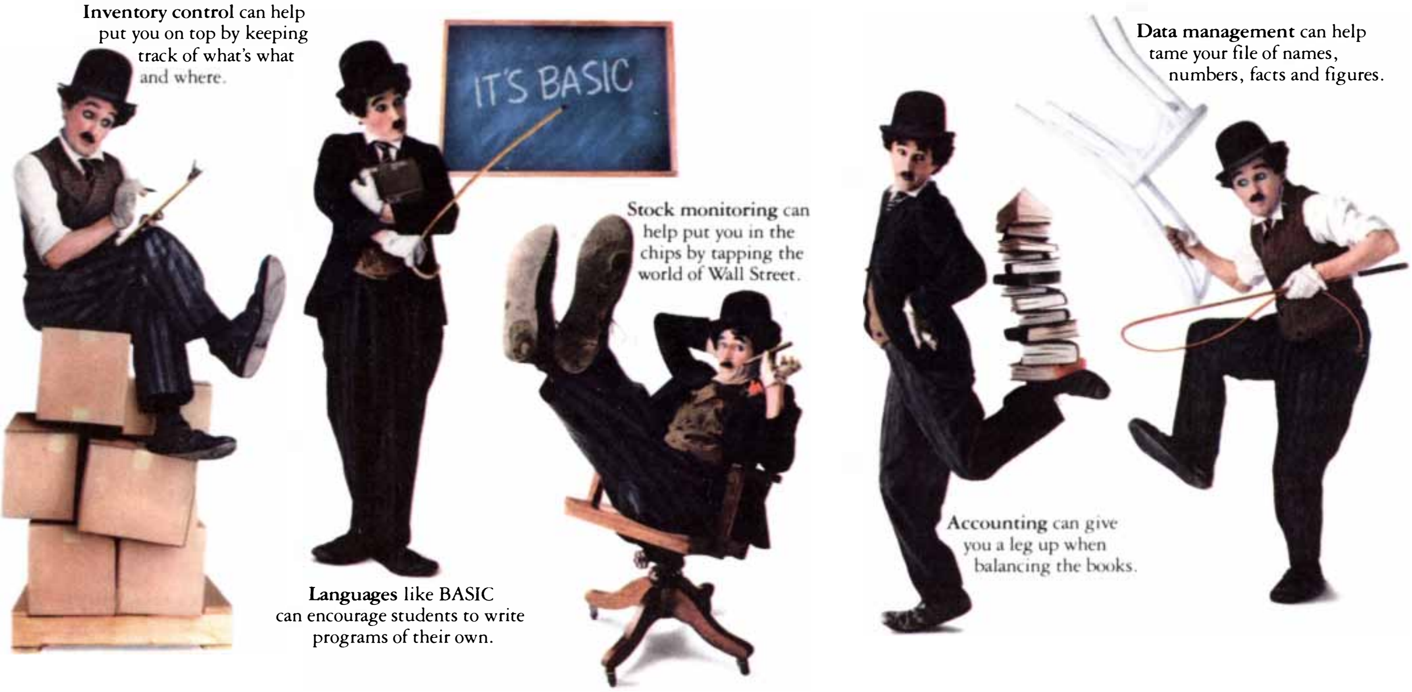
Whether these savings are actually achieved depends on whether the rate incentives (.5 cent per piece for presorted batches of 500 or more, which already receive a three-cent discount, and .9 cent off the normal 20-cent rate for nonpresorted batches of 250 or more) are enough to induce businesses to convert to ZIP + 4. Participation by businesses is crucial, because they send 83 percent of all first-class mail. Since personal correspondence accounts for only 6 percent of the total (the rest is business reply mail), most of it not machine-readable, the Postal Service does not particularly care whether individuals use the 32 million nine-digit codes. In fact, it has no plans for telling households what their codes are. Nor do individuals have an incentive to put ZIP + 4 on their letters: the extra four digits will make mail sorting more efficient, but they will not, according to the Postal Service, speed up delivery.

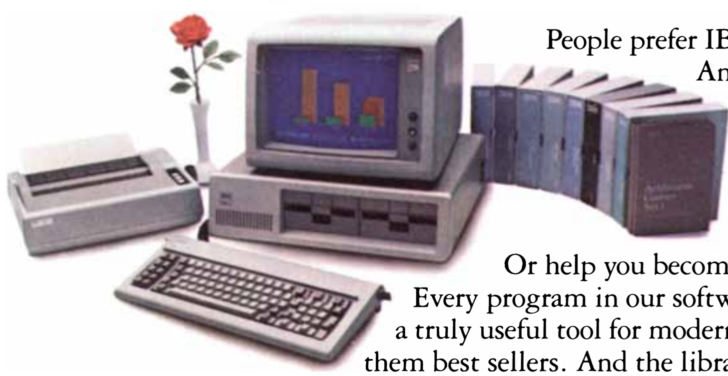
Cull of the Wild

The phenomenal growth of agricultural productivity over the past few decades can be attributed in large part to the application of modern concepts of genetics to the breeding of superior crops and livestock. So far most of the gains have come from the exploitation of the genetic diversity present in species that have long been domesticated. In recent years, however, it has become apparent that further improvements in the yield and quality of many agricultural products may depend increasingly on the introduction of genes from wild plants and animals. Accordingly greater attention is being paid to the conservation of wild genetic resources, both in natural habitats and in artificial ones.

Wheat is a case in point. Between 1930 and 1975 wheat yields per hectare in the U.S. rose by 115 percent, with roughly half of the increase being due to genetic advances. Nevertheless, it has been argued that the genetic material of the cultivated wheats has already been exploited for breeding purposes almost to its full capacity. Moreover, the very

What's on the best-seller list in IBM Personal Computer software?





People prefer IBM Personal Computer software for a variety of reasons. And one of the best reasons is its variety.

Because, for just about anything you want the IBM Personal Computer to help you do, there's software to help you do it. Software to help improve productivity, efficiency and planning. To help teachers teach and students learn.

Or help you become an even more astute gamesplayer.

Every program in our software library makes the IBM Personal Computer a truly useful tool for modern times. That's why a lot of buyers like you have made them best sellers. And the library is still growing.

So the best may well be yet to come. IBM®

For more information on where to buy IBM Personal Computer software and hardware, call 800-447-4700. In Alaska or Hawaii, 800-447-0890.

success of modern plant-breeding practices has had the unintended effect of drastically reducing the range of genetic variation of the cultivated wheats.

In general, crops have already been improved significantly by the contribution of wild genetic resources, whereas livestock have not. The most important benefit obtained from the wild so far has been disease resistance. In the case of wheat, for example, several commercial cultivars of the common bread wheat *Tritium aestivum vulgare* have benefited from the transfer of genes for resistance to certain strains of fungal disease from distantly related wild species. Similarly, in the case of rice much of the Asian crop now gets its resistance to two major diseases (blast and grassy stunt virus) from the wild rice species *Oryza nivara* (see "Rice," by M. S. Swaminathan, page 80). The need for new resistant cultivars is a continuing one, since most pathogens evolve fairly rapidly.

According to a recent review of the worldwide status of the exploitation of wild genetic resources, published by the International Institute for Environment and Development, the two cereals cited above are among 24 major crops that have already been improved by the introduction of germ plasm from wild species. The other beneficiaries are maize, barley, oats, potatoes, cassava, sweet potatoes, sunflowers, oil palms, sesame plants, tomatoes, peas, carrots, grapes, apples, pears, strawberries, sugarcane, sugar beets, tobacco, rubber, cacao and cotton. The report points out that some species, such as the tomato (*Lycopersicon esculentum*), probably could not be grown as commercial crops without the genetic support of their wild relatives.

In the view of the authors of the report, Robert and Christine Prescott-Allen, the results achieved so far in this direction "are substantial enough to demonstrate the enormous potential of the genetic diversity of wild plants and animals to improve the yields and quality of domesticated crops and livestock and to provide for the more rapid domestication of new crops and livestock." The exploitation of wild genetic resources, they point out, "is a means of benefiting materially from wildlife without harming the donor species." As with wild species, the main threats to wild genetic resources are identified as "loss of habitat, overexploitation, and competition and predation by introduced species." With wild genetic resources, however, "the impacts of these threats are more difficult to detect. Valuable gene pools of widespread species may disappear undetected, either because not enough is known about the distribution of genetic variation within the species, or because the very abundance of the species masks the disappearance of its constituent gene pools."

The only prudent policy, the Prescott-Allens conclude, is a two-track approach, combining in situ conservation (defined as the maintenance of a wild gene pool in its native habitat) with ex situ conservation (the maintenance of the resource outside its native habitat, for example in zoos, botanical gardens and germ-plasm collections). "Fortunately," the authors report, "a convergence of opinion appears to be emerging among genetic resource users and the wildlife conservationists and protected-areas community.... New ways of maintaining and new ways of using the newest resource are being developed, as they must be, side by side."

Da-da-da and ba-ba-ba

In the four to six months before the onset of recognizable speech almost every child begins to babble. The preceding gurgles and unarticulated glides give way to relatively stable vocalic sounds punctuated by the repetitive stopping and release of air in the vocal tract; the rhythmic and playful sounds "da da da" or "ba ba ba" that result are familiar in the nursery. Now Rachel E. Stark and Jennifer Bond of the John F. Kennedy Institute in Baltimore have closely studied babbling in three infants; they conclude that the behavior represents an abrupt advance in linguistic development and may be neurophysiologically related to other developments such as rhythmic waving and shaking that begin at roughly the same age. They presented their results recently at a meeting of the Acoustical Society of America held in San Diego, Calif.

Although the distinction between babbling and its precursor sounds has been recognized for many years, there is renewed interest in the phenomenon among linguists because of the limited sound repertory that is observed. Until recently it was widely believed babbling infants explore the entire range of speech sounds employed in the languages of the world. It is now known, however, that many speech sounds are almost never heard in babbling, and certain others have a high frequency no matter what the language community into which the child is born. Work on babbling has been carried out in at least seven languages, among the children of speakers of Arabic, Chinese, English, Japanese, Norwegian, Spanish and Yucatec. In all these languages the most frequently babbled sounds are nasals such as *n* and *m*, glides such as *y* or *w* and single stop consonants such as *d*, *t* or *b*, followed by a vowel formed with the tongue near its "home" position, such as the eh sound in "bet" or the aeh sound in "bat." Vowels formed with the tongue in an extreme position, such as the aw sound in "ought" or the front, rounded vowel sound in the French word "tu," are seldom heard, and consonants such as *l*, *r*, *f* and *v* and consonant clusters such as *st* are rare. Moreover, according to D. Kimbrough Oller of the University of Miami, the sounds favored during babbling are those favored by the child a few months later when meaningful speaking begins.

In the study by Stark and Bond both audio and video recordings of the three infants were made in each infant's home. The audio recordings were analyzed for various acoustic characteristics, such as the pitch contour, the frequency spectrum of the signal, the timing of the repeated utterance and the timing of the consonant release. The video recordings made it possible to transcribe the sounds emitted both before and after the onset of babbling according to place and manner of articulation. Stark and Bond found not only that babbling is rather stereotyped behavior but also that it seems to mark a sharp discontinuity in linguistic development. Utterances emitted before the onset of babbling were broken up into fairly large chunks by silent intervals or by a breath. Although there were changes within the chunks, none of them gave rise to a clear separation of the chunks into syllables. With the sudden onset of babbling by the time the child was six to eight months old, the chunks were clearly divided into small segments or syllables; the syllables were identified by rapid changes in the intensity of the acoustic signal and in the pitch contour. Two to four syllables, each lasting for about half a second, were found in each chunk, and the chunks were grouped together in pairs or in triplets.

It is not yet clear from such studies exactly what role babbling plays in the acquisition of language. Oller maintains that babbling manifests an emerging capacity for controlling the timing and articulation of speech sounds, although he points out that babbling may not be a necessary stage of language development. Children who are unable to babble because of surgical tracheotomy can learn to speak after the trachea is reclosed without necessarily babbling. Other investigators suggest that children who babble playfully with adults are learning to take turns, which is a prerequisite of conversation. Stark and Bond point out that at roughly the same age as the onset of babbling the child has also become preoccupied with objects, and object interactions often take the form of repetitive behavior such as waving or banging. All such repetitive behavior may reflect development of the motor cortex of the brain. Because there is less variation among children in babbling than there is in other kinds of behavior, some linguists believe it should be possible to detect neurological impairment early by finding irregularities in babbling.

AUSTRALIAN TECHNOLOGY: PAST, PRESENT, FUTURE

A free ride on Qantas and TWA? Nothing's free.

The Frequent Flight Bonus[SM] Program.

Latest in a long line of cheap stunts to lure business travelers to Australia. And louse up my peace and quiet.

Qantas has joined with TWA (an airline that should know better) to offer free discounts. And free 2-for-1 tickets. Trips to places like New York, London, Paris, Rome, Athens, Tahiti and (unfortunately) Australia and The Great Barrier Reef.

Even free round-the-world First Class tickets for two!

What the fine print doesn't tell you is that they won't pay a cent toward your passport. Or that suit you want tailored on Saville Row. Or the opals you can't resist in Australia. Or that side trip to the Greek Islands.

So before you jump willy-nilly into this Frequent Flight Bonus Program, pick up an application at any Qantas or TWA office and read it. Every word. Every detail.

You'll probably decide–as I have–to give both airlines the big thumbs down.

Koalas Against Qantas and TWA.

The Australian Airline.

Qantas and TWA Frequent Flight Bonus Program. You earn credit for every mile you fly anywhere on the worldwide route networks of Qantas and TWA. This program available only to U.S. residents.

Australian Technology: Past, Present, Future

Researched and written by Jane Ford, with John Costello
Edited by Michael Knipe

Contents

Australia is fast becoming aware that vast mineral and agricultural wealth is no longer enough to guarantee its future. The importance of scientific and technological expansion has become an issue of national concern. The following study presents a wide-ranging review of technology and its potential in Australia. The report contains two primary parts. A general survey, divided into past, present and future sections, provides historical background and organizational information on the nation's scientific enterprise. A series of focus articles examine specific companies, projects, regions, technologies and issues.

The Survey

Focus Articles

Conceived and produced by:	John M. F. Keeney
Marketing and coordination by:	Stephen P. Smith, Managing Director, International Media Representatives (IMR), 2nd Floor, Cordell Building, 160 Sailors Bay Road, Northbridge NSW 2063, Australia TEL: (2) 958 1699, TELEX: 74473 IMR
Public Affairs Consultants:	Canberra: Jonathan Gaul and Peter Sekuless, Canberra Liaison Pty. Ltd. Perth: David Deane Spread, Deane Spread Associates Brisbane: Robert Baudino and Peter McCall, B & M Public Affairs Consultants (Aust.) Pty. Ltd.
Front Cover:	Vernon Oakley Design, London

The Past: An agricultural and mining economy is founded and early innovation flourishes

Australia's development has resulted from the application of capital and labor to the natural resources of the continent. However, supplies of capital and labor have not always been sufficient, and development has been periodically thwarted by the severity of the climate. In exploration and development a prolonged battle has been waged against severe droughts and an extreme range of temperature.

Innovation has been an essential element in helping Australia to overcome such obstacles since its earliest pioneer days.

The economic impulse that opened up the continent originated in the wool industry, which for more than a century has supplied Australia's most valuable export. The first settlers, however, were interested in sheep as food rather than in their wool. When John Macarthur, both an officer in the New South Wales Corps and a farmer at Parramatta, suggested breeding sheep not for their meat but for the fleece, he met with derision from his fellow colonists. But when in 1797 Macarthur imported a small flock of Merinos, the more discerning farmers realized the potential value of the wool.

Governor King, an advocate of the development of the wool trade, wrote prophetically in 1800: "The introduction of a breed of Spanish sheep into the flocks of individuals has so much improved the fleeces that there is a promising appearance of a great quantity of wool being produced in a few years." In 1807 the first consignment of Merino wool, weighing 245 pounds, was shipped to England, where its fine quality attracted wide interest.

The Australian multimillion-dollar wheat export industry had its origin when William Farrer, Australia's pioneer of scientific wheat breeding, realized that English wheat varieties were potentially high yielding in Australia but matured too late to survive the long summer droughts. In 1886 he began breeding new varieties. By crossing Indian and English varieties he produced the breed called Federation, which combined early maturity with high yields. Because the wheat is grown under drier conditions, it is less susceptible to attack by rust.

Innovative farm machinery also vitally spurred Australia's development. In 1843 John Ridley, an Adelaide miller, invented the stripper: a combination grain-gatherer and thresher. The machine effectively combined two out of the three steps in the harvesting of grain; winnowing was still done separately with a hand-turned machine.

Thirty years later R. B. Smith's stump-jump plow, which became Australia's most well-known agricultural invention, further simplified farming. In 1884 Hugh Victor McKay, then 18 years old and working on his father's farm in Drummartin, Victoria, revolutionized the grain industry by constructing the first commercially successful combine harvester: it stripped, thrashed, winnowed and bagged the wheat in one continuous movement.

Captain Cook's arrival in Botany Bay aboard the Endeavour in 1770 marked the beginning of a strong scientific tradition in Australia. Cook had set sail at the behest of the Royal Society of London to plot the transit of Venus from Tahiti and to discover a new southern continent.

He did not discover a new continent but instead managed to explore the entire eastern coast of Australia, claiming the region for Britain. He amassed a superb collection of bizarre and primitive plant and animal specimens, which aroused enormous interest on his return to England.

Britain established the first penal colony at Port Jackson in 1788—now the site of Australia's largest city, Sydney. For 60 years or so science and technology languished under convict settlement conditions but in the 1840's penal transportation came to an end and innovative, educated settlers poured into the colony, fleeing Victorian England determined to make a new life.

The new settlers erected universities, museums, herbaria and astronomical and magnetic observatories, beginning an exciting period of innovation. They soon realized that the vast distances and isolation of Australia necessitated improved communication and transport systems. A 10,000-mile network of railway tracks was built across the continent and a transcontinental telegraph opened between Adelaide and Darwin in 1872, only 10 years after the continent was crossed for the first time from south to north. The telegraph was soon connected to the overseas cable, linking Australia to London and the rest of the Western financial and commercial world.

To farm the poor soils and drought-prone land new breeds of crops and animals were needed, and the great mineral wealth could only be tapped through new exploration and mining methods. Such were soon developed.

AUSTRALIA—FACTS AND FIGURES

Figures for 1981–82 unless indicated

AREA: 7,682,300 sq. km.

POPULATION: 15.2 million (70% urban)

STATES AND TERRITORIES	Area ('000 sq. km.)	Population ('000)	Capital	Population ('000)
New South Wales	802	5,305	Sydney	3,281
Victoria	228	3,994	Melbourne	2,804
Queensland	1,727	2,421	Brisbane	1,086
South Australia	984	1,331	Adelaide	952
Western Australia	2,526	1,334	Perth	918
Tasmania	68	430	Hobart	170
Northern Territory	1,346	127	Darwin	57
Australian Capital Territory	2	230	Canberra	246

CONSTITUTION AND GOVERNMENT

Federal System

Seats (1983)	House of Representatives	Senate
Labor	75	30
Liberal	32	24
National Party	17	4
Australian Democrat	—	5
Independent	—	1
Vacant	1	—

NATURAL RESOURCES

Mining ('000 metric tons)	
Iron Ore	86,236
Black Coal	100,428
Bauxite	24,987
Mineral Sands	1,901
Metallic Contents of Minerals	
Copper	249
Lead	654
Zinc	591
Nickel (1979)	72
Agriculture ('000 metric tons unless indicated)	
Wool (million kg.)	716
Wheat	16,330
Meat (excl. canned)	2,598
Sugar	3,434

NATIONAL ACCOUNTS

GDP	$ 148,235
Gross Fixed Capital Formation (1982: Percentage of GDP)	25.1

TRADE

Principal Exports (% of total value)
Coal 12, Wool 10, Wheat 9, Metals 8, Meat 7, Iron Ore 6.

Destination (% of total value)
Japan 28, U.S. 11, EEC (excl. U.K.) 8, ASEAN 8, New Zealand 5, U.K. 3, China 2.

Source: Australian Bureau of Statistics
Bureau of Agricultural Economics
Bureau of Mineral Resources
OECD

At the time of writing 1 Australian dollar equals .91 U.S. dollars.

All dollars are Australian dollars unless otherwise indicated.

AUSTRALIA

Australian Information Service

A string of meteorological stations set up across the continent and meteorological services established in each colony helped to keep track of the weather conditions.

The discovery of gold in 1851 provided an unexpected stimulus to Australia's development. Between 1840 and 1850 the population increased by 200,000 and in the following 10 years to 740,000, bringing the total to more than one million. The number of gold miners in the country peaked at 150,000 in 1858 and by 1865 had declined to 80,000. Company mining gradually replaced individual activity as equipment and processes became more complex.

Development of agriculture was slower than in America where the climate was more temperate, the land more suitable for farming and transport far better. Australia has few navigable rivers, and it was not until railways reached the wheat lands across the Great Dividing Range that the farmer could come into his own. Wool could be more cheaply transported and between 1861 and 1884 the area under crop in New South Wales increased from 300,000 to 700,000 acres, which were owned by a mere 40,000 farmers. Pastoralists improved their breeds by increasing the weight of the fleeces, began using better methods for washing, pressing and sorting the wool and marketing it more efficiently.

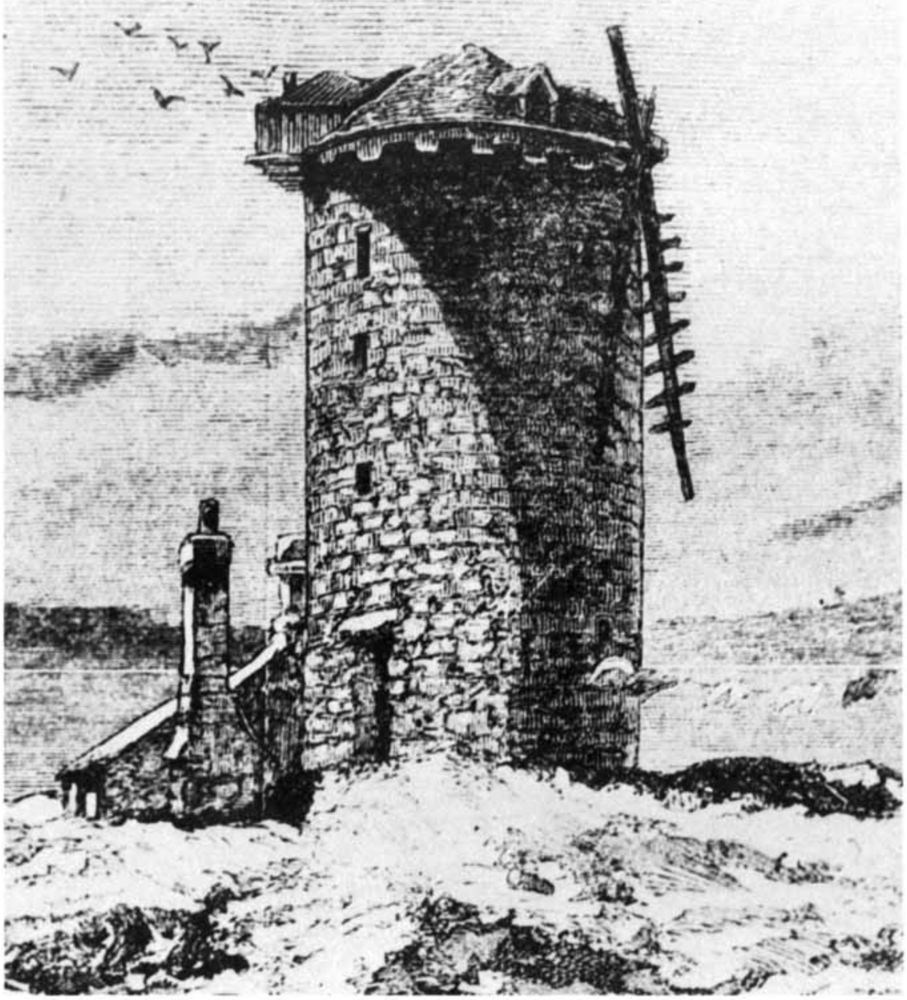
Farming made earlier progress in Victoria, where the climate is milder and transportation was more efficient. The area under crop increased from 400,000 acres in 1860 to 1.5 million acres in 1880. In South Australia the wheat-growers succeeded even better. Between 1862 and 1880 their crop constituted more than half the total wheat production of the continent; in 1881, more than eight acres per capita were planted with wheat.

Gradually, the need for a scientific study of agriculture was realized. The Roseworthy Agricultural College in South Australia was founded in 1884, followed in 1885 by the Dookie College in Victoria.

Pioneering Investment

Considerable capital investment in industry, mining, communications and other public works brought about additional development of Australia. In the 1860's Victoria protected its infant manufacturing industries with a tariff; under its shelter textiles, clothing, leather goods, foods, agricultural implements and foundry products were developed or manufactured.

Australia's great distance from Europe and America also provided considerable market protection and made possible the development of industries connected with rural production. Many of the major financial houses, insurance and pastoral companies and merchandising concerns that still dominate the Australian commercial scene were developed in this period. Mining



Australia's first windmill, built at The Domain, Sydney

was another important company activity. The development of the railways and the steamship increased the demand for coal, bringing prosperity to the collieries around Newcastle.

The greatest investors in the country were not Australian private businesses but various colonial governments. Australian companies began to build the railways, with high hopes in Victoria and New South Wales, but failed to complete the lines, and so the colonial governments took over construction.

The early difficulties were partly financial and partly technical. In the eastern colonies, for example, the railroad lines had first to cross the rugged mountain barrier, which obstructed the engineers and prevented local traffic, increasing costs and reducing revenue. But in the 1870's a period of optimism began, and so the railways rapidly advanced. In 1861 there were fewer than 250 miles of lines, in 1870 only 1,000 miles, but by 1890 approximately 10,000 miles were open at a cost of more than £100 million.

Telegraph lines began to cross the continent; by 1889, 40,000 miles were erected. In the cities tramways and waterworks were built, and in the 1880's the great irrigation works in Victoria were begun. Between 1860 and 1880 Australia's population doubled, rising from 1.1 million to 2.2 million. It rose by another million over the next 10 years, thanks to rapid natural increase and vigorous immigration programs.

The country prospered. But all the newcomers had to be housed, fed and clothed, and the governments extended transport facilities and public utilities to provide for them.

By the late 1880's, however, there were signs that the boom was coming to an end. Metal, wheat and wool prices continued to fall. The pastoralists were already in difficulty and the great strikes of 1891, led by the newly powerful labor movement, brought an unwelcome notoriety to Australia's problems. The confidence of British investors waned, and for the first time in many years the colonial governments failed to raise the money they needed on the London market. The banks had to restrict credit, and the governments had to economize. Public work gradually came to a standstill, causing unemployment to increase. Neither city speculators nor many pastoralists could repay their overdrafts, yet many depositors in land banks, building societies and even trading banks wanted their money in gold, for paper money was not yet in use as legal tender.

The building societies and then the land banks started to go bankrupt, and a run on the trading banks occurred. In April, 1892, the Commercial Bank of Australia suspended payment, and by the end of June of that year, 22 other banks had either suspended payment or failed; only 10 banks survived.

Recovery began with the expansion of wheat growing. The extension of the railways, low wool prices and in particular improved farming techniques led to the doubling of the area under wheat cultivation in New South Wales between 1894 and 1899 and a second doubling by 1911, when over three million acres were planted, producing a crop of 38 million bushels. This is roughly 10 times the crop of 1891, greater than that of any other state in Australia at the time and, as a consequence of the discoveries of William Farrer, with an increased yield per acre. Australia soon assumed a world role as a major wheat exporter.

In the 1890's development of Western Australia boomed. The discovery of major gold deposits at Coolgardie and Kalgoorlie resulted in the colony's becoming the country's major gold producer. The population rose from 40,000 to 180,000 in less than a decade, and as the diggers turned to the land in a short time more than one million acres were under cultivation.

Industrial development suffered a setback as a result of the financial crisis of the 1890's, but progress resumed when Australia gained self-government in 1901. The six colonies of Victoria, New South Wales, South Australia, Western Australia, Tasmania and Queensland were made into states under the new federal government of Australia. Import duties were imposed with the avowed intention of protecting Australian manufacturers from the competition of countries where the workers' wages and the conditions of work were inferior. One result of this was the near-doubling of the number of factory workers in Victoria and New South Wales between 1901 and 1914.

Nevertheless, much of Australian industry remained tied to primary production and to consumers' immediate needs: clothing, food and drink, agricultural machinery, woodwork, vehicles, books and printing. One notable development at this time was the construction of an efficient modern iron- and steelworks at Newcastle.

After World War I

World War I intensified industrial development. In 1916 the first major industrial enterprise was established in Tasmania: the Electrolytic Zinc Company, which used hydroelectric power supplied from a new government power project.

After World War I profits in agriculture dropped markedly due to falling world prices and increased costs. Industries that had flourished immune from foreign competition demanded protection, and during the 1920's tariffs were steadily raised. Not all manufacturers were efficient, however, and most were handicapped by having to produce on a small scale. As a result the economy shifted to increased tertiary or service income earned from transport and distribution, construction, professions such as law and medicine, entertainment and other consumer services, which by 1928 comprised 56 percent of the national income.

The Depression

The world depression of 1929 brought near-chaos to the Australian economy. The government had been raising almost £30 million a year from overseas; the sudden cutoff of funds and the decline in export prices reduced the real national income by more than 10 percent and threw one-third of the labor force out of work.

With the recovery of world prices in the 1930's manufacturing expanded considerably. Despite a modest population increase between 1930 and 1940, the real value of secondary production was one-third above its predepression level.

World War II, like its predecessor, stimulated the Australian economy; the greatest advances took place in the metals and engineering industries. Tremendous postwar immigration brought more than half a million people to Australia between 1948 and 1953, seriously straining public services and housing. However, as world prosperity increased and the prices of food and primary products went up, Australian farmers experienced a period of extraordinary prosperity.

Early Science

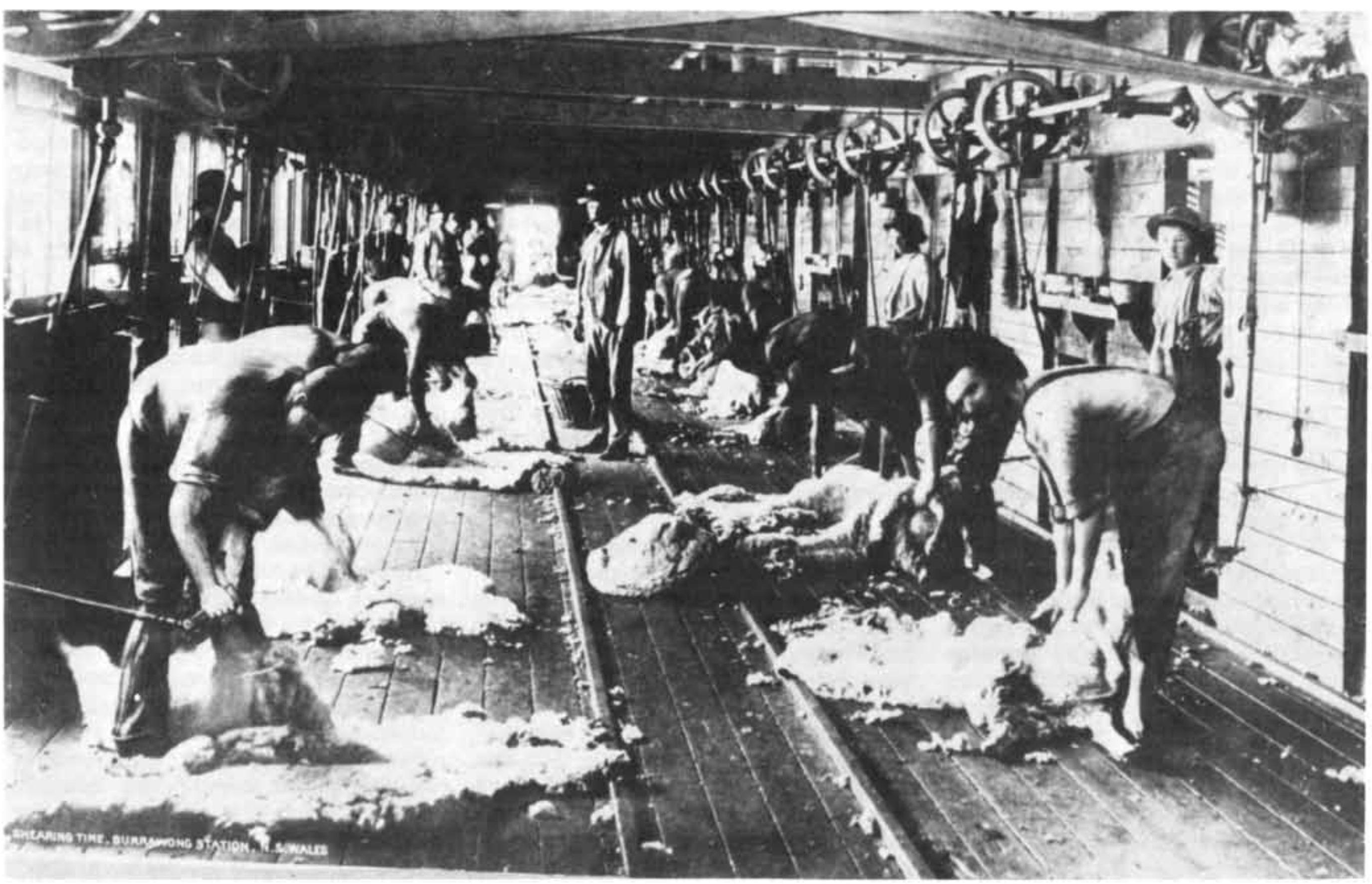
During the early years of Australia's development, numerous small observatories were built to observe the southern skies and in particular the center of the Milky Way. In the 1860's the Victorian government decided to build a major observatory and constructed the Great Melbourne Telescope—then the largest equatorially mounted reflector in the world. Well over a century later, substantially refurbished, it is still in operation at Mount Stromlo Observatory, near Canberra. This early interest in astronomy has led to Australia's remaining one of the world leaders in the field.

Naturalists were attracted to the country's unique plant and animal life, the majority of it long extinct in other parts of the world but preserved due to Australia's isolation. Native marsupials, such as kangaroos, koalas, possums and wombats, the hundreds of different species of native parrots and cockatoos and many of the 1,100 or more species of eucalyptus and wattles were collected and classified. Today the process of identifying Australia's wildlife is still going on; some scientists claim hundreds of species and subspecies may yet be discovered.

Geology was another colonial interest, because of the rich mineral deposits and also because of an extensive debate on the age of the continent. Australia is now acknowledged as one of the oldest land masses in the world. The igneous, sedimentary and metamorphic rocks of the Commonwealth, range in age from the earliest pre-Cambrian (three billion years) to recent geological times (one million).

Government funding for science started with support for geological and mineralogical surveys. The surveys were meant to be utilitarian searches for metals, in particular gold, but geologists also produced systematic geolog-

Shearing time at Burrawong Station, New South Wales circa 1900



Moving a compressor used to supply power for drills in early goldfields in Australia

ical, mineralogical and topographical maps, which formed the basis of the present-day delineation of the country's mineral resources.

The country is now the largest bauxite and alumina producer in the world, with total identified resources of 6,284 million metric tons, and is the largest exporter of iron ore. It also has vast deposits of black coal, with identified recoverable amounts of 300,000 million tons, and extensive natural-gas reserves both on- and offshore. In comparison oil reserves are poor, and more than 30 percent of the country's needs must be imported.

Distinguished Innovators

Many innovators flourished in the heady atmosphere of growth and discovery, often developing inventions at the same time, if not before, their counterparts in Europe and the U.S. One such innovator was Henry Sutton, a music shop proprietor, who designed an electric continuous-current dynamo as early as 1870, constructed more than 20 different types of telephone systems at about the same time as A. G. Bell and carried out the first experiments with heavier-than-air materials for flight in the 1870's.

Another was Lawrence Hargrave, who, despite his isolation from mainstream science and technology, was at the forefront of aeronautical research in the 1880's and 1890's. In 1889 he developed the radial rotary airscrew engine, which formed the basis of the first engines used to power European planes in the early 20th century.

Australia achieved world recognition for the first feature-length narrative film, which perhaps presaged the nation's current vigor in motion picture production. In 1906 a Melbourne impressario, Charles Tait, produced a film of Ned Kelly that ran for nearly 90 minutes. A commercial success in a time when other films ran for only about 20 minutes, it was also the first film to include interior photography.

Government Activity

The organization, sponsorship and goals of science began to change after Australia gained self-government in 1901. A number of Commonwealth government scientific organizations were established, such as the Bureau of Meteorology in 1907, which is now Australia's major forecasting and meteorological research body. In 1926 the Science and Industry Research Act was passed, leading to the most significant development in Australian science—the establishment of the Council for Scientific and Industrial Research (CSIR), which in 1949 became the Commonwealth Scientific and Industrial Research Organisation (CSIRO).

Since its inception in 1926 and particularly since World War II the organization has been responsible for many new scientific developments, particularly in mining and agriculture, including the downstream processing of farm products.

A self-twist spinning machine, 15 to 20 times faster than conventional machines, has revolutionized wool knitting around the world. Another process, Sirospun, has almost halved the cost of weaving yarn. Other processes have made it possible for wool to be shrink-proofed, mothproofed, permanently pleated and given wash-and-wear properties.

Important advances in the control of livestock diseases have been achieved, new cattle for the Tropics have been bred and new pasture plants for eastern and northern Australia have been developed, along with novel methods of biological control. CSIRO has also significantly contributed to the success of the Australian mining and mineral-processing industry with new exploration and processing technologies.

Revolutionary Mining Exploration

Forsayth Oil & Gas NL, a little-known Western Australian exploration and investment company capitalized at only $4.5 million, looks set to move from obscurity into early gold production following its recent huge mining tenement purchases around Lawlers in the Yilgarn, Western Australia.

Forsayth now owns 100 percent of the mining licenses covering an area of 175 square kilometers in which a geological appraisal indicates excellent potential for the development of two gold-mining operations and predicts the discovery of gold, silver, platinum and base-metal deposits. Initial drilling at the old Great Eastern Mine indicates a reserve of 340,000 metric tons at a grade of four grams per ton to a depth of 30 meters. Silver assays of up to 42 grams per ton and platinum assays of up to 3.4 grams per ton have been obtained previously.

Results of percussion and diamond drilling previously completed by Texas Gulf in the vicinity of the old Great Eastern Mine at Lawlers have shown:

1. An in-ground resource of at least 1.36 million grams of gold that could be extracted by open-pit mining operations to a depth of 30 meters.
2. An additional in-ground resource of two million grams of gold to a depth of 250 meters below the abandoned workings of the Great Eastern Mine, which would require an underground mining operation.
3. Potential for additional reserves in the vicinity of the abandoned Caroline and Donegal mines.

Surface sampling and pitting has established the existence of areas of gold mineralization over a length of 20 kilometers within the tenement area, indicating excellent prospects for additional discoveries.

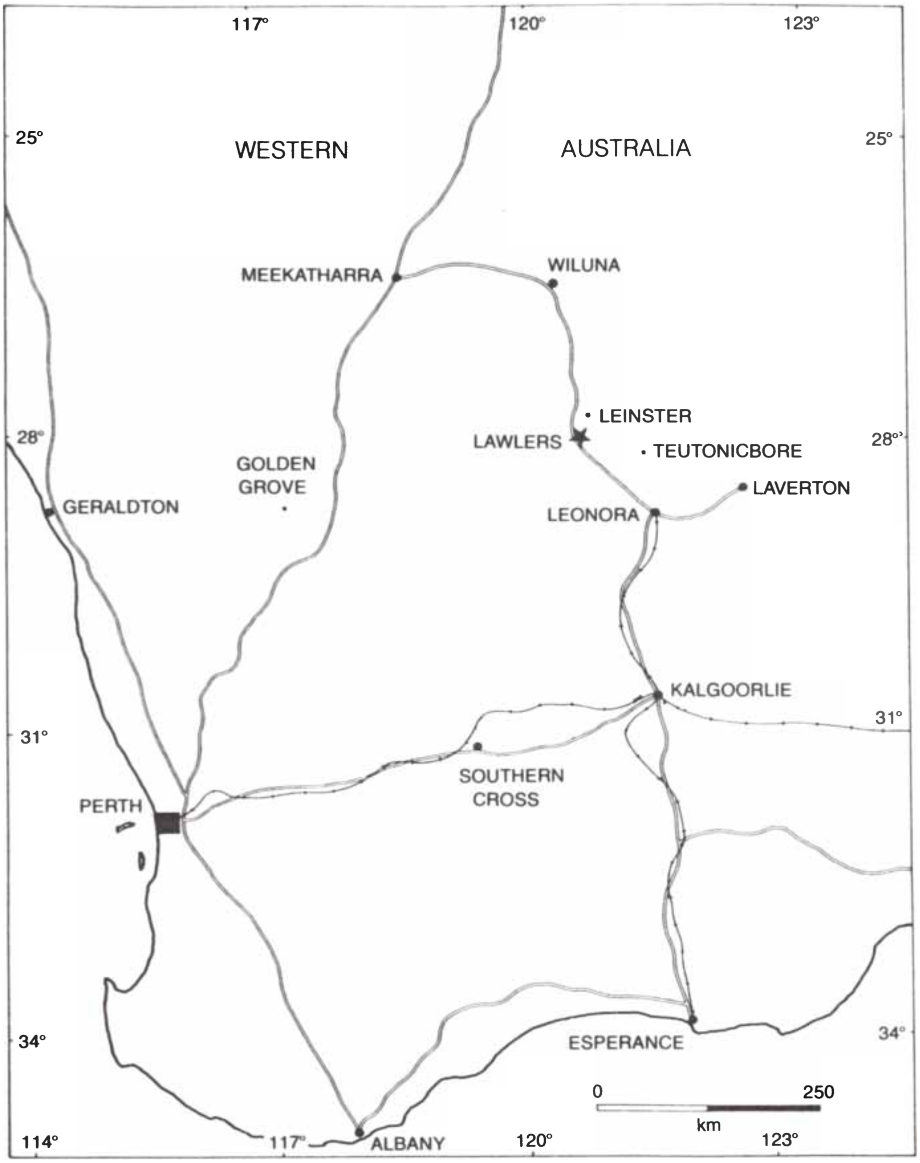
In the southeastern area, diamond and percussion drilling has revealed platinum mineralization of economical grade, up to 3.4 p.p.m. platinum and 3.3 p.p.m. palladium in composite samples representing drill-hole intersections of 25 feet (7.6 meters). The extent of these deposits is unknown, but additional exploration is warranted. Geochemical analysis of the gossans shows values of up to 1.52 percent copper, 8,950 p.p.m. zinc, 3,838 p.p.m. lead and 42 p.p.m. silver.

The in-ground resource of one million grams of gold equal to about $14 million has already attracted the attention of Seltrust, which would like to treat the ore at its nearby Teutonic Bore mine, and of a Perth private company, which has offered to spend up to $2 million on an ore treatment plant on a joint-venture basis, with the promise of gold production within eight months.

Of great significance to Forsayth's tenement purchases is the fact that the exploration was not completed by conventional prospecting and explora-

The Landsat image of the Lawlers region produced by satellite from approximately 600 miles above the earth's surface. Well-defined lineaments and faults are visible to the human eye.

Department of Lands and Survey, Western Australia



Western Australia showing the Lawlers region.

tion methods and could not conceivably have been so completed.

A regional geological appraisal of the eastern goldfields of Western Australia based principally on the geological interpretation of Landsat satellite imagery indicated that the Lawlers district has potential for the discovery of economical mineral deposits. Detailed analysis within this district using the computerized lineament analysis techniques developed by M. J. Longman & Associates showed three areas of interest, one area at Lawlers and two extending to the southeast.

The primary aim of Longman's investigation was to locate, using known mineral deposits as models, areas where the geological environment was potentially favorable for the development of the following three major gold mineralization types:

1. Kalgoorlie Type. Gold associated with quartz veining and sulfide zones with a late-stage differentiate of a basic sill.
2. Kambalda Type. Gold associated with the sulfide or graphite-rich metasediments within a sequence of ultrabasic or basic flows.
3. Porphyry Type. Low-grade, high-tonnage deposits closely associated with late-stage granitic intrusives with copper, molybdenum or tungsten affinities.

Of the areas studied to date the Lawlers district is unique in that the criteria required for each type of deposit exist, with known mineralization confirming the presence of all three geological environments. The regional geological structure is dominated by a trend of north-to-northwesterly lineament zones. Visual interpretation of the lineament patterns visible on the Landsat image suggested that the Lawlers area is close to the intersection of three major lineament zones.

Because the regional structural features indicate the region could be favorable for the implacement of structurally controlled ore zones, a computer lineament analysis was undertaken of the greenstone belt extending from Lawlers to Camel Bore. The analysis, using Longman's techniques for lineament analysis of Landsat data, confirmed that the Lawlers mining area occupies an area of high lineament density dominated by major east-west structures.

Ore zones extend southeast from Lawlers for a distance of four kilometers. A second zone with similar characteristics occurs two kilometers southwest of Camel Bore and extends for a distance of five kilometers to the west-northwest. A third zone extends southeast for a distance of five kilometers from two kilometers northeast of Table Well. All three zones parallel the granite greenstone contact, and interpretation of the computer-developed lineaments suggests that any area within one kilometer of the zones of high lineament density would be favorable for mineralization.

Longman's theories and interpretation were in the event supported by an earlier drilling program by Texas Gulf, which abandoned the Great Eastern Mine area when it appeared its large-scale target tonnages would not be met. Its complete exploration program has, however, proved the potential for a medium-grade operation and saved Forsayth about $700,000 in the process. This has put the area well beyond the exploration stage.

While Lawlers is undoubtedly the feather in Forsayth's cap, the company is not without its other attractions:

1. Cash and quoted investments of $3.5 million, equal to 22¢ per share;
2. A 15 percent interest in a highly prospective oil permit in the Bass Strait, and an interest in ATP 298 in Queensland due to spud-in in mid-October;
3. A small gold joint venture being negotiated with a major mining company;
4. Australian rights to Longman's techniques.

Even with the price of gold languishing at about US $400 per ounce, Forsayth will not be ignored much longer, particularly when infill drilling results and ore reserve figures are expected from Lawlers in the near future.

The Present: The government leads research expenditure as private companies specialize

CSIRO is now Australia's largest scientific research establishment. The organization employs over 7,500 people in more than 100 laboratories around the country, consumes more than a third of the total government science budget and carries out research in all areas but nuclear energy.

Australia's defense research effort began in 1916 with the establishment of an arsenal and a supporting defense laboratory. This effort gradually expanded through the establishment of defense research laboratories, mainly based at Salisbury in South Australia, and through the establishment in the 1970's of the Defence Science and Technology Organisation (DSTO), which now employs over 1,000 professional scientists and engineers.

Another major milestone was the creation of the Australian Atomic Energy Commission (AAEC) in 1953 to help with the exploration, mining and treatment of uranium, the use of nuclear power and the use of radioisotopes for medical purposes.

The organization had its heyday in the 1960's and early 1970's, when there was a strong possibility that Australia would build a nuclear power station and move into enrichment and conversion of uranium oxide.

However, under the policies of the new Labor government, the nuclear research effort is being wound down and work is concentrating on the medical and environmental uses of radioisotopes and on nuclear waste disposal.

Universities also provide research facilities. There are now 19, twice as many as 10 years ago, and despite the increase there are still quotas on student admissions. However, enrollments in engineering and the physical sciences are low, causing concern over future manpower, particularly for high-technology industry and resource development.

Tertiary education is backed by more than 50 vocational colleges and more than 200 technical colleges.

Overall, Australia's science and technology effort is still strongly biased toward the colonial model of astronomy, natural sciences, meteorology, agriculture and geology, although it is also extremely strong in medical research, the development of scientific instruments and mining engineering.

The industrial research effort in Australia is extremely low in comparison with that of other OECD countries: only .2 percent of the gross domestic product was spent on research by industry in 1978–79, compared with 1.9 percent in countries of relatively comparable population such as Switzerland and Sweden.

However, government spending on R & D as a percentage of GNP compares extremely well with that of other OECD countries, at .8 percent. For both Canada and Switzerland the figure is only .6 percent.

This imbalance is accentuated by the relatively large sums that still go to agricultural research, despite the sector's declining input to overall production, contributing only about 10 percent in recent years. However, it still accounts for about 40 percent of export income.

The imbalance is highlighted by CSIRO, which still spends more than a third of its annual budget of over $300 million on agricultural research. In comparison, mineral and energy research receives less than a sixth of the total budget and high-technology about a tenth.

Moves are under way to channel more funds into the manufacturing and high-technology sectors, but a change in direction can only be slow and many argue that CSIRO should continue to concentrate its effort on the primary industry sector, which has served the nation so well. However, the minister for science and technology, Mr. Barry Jones, is also pressing for CSIRO to change its emphasis and, he says, "come out of the closet."

The disparity has prompted government action to try to boost industrial research through a number of grant schemes, in particular the Industrial Research and Development Incentives Scheme (IR&DI), operated through the Department of Science and Technology. It had a budget of $49 million in 1982–83 and made grants to a wide range of industrial research projects, many by emerging small companies. Its budget for 1983–86 is $71.6 million.

Efforts are in progress to establish more specific industrial research funding schemes, particularly in the high-technology area. For instance, earlier this year the National Biotechnology Program was launched specifically to provide funds for the commercial scale-up of biotechnology research.

A similar energy research scheme, operated by the National Energy Research, Development and Demonstration Council (NERDDC), has been in existence since 1978 and has made grants totaling almost $100 million in a wide range of fields from coal exploration to wind power.

At the hub of Australia's research effort is the Department of Science and Technology, formed from two separate departments in 1981. The department, with a total budget of $574.2 million for this financial year, is responsible for a number of Australia's major scientific bodies such as the Bureau of Meteorology, the Antarctic Division, the Australian Government Analytical Laboratories and the Patent, Trade Marks and Design Office.

The Department of Science and Technology is also the main source of policy advice to the minister and is responsible for the government's major industrial innovation and technology development programs. It administers a number of grant schemes, including the Australian Research Grants Scheme (ARGS), which funds basic research, the Marine Sciences and Technologies Research Grants Scheme, the Industrial Research and Development Incentives Scheme and the new Australian Biotechnology Scheme.

Australia's major tropical marine research body, the Australian Institute of Marine Science (AIMS), based at Townsville, Queensland, also reports to the department. AIMS was established in 1972 to carry out a wide range of multidisciplinary research projects, particularly on the Great Barrier Reef. It now has a budget of $6.9 million and its main projects include work on marine food chains, reef-building organisms and coral reefs, tropical oceanography and marine pollution.

The Department of Science and Technology administers Australia's space programs, which include NASA's deep-space communications complex at Tidbinbilla, near Canberra, and the nearby Orroral Valley tracking station, which is to close shortly as part of NASA's rationalization of space facilities around the world.

Tidbinbilla's main facility is a 64-meter antenna supported by a 26-me-

(continued on page A16)

The Microbee

In the battle with the international giants in the highly competitive microcomputer business, one Australian company has been remarkably successful. With almost 20,000 Microbee computers delivered to date, Applied Technology is currently scaling up its production facilities in the New South Wales town of Gosford to handle an output of 2,000 computers monthly. Additional company-owned support facilities are now operational in Perth, Western Australia, and Melbourne, Victoria.

The key to Applied Technology's success was in the specification of its Microbee microcomputer system and the low price it charges for the product. The company was one of the first to recognize the importance of the price/performance factor in key but cost-sensitive markets such as education. Indeed, at the time of this writing, the Microbee has been selected as an approved computer for schools in New South Wales, Queensland and Western Australia and is in widespread use in schools, universities and technical colleges throughout Australia and New Zealand. The success of the Microbee has in turn started to generate a stream of quality educational software utilizing the extensive graphics, sound and input/output features of the machine. The Microbee is thus able to compete favorably with imported microcomputers.

Largely the brainchild of its developer, Mr. Owen Hill, the Microbee was originally sold in kit form. The machine appealed to hobbyists, and it quickly acquired a reputation for being a reliable and surprisingly powerful system. As demand built up the company began marketing fully assembled units using the information gained from the kit customers to fine-tune the final product.

The Microbee is based on a Zilog Z8OA microprocessor, one of the most widely used eight-bit microprocessor chips in the world. The architecture of the computer is another virtue: It essentially consists of a compact keyboard/visual display/computer controller unit that can be expanded to various configurations by adding various types of memory boards. The Microbee was one of the first computers to use battery-backed, CMOS static memory, and the 16K and 32K units are supplied with BASIC, WORDPROCESSING and COMMUNICATIONS software in read-only memory. A user can write a file on the Microbee under WORDBEE, disconnect the computer from the power supply and then take it to the office the next day and print-out or edit the same WORDBEE file. The 64K disk version uses CP/M as an operating system and is supplied with world-class software such as WORDSTAR, MULTIPLAN and BASIC bundled into the incredible $1,500 hardware/software package. The choice of CP/M gives Microbee users access to what is probably the largest base of applications programs worldwide.

Applied Technology recently demonstrated several new products that are likely to ensure that it remains at the forefront of computer technology. These included a fully portable machine using LCD display; powerful and technically advanced peripherals such as the LOW COST DIGITAL MODEM (BEEMODEM); BEETALKER, a fully integrated speech synthesis unit, and BEECOMPOSER, a polyphonic music generator with a built-in editor that enables the user literally to write his sheet music as he composes.

The following of the Microbee is so large that two leading magazines publish monthly Microbee columns on hardware and software developments, and there are eight very active Microbee user groups around Australia with a collective membership exceeding 3,000.

The Microbee is sold internationally under various licensing arrangements. Although the product has been released for worldwide resale for less than six months, it has already captured a significant market share in Denmark, Norway, Sweden and Finland, in Israel and other Middle Eastern countries, and particularly in South Africa. Interest has been shown in West Germany, Hong Kong, the Philippines, India, Indonesia, Singapore and the U.K. Applied Technology is planning to release a version of the Microbee into the U.S. and U.K. markets in 1984, which will be the real test of this product of Australian high technology against the might of the established.

Three Notable Inventions

The Australian Department of Science and Technology provides support for research and development through three main funding schemes:

- Basic research grants through the Australian Research Grants Scheme (ARGS).
- Mission-oriented research grants, e.g. through the Marine Sciences and Technologies Research Grants Scheme.
- Industrial research grants through the Australian Industrial Research and Development Incentives Scheme (AIR & DI).

The following are examples of projects supported by the department through the AIR & DI scheme.

Sirofloc

A new water purification technique, which decolorizes and clarifies turbid water cheaply and effectively, has been commercially developed in a joint project between private industry and the department.

The technique, known as Sirofloc, was invented in 1977 by CSIRO scientists, and has been developed by Austep Pty. Ltd., a joint-venture company formed by Clough Engineering and Davy McKee Pacific Pty. Ltd. in 1979.

Two commercial-scale plants have been built, one in Western Australia and the other in Tasmania. The first, a 35 megaliter plant at the Perth Metropolitan Water Board's Mirrabooka site opened in 1981, and the second, a 20 megaliter plant at Bell Bay in Tasmania, was commissioned last year.

Conventional plants for water purification require the use of flocculants such as alum, which are much slower and less efficient and require large sedimentation tanks.

The Sirofloc process is fast and efficient and uses small particles of magnetite to rapidly remove impurities and color from water.

Impurities in the water are attracted by electrostatic forces onto the magnetite. The particles are then magnetized causing them to aggregate into large clumps that settle out very quickly, leaving behind clean water that is ready for consumption.

Magnetite is recovered for reuse by stripping it from its accompanying sludge using magnetic drum separators and an alkali wash. This reactivates the magnetic particles so that they attract and hold colloidal matter, algae, bacteria, dissolved colored material and other impurities.

The process is particularly effective for purifying water that is hard to treat by conventional methods, such as reservoir water that is colored by dissolved organic matter. The humic acids responsible for the color readily bind to the treated magnetite.

Austep is also commercially developing two other CSIRO water-treatment processes—Sirotherm, for the desalination of brackish water, and magnetic dealkalization, for the removal of temporary hardness from water.

The department has contributed $4.6 million to the development of the three processes. The overall aim is to cut the cost of water treatment and make the most efficient use of one of Australia's most scarce resources.

The Sirofloc plant at Bell Bay, Tasmania is compact and produces 5,280,000 gallon/day of treated water. Major equipment includes contact tanks (left) and clarifier (right).

Bionic Ear

Commercial development of an Australian-designed bionic ear is in its final stages with full-scale clinical trials in both Australia and the U.S.

Work on the bionic ear, or more correctly cochlea implant, has been supported by the department since 1979. A total of $4.7 million has been committed to the project.

It uses sophisticated state of the art electronics and a multichannel implant in the cochlea to convert normal sounds into electrical impulses inside the ear. This helps nerve deaf people with a profound or total hearing loss to recognize speech and hear again.

It was originally developed by Professor Graeme Clark of the University of Melbourne. Initial clinical testing of prototype devices was carried out in 1978–79. In 1981, a leading Australian biomedical company, Nucleus Ltd., was awarded a contract by the department for the commercial development of the implant.

Last year six of the latest models of the device were successfully implanted in Australia with encouraging results.

Two cochlea implant clinics have been established at Iowa University and the Baylor College of Medicine in Houston, Texas, with more planned. Nucleus Ltd. is now planning full-scale commercial production of the implant.

The device consists of a speech processor-transmitter weighing only 200 grams, which fits into a patient's pocket. This receives sound vibrations through a microphone and converts them into a coded radio signal. These signals are transmitted from a small coil behind the wearer's ear through the skin to the cochlea implant, which has been fitted surgically in the mastoid bone behind the ear.

The implant converts the radio signal to electrical impulses which are fed into the multielectrode array in the inner ear. This consists of a bundle of ultra miniature wires that make contact to the auditory nerve endings within the cochlea via a series of electrode rings spaced along the array.

A special connector has been developed that, through a simple surgical procedure will allow the whole electronics package to be replaced when new, more advanced electronic techniques are developed in the future.

The world market for a bionic ear is not known although it is estimated that over four million people are profoundly deaf. A Nucleus survey estimated that 0.1 percent of the population could benefit from the device.

The complete cochlea implant system including implantable prosthesis microphone headset and speech processor.

Interscan

A new Australian-designed microwave landing system should be introduced into most world airports toward the end of this century.

The system, known as Interscan, was chosen by the International Civil Aviation Authority (ICAO) in 1978 as the international standard for landing systems into the 21st century.

Interscan Australia Pty. Ltd., a wholly owned subsidiary of the Australian Industry Development Corporation, and its U.S. partner, Wilcox Electric Inc., have been commercially developing the system with the aid of nearly $14 million in grants from the Department of Science and Technology.

This year, the partners bid for a contract for 208 of the systems for the Federal Aviation Authority (FAA) for use at small U.S. airports, and if successful, they will be well placed to develop a substantial worldwide market.

Already the Australian government has decided to install the system at three country airports to show its capabilities, and a production plant has been set up in Sydney.

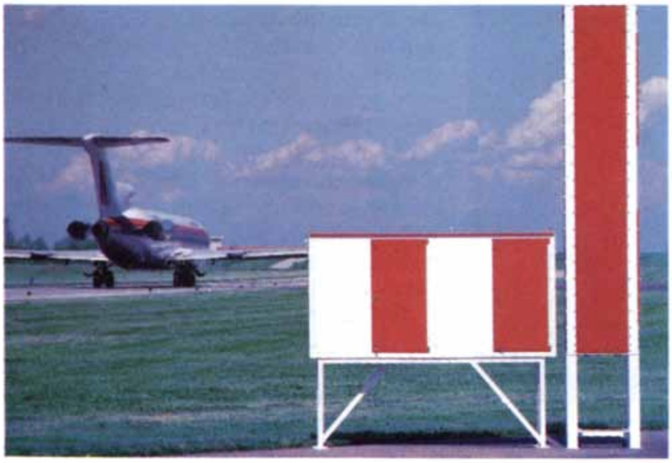
Interscan was originally developed in the 1970's by CSIRO and the Department of Aviation and is based on the principle of microwave beams that intercept and pinpoint the aircraft in the sky.

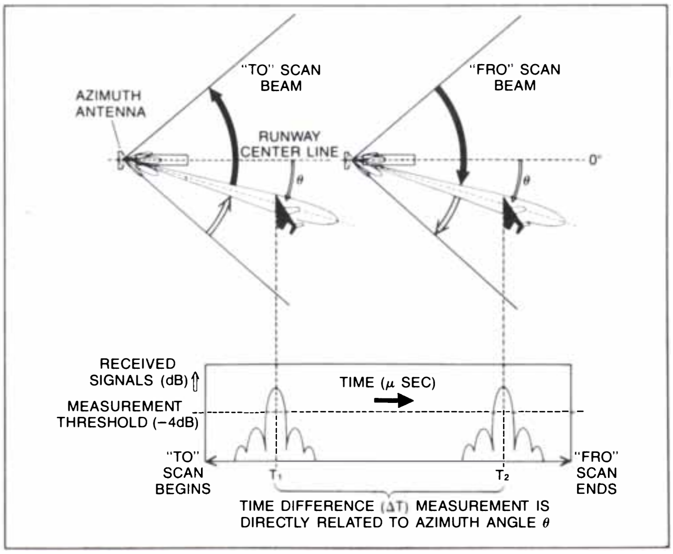
Two ground antennas radiate narrow fan-shaped microwave beams, one scanning horizontally and the other vertically. Each beam produces two pulses in the aircraft's receiver and by measuring the time interval between the two pulses the aircraft's azimuth and elevation angles relative to the runway can be accurately measured.

Precision measuring equipment aboard the aircraft allows the crew to determine its distance from a beacon near the runway. This information, which is updated many times per second, pinpoints the position of the aircraft and allows the pilot to approach the runway from any direction up to 20 nautical miles from the airport.

Interscan has many advantages over continental instrument landing systems as it allows curved approaches, facilitates safe landings in poor visibility, optimizes approach paths and operation of aircraft, enables a wide variety of flight paths for both landings and takeoffs and can be installed at airports with difficult terrain.

The Interscan microwave landing system. The 1½ degree elevation antenna and associated electronics cabinet on display at Kansas City International Airport (left); the scanning principle (right).





ter dish, and the station has taken part in many deep-space missions, including the Mariner, Pioneer, Viking and Voyager flights.

Another element of Australia's space program is the Australian Landsat station with a data acquisition facility at Alice Springs and a processing facility in Canberra. It has been modified recently at a cost of $600,000 to receive some data from the Landsat 4 satellite, but there is disagreement as to whether an additional $10 million should be provided to upgrade it to receive high-resolution data from the new series of Landsat and French Spot satellites.

The minister for science and technology is also responsible for CSIRO. The organization is split into five institutes and a bureau. These include the Institute of Biological Resources, Energy and Earth Resources, Industrial Technology, Animal and Food Sciences and Physical Sciences and the Bureau of Scientific Services. Within this structure are 40-odd divisions and units concentrating on specific areas of research from radio astronomy to manufacturing technology.

Current high-priority research areas include biotechnology, plant pathology, energy, manufacturing industry, water and soils, information technology and oceanography.

Under pressure from the minister for science and technology the organization is redirecting its work into a number of specific areas of industrial research, including the use of computer technology and microelectronics for industrial processes, flexible manufacturing systems, advanced-process and quality-control technologies, advanced materials, mineral-processing technologies, environmental-monitoring and mineral-exploration instrumentation, communication technologies and agricultural chemicals, veterinary vaccines and medical pharmaceutical products.

CSIRO has also established a new contract research company, Sirotech, to help transfer CSIRO innovations from the research laboratory into commercial manufacture.

However, CSIRO is not the prime source of science and technology policy advice to the government. Since 1977 this has been the role of the Australian Science and Technology Council (ASTEC), which is responsible directly to the prime minister.

ASTEC is headed by Professor Ralph Slatyer of the Research School of Biological Sciences at the Australian National University and a former Australian ambassador to UNESCO.

Apart from the Department of Science and Technology, another 12 or so federal government departments are involved in science and technology initiatives—accounting for total federal government spending on science of $1,552 million in 1982–83.

The Defence Science and Technology Organisation (DSTO) within the Department of Defence, is Australia's second-largest research organization after CSIRO, with a budget of $145 million for this financial year—about 3 percent of the total defense budget. It employs about 1,000 professional scientists and engineers and an additional 3,400 technical and support staff at 10 research establishments around Australia. Research is mostly in the physical sciences and engineering, but is also undertaken in food sciences, environmental psychology and operational research.

Another department heavily involved in science and technology is the Department of Resources and Energy, spending $89.2 million over the last financial year. It is responsible for both the Australian Atomic Energy Commission (AAEC) and the Bureau of Mineral Resources (BMR).

The bureau, with a budget of $23.1 billion for 1983–84, is the country's main geoscience research organization. It carries out a wide range of geological studies both on- and offshore and in Antarctica. It is organized into five divisions: continental geology, geophysics, marine geoscience and petroleum geology, and petrology and geochemistry; the fifth division is devoted to the assessment of Australia's mineral and energy resources.

The AAEC is a statutory authority with a budget of $38.3 million this financial year. With a staff of about 1,000 it is the largest energy research body in Australia.

In the past the AAEC has concentrated almost a quarter of its staff on uranium-enrichment research, but this has now been wound down in line with the Labor government's policy on uranium mining and use. Research is now being concentrated on waste disposal, radioisotope use and production, nuclear physics and fusion and environmental programs.

The government has plans to split the commission into three separate authorities in the future: one responsible for regulatory matters, including environmental protection, health and nuclear safeguards; a second, a government corporation to produce and market radioisotopes, and the third, a nuclear research science authority.

The Department of Health is responsible through the National Health and Medical Research Council (NH and MRC) for the majority of medical research funding in Australia, which this year has a budget of $38 million, 28 percent more than last year.

Last year the Australian Science and Technology Council recommended great revisions in the funding, coordination and planning of medical research, pressing for the abolition of the NH and MRC and the establishment of a Medical Research Council and a National Health Council. This action was strongly opposed by the Australian Medical Association and no decision has been made.

The state governments are all heavily involved in science and technology development. In the past this has been mainly limited to agricultural research and extension services and medical research, but more recently it has extended to the support of high-technology development.

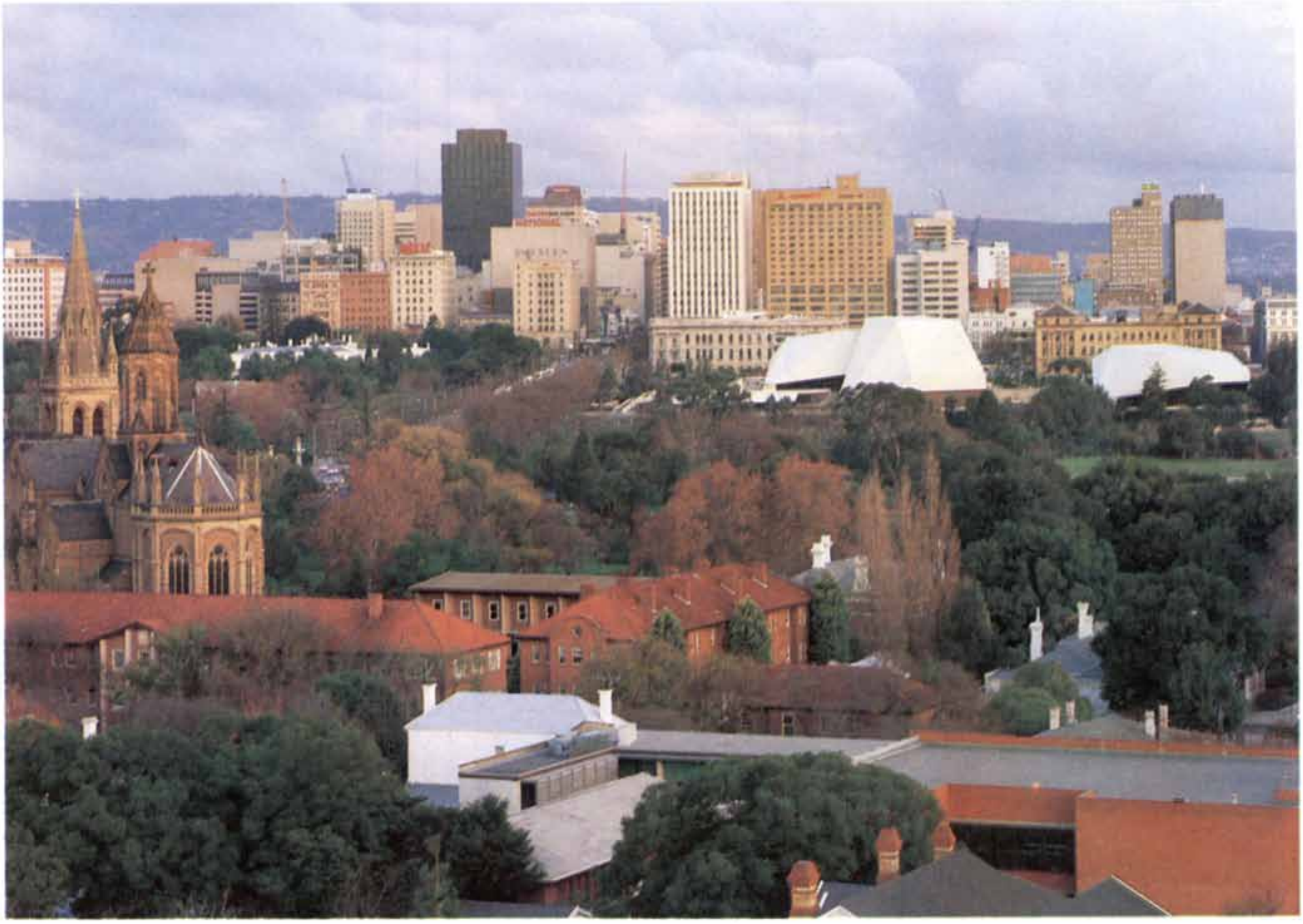
All of the states are moving strongly into this area by establishing technology ministries, technology parks, innovation centers, special technology funds and information services.

The universities receive a great deal of their funding for basic research through the Australian Research Grants Scheme (ARGS). However, the funding for this, $22.4 million for 1984, is considered far too low to support an

(continued on page A20)

Tidbinbilla tracking station, near Canberra.

Determined High-Technology Development



Adelaide, South Australia: ". . . possibly the last, well-governed and moderately contented metropolis on earth." The New Yorker

South Australia likes to call itself the Festival State. It is also the central state and the self-styled high-technology state of Australia. To the east of South Australia lie the more populated states of Victoria and New South Wales. To the north is the vast Northern Territory, and to the west lies the even larger and emptier state of Western Australia.

Of South Australia's total area of 380,000 square miles, about two-thirds in the north is desert or near-desert. Recent discoveries of major natural resources include the Cooper Basin natural-gas fields, which now supply parts of New South Wales, and the massive copper, gold and uranium deposits at Roxby Downs.

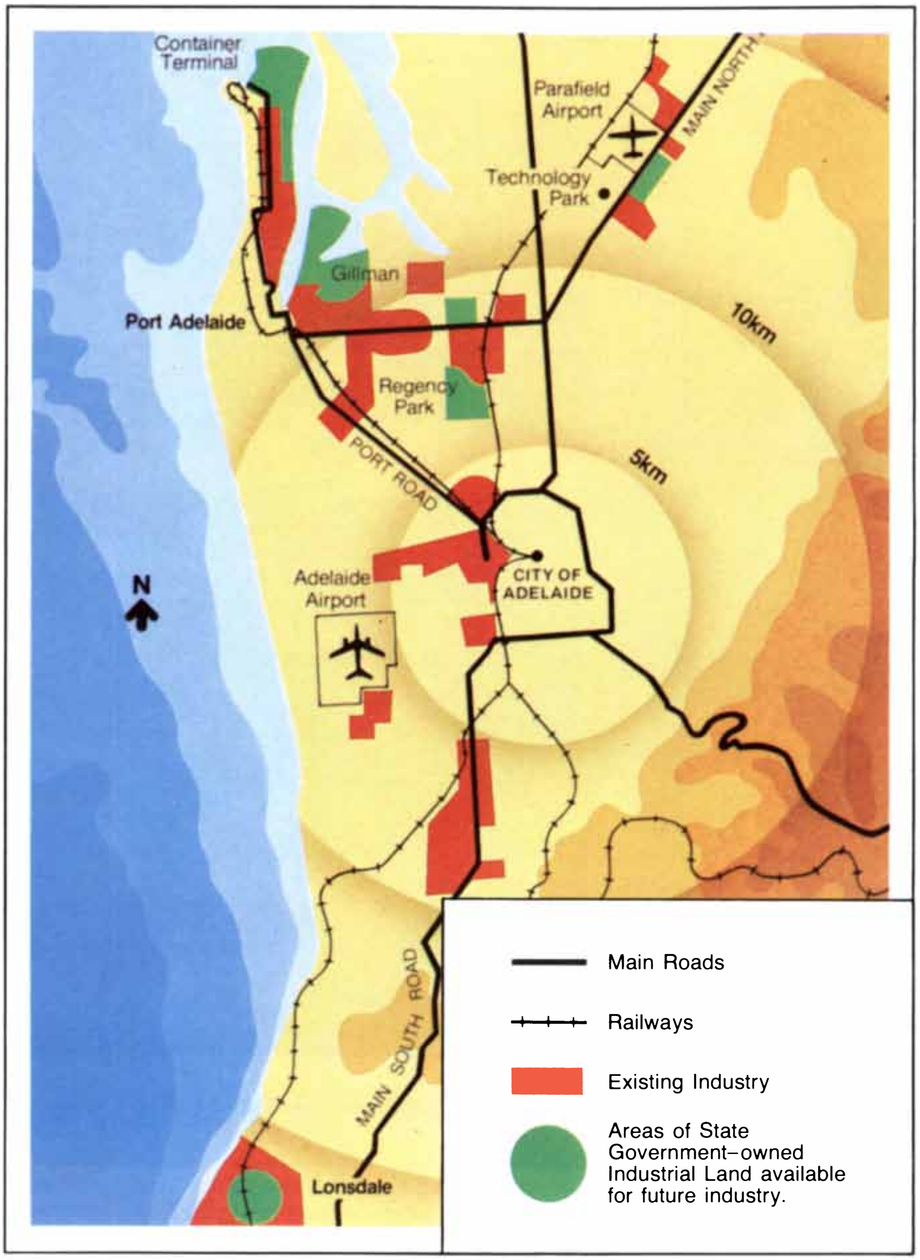
Historically, South Australia has been highly dependent on agriculture for its wealth. Expansive grazing lands in the center of the state support a large population of sheep and cattle. In the south are temperate farming lands, and around the capital city of Adelaide are important wine-growing regions.

During the 1950's and 1960's, South Australia made a determined bid to reduce its dependence on agriculture and increase industrial activity. Incentives were offered to manufacturers to relocate to South Australia and a policy of encouraging migrants was actively pursued. Major automotive and white-goods manufacturing operations were established, and South Australia is now the third-largest manufacturing state in Australia.

In the 1980's South Australia has been taking the next step in socioeconomic development. Because the major service-based industries are in Sydney and Melbourne, Adelaide had to follow a different route. Rather than chase tertiary industry wholesale, South Australia is embarking on an ambitious program to attract high-technology companies. The move to high technology is designed to broaden the state's traditional base, much of which is facing intense competition, while at the same time improving the competitiveness of existing industry.

South Australia has had a headstart in the battle to attract high-technology companies; its vast empty interior attracted the U.K. and Australian governments to set up a rocket range and nuclear bomb testing site there in the late 1940's and early 1950's. These moves into advanced technology occurred at about the same time the state was expanding its manufacturing base. From this defense-related activity sprang a diverse range of high-technology enterprises, which the state is determined to capitalize on.

About a 20-minute drive from Adelaide lies the largest electronics research establishment in the Southern Hemisphere. The Defence Research Center, Salisbury (known as DRCS), forms part of the Australian government's Department of Defence. DRCS employs 2,700 people in three major research laboratories: electronics,



The major industrial areas, Adelaide, South Australia.

weapon systems and advanced engineering, working on indigenous defense equipment development and modification and evaluation of overseas equipment. This center has advanced skills over a wide range of high-technology areas such as optics, infrared, radar, aeroballistics, fluid dynamics, propulsion, guidance and control systems, electronic warfare, underwater detection, communications, precision mechanisms and many others.

DRCS involves industry in projects at an early stage. As a result, a number of private companies have developed around the organization. Some of these companies are almost completely devoted to defense work whereas others are also involved in commercially oriented high-technology work. Major companies such as Thorn-EMI Electronics, British Aerospace and Fairey Australasia provide equipment to U.S. and European defense contractors. This has given South Australia more of an advanced technology flavor to its industry, concentrated in innovative R&D and complex systems development and manufacture, than any other state in Australia.

The South Australian government is keen to exploit and build on the advanced level of existing high-technology expertise. The areas of expertise are considerable, but they cover in the main defense-related technologies, electronics and optoelectronics, biotechnology, mining-related technology, agriculture-related technology and remote sensing. In addition to highly capable, indigenous high-technology companies, South Australia boasts such international names as British Aerospace, Philips, Raytheon and Texas Instruments.

In the area of biotechnology, South Australia has extensive expertise particularly through the University of Adelaide with its National Centre for Gene Technology and the internationally known Waite Agricultural Research Institute, Flinders University, Roseworthy Agricultural College, the South Australian Institute of Technology and various CSIRO divisions based in Adelaide. The state government is in the process of formulating a South Australian Biotechnology Advisory Committee to provide an important link between the scientific community and industry to assist in the commercial development of biotechnology.

Adelaide is also home to Australia's largest manufacturer of H.F. communications equipment, Codan Pty. Ltd., which is diversifying into small earth terminal design and production. Codan has received a significant federal government research and development grant to aid its efforts in innovation.

The state government has invested $4 million in developing Australia's only industrial/office park complex designed to meet the needs of high-technology manufacturers and organizations involved in research and development, and established a separate corporation to promote the development of high technology. Set in a landscaped area near DRCS and next to the main campus of the South Australia Institute of Technology, Technology Park Adelaide appears ideally suited to meet the needs of technology-intensive companies.

The park will include a multitenant complex for small companies, and this building will also house an innovation development advisory service. The technologies of future tenants span defense electronics, plastics technology, robotics, mining technology and communication.

The size of Adelaide and its excellent infrastructure is conducive to a flow of high-tech start-up ventures, a number of which have already committed to occupancy of Technology Park Adelaide.

The first separate tenant will be Duntech International, involved in antenna design and audio equipment. However, Technology Park has received a substantial boost from the recent announcement by British Aerospace, a major multinational company, to establish its Southern Hemisphere headquarters at the park, including design and engineering facilities and the progressive addition of production facilities.

British Aerospace will purchase 3.2 hectares of land at Technology Park, and the expansion is likely to lead to a substantial employment boost as a result of offset contracts from the FA-18 and Orion programs. British Aerospace sees the park as an excellent base location for future work associated with space projects in Australia. The company is presently involved in the design specification of the electronics package for Australia's contribution to the Starlab space satellite.

Encumbrances on intending purchasers or lessees of property within Technology Park will ensure that the integrity of the development as a high-technology center is kept.

The executive director of Technology Park Adelaide Corporation, Mr. Barry Orr, expects tenants of the complex to benefit from the location, the sharing of commercial services and the research resources of the South Australian Institute of Technology. Also supporting Technology Park Adelaide are several other major academic institutions. These include Flinders University, which houses the Institute for Atomic Studies, and the Waite Agricultural Research Institute at the University of Adelaide.

One of the most exciting areas being investigated in South Australia is the CSIRO project on the development of very-large-scale integrated circuits (100,000 circuits per chip) by Dr. Craig Mudge.

Adelaide is also the base of the Australian Mineral Development Laboratories (AMDEL). Similar to the U.S. Colorado School of Mines, AMDEL is Australia's largest mining-related contract research and technical consulting organization.

Aside from the support available from the well-established tertiary institutions, Technology Park Adelaide is also likely to benefit from the existing manufacturing structure in the state. Heavy involvement in automobile and white-goods manufacturing has developed skills in areas that include molding and extrusion of plastics, toolmaking, presswork and aluminium casting, electroplating and sheet metal working capabilities.

Of considerable interest to any company attracted to South Australia is its economic and social advantages. Although labor costs in the state and elsewhere in Australia are similar to those in other advanced Western countries including the U.S., South Australia has a marginally lower cost structure and a significantly lower rate of industrial disputes. This brings the overall labor rates in the state well below those in the other Australian states.

Housing costs are also generally lower than in other parts of Australia and considerably lower than, for example, those in the Santa Clara Valley region of California. The median price of a quality family-sized home on its own block of land in the Adelaide suburbs is $55,000. The population of the metropolitan Adelaide area is about 900,000, and the region has a well-planned and well-maintained road and public-transport system. Surveys have shown that 70 percent of the working population of the Adelaide area live within 30 minutes' commuting time to the central business district. For those living to the north of the city, the traveling time to Technology Park would be less.

Adelaide was once described by *The New Yorker* magazine as "possibly the last well-planned, well-governed and moderately contented metropolis on earth." It certainly has less of the bustle of Sydney or Melbourne, more sophistication than Brisbane and perhaps a touch less brashness than Perth. Adelaide is on a coastal plain surrounded by a low range of hills. The climate generally is conducive to comfortable living.

The city's cultural activity is widely noted. Adelaide has wholeheartedly adopted the idea of having regular festivals. Every second year it organizes a festival of concerts, opera, ballet, drama, art, literature and light entertainment that is achieving a standard of world renown. Professionals delight in Adelaide's clean environment, wealth of attractions and excellent quality of life. These features give some justification to the city's potential for further high-technology development.

adequate level of pure research in the country.

Recently the chairman of ARGS, Professor Peter Sheehan, has called for a substantial increase in funding to at least $40 million even if only the most promising projects are to be supported.

The colleges and institutes have been more successful in attracting funds from industry. There are now 22 Tertiary Consulting Companies—only three based in universities—carrying out work for private enterprise. Some have been so successful that they now earn fees in excess of $2 million a year for consulting.

Other organizations involved in policymaking are a number of professional bodies such as the Institute of Physics and the three science academies. These include the Australian academies of science, technological sciences and social sciences.

A number of research associations, such as the Australian Welding Research Association and the Sugar Research Institute, support research in their particular industry area. Currently, the government supports four out of the 14 research associations through a special budget allocation.

Australian Company Profiles

CSL

Commonwealth Serum Laboratories (CSL) is a government-owned pharmaceutical and biotechnology research, development and manufacturing organization, based in Melbourne.

It was established by the government in 1916 to provide essential vaccines, sera and other biological materials and since then has expanded to become Australia's major manufacturer of pharmaceutical products.

AMDEL

Australian Mineral Development Laboratories (AMDEL) is an independent, commercial organization, based in Adelaide, that carries out research and development, design, manufacturing and consulting, mostly for the mineral and mining industry but also for government and other sectors of industry both in Australia and overseas.

AMDEL has a wide scope of activities, ranging from mine planning, chemical metallurgy and forensic investigations to the design and manufacture of high-technology instrumentation, environmental studies and computer services. The organization has laboratories in Adelaide, Perth, Melbourne and Townsville and has plans for another facility at Darwin. The staff of 320 is headed by Mr. Brian Hickman, managing director, and is split into a number of divisions including analytical chemistry, applied technology, mineral and materials sciences, fuels and geological services.

AMDEL is responsible for the development and commercialization of a wide range of advanced instrumentation. Recently this has included Sirolog borehole logging probes (initially developed by CSIRO), programmable mineral analyzers, a xanthate monitor and monitors of lead in the air.

In 1981–82 a total of 17 percent of AMDEL's income came from overseas, substantially bolstered by a $1 million sale of instrumentation to the Soviet Union.

Mount Isa Mines Ltd.

Mount Isa Mines Ltd. (MIM) operates one of the largest copper and zinc-lead-silver mines in the world at the isolated township of Mount Isa in the far west of Queensland. The site also produces copper anodes, crude lead and zinc concentrate, which are further processed by associated companies in Australia and the U.K. Group annual revenues are more than $750 million. Operating more than 1,000 kilometers from the nearest major town, MIM has had to develop a high degree of self-sufficiency. In doing so, it has also developed a number of techniques that have found ready acceptance in other major mining companies elsewhere in the world.

Where MIM is not losing a competitive advantage, positive efforts have been made by the company to market its innovations. One of the best examples of this was a process control software program developed by the company for enhancing the control of various processing stages of the ore it mined. The value of the package was recognized by Hewlett-Packard Co., the major U.S.-based computer and instrumentation company.

MIM developed the software in response to its own need for a system suitable for direct digital control, supervisory set point control and model-based control system applications. The company also needed a system that was easy to use and could be connected to a wide variety of instrumentation. MIM decided to implement the system on a Hewlett-Packard computer.

HP bought and further developed the software, adding a few features and slightly altering the configuration to achieve a more flexible package. Now known as Process Monitoring and Control 1000, the package has been acquired by Exxon, Northern Telecom, General Motors, Kodak and Rockwell in the U.S. General Electric has acquired PMC/1000 for divisions in the U.S. and Canada, and Agfa-Gevaert is also using the program in Belgium.

MIM and its associated companies have also sold technology to U.S. and European metal-producing companies and have licensed for manufacture and sale specialized environmental monitoring devices for use in the lead industry. Also in the lead industry, MIM has with Australia's CSIRO jointly patented a direct lead smelting process. The process is at the stage in which a pilot plant is being built. MIM believes the technology will be commercially viable in the early 1990's, when it will reduce energy costs, increase metal recoveries and allow lower-grade ores to be mined.

CRA Ltd.

CRA Ltd. rides near the top of Australia's list of largest publicly owned companies. With sales of about $2.5 billion, its interests reach from the mining and processing of base and precious metals to biotechnology research. While CRA is often regarded as simply a resource development company, its vice-president of development, Dr. Geoff Bone, considers it a manufacturing concern.

CRA's history as a processor of raw minerals goes back to its origin as the Zinc Corporation Ltd., which was formed in 1905 to treat zinc-bearing tailings at Broken Hill, in the far west of New South Wales. The town is the site of one of the largest discoveries of zinc, lead and silver.

Broken Hill was arguably the first place in the world to develop the flotation process for the removal of metal-bearing compounds from raw mineral aggregates. A subsidiary of CRA, Broken Hill Associated Smelters, developed the first continuous lead refining process at Port Pirie, South Australia. The key to the continuous process is the addition of zinc to the impure lead at near melting point. The zinc and silver combination floats to the outer shell of a container called a desilvering kettle, and the lead is collected from the center of the vessel. In recent years Broken Hill Associated Smelters has also developed a process for removing bismuth from lead, thus improving the purity of the metal.

CRA's majority-owned subsidiary, Comalco, mines bauxite in the far north of Australia and processes it into aluminum products. Completing the company's portfolio of activities is its participation in a diamond mining operation in Western Australia.

Dr. Bone notes that most research in Australia tends to be of a "hands-on" nature. "In the future we are also likely to continue to focus on problem areas through the adaptation of existing technology rather than break new ground," he says.

(continued on page A22)

Telecommunications Pioneer and Leader

Australia is the fourth-largest user of the global communications satellite system.

Australia is fast becoming the hub of an extensive communications network in Southeast Asia. International companies that have interests throughout the region are leasing private voice and data links that reach out from a sophisticated, computer-controlled switching center in Sydney.

This new development in high technology has been won by Australia's Overseas Telecommunications Commission (OTC) against strong competition from the communications carriers of several other countries. Its success reflects Australia's highly developed communications infrastructure, and the country's long history of stable government.

The network provides secure, dedicated links between companies' Australian operations and their Southeast Asian affiliates in the Philippines, Indonesia, Singapore, Malaysia, Thailand, Hong Kong and other regional sites. Further connections are being established in some cases to link the hub in Sydney to the U.S. or to Europe to provide a worldwide network.

OTC expects this type of international networking for intracorporate traffic to grow, and it is actively working with several major multinational companies on a number of more specialized networks for the region, including an extensive international facsimile hookup.

Research done by OTC indicates that business managers in Australia perceive the greatest challenge in communications to be how best to establish corporation-wide links, and that these improved internal communications will be a vital factor in forging a competitive edge.

Australia has long been a major user of international telecommunications, partly because of its isolation and partly because of the diverse makeup of its population resulting from the heavy immigration programs and the population boom in the post–World War II period.

OTC was formed to handle international communications in 1946, at the end of the war, because of public agreement that a nation's strategic communication links should be under government control. Thus the government acquired the private interests in the existing radio network and undersea cables. Since that time OTC has provided all of Australia's public overseas communications, progressing from the era of the telegraph cable to submarine telephone cables, to satellites and now fiber optics.

OTC is the fourth-largest shareholder of the global communications satellite system, based on "user pays" arrangements. Australia is among the highest per capita users of telecommunications in the world, behind only the U.S., Britain and France in total use of the system.

In every sense, Australia is part of the global trend toward the information age. Social use of the telephone in Australia has been growing at a steady 30 percent per year, slowing only slightly during the recent economic downturn, and new business services have been introduced for high-speed transmission of data across the country or across the world.

Since its formation, OTC has grown in stature in the international telecommunications community. It initiated the permanent leasing of international television broadcast transponders. Two major television networks in Australia now have 24-hour access to television program material, such as news reports, from North America, keeping Australians informed of the latest developments overseas.

Through its Intelsat involvement OTC is abreast of the progress being made in satellite communications, but it also shares the ownership and operation of substantial submarine cable systems around the world. Currently under construction is the $400 million ANZCAN cable linking Australia and New Zealand to Canada via Norfolk Island, Fiji and Hawaii. Australia is the major partner in this project. It is also involved in the planning and development of a cable system across the Indian Ocean, which is due to be completed in 1986.

OTC is test-marketing some novel forms of telecommunications. Together with the U.S., Japan, Canada and the U.K., it is introducing an international videoconferencing technique that dramatically reduces the bandwidths and cost involved in satellite television transmission. The technique involves digitizing information and then transmitting only the information that changes from one picture frame to the next, not full frame information. OTC has also recently introduced the first electronic messaging services in Australia, based on the U.S. Dialcom facility.

OTC's progressive attitude toward service innovation and its experience in the telecommunications forums of the world will ensure that the most modern, convenient services will continue to be available for business and social use in Australia.

The Broken Hill Pty. Co. Ltd. (BHP) is Australia's largest publicly owned company and has developed beyond a mining company to an integrated steel and metal products manufacturer. Its interests include oil and gas exploration, coal mining and the creation of a new type of internal combustion engine.

BHP's manager for new technology, Dr. John Parrott, believes Australia is evolving beyond its historical role as the "quarry of the world." "We are now managing our own resource deposits and applying technology to that management process," he says. BHP has developed novel processes with the $26 million budget it allocates to research and development.

A central research laboratory is maintained at Shortland, New South Wales, about nine kilometers from the company's Newcastle steelworks. The major aim of the laboratory is research into the raw materials and processes used in all of BHP's operating divisions. Of particular interest is work being done on the hydrometallurgical and electrometallurgical processing of ores and concentrates.

In the Melbourne suburb of Clayton, BHP operates a research laboratory devoted to work on the maintenance and expansion of the company's existing markets. Nearly half the research effort at Clayton is devoted to improving the properties and quality of existing steels produced by BHP and to the development of new steels.

BHP scientists are involved in the research being conducted on the conversion of natural gas and light hydrocarbons into liquid transport fuels. (Australia's natural gas reserves are known to exceed 9×10^9 cubic meters.)

Dr. Parrott says the basic technology used by BHP in commercially developing liquid fuels from natural gas was researched and patented by CSIRO. The process does not use the production of methanol as an intermediate step but rather synthesizes hydrocarbon molecules directly from methane via butane. "The work in this area was a spinoff from our coal-to-oil project," Dr. Parrott says. BHP is spending about $1 million this year on the gas-to-oil project. Although BHP admits it may be too early to define the commercial prospects for the process—production costs are an inhibiting factor—the company believes the process has exciting long-term potential.

ERA Computer Corporation Pty. Ltd.

Dr. Bill Caelli left the Australian subsidiary of Control Data Corporation, the major computer company, in the late 1970's with the idea that Australia could play a significant role in the computer industry.

The company he founded, ERA Computer Corporation Pty. Ltd. (known as Eracom), began making small microprocessor-based commercially oriented computer systems. But it is not in this area that Eracom is likely to find its biggest success. While working with microprocessors, Dr. Caelli began investigating the security problems associated with the transmission of data over public telecommunications networks.

Eracom believes it has gone some way toward solving the problems of keeping data secure with the development of an encryption device claimed to be better than any similar device in the world. The device has attracted the attention of several companies in the U.S., which are ready to market the product there and, if necessary, build it under license. Also giving the product some attention is the world's largest computer company, International Business Machines Corp. (IBM). This company's Australian subsidiary recently asked Eracom to jointly bid with it for a major computer and communications network required by the police department in the Australian state of Queensland. (Eracom is now based in the southern Queensland resort city of Gold Coast.)

While leading in the development of data encryption devices, Dr. Caelli also has strong ideas about high technology in Australia. "Australia has to look for viable and well-rounded participation in new industries rather than concentrating on some esoteric market niche that may be here today and gone tomorrow," Dr. Caelli recently told computer engineers. He says that the information technology industries are ideal for viable job creation because they include computer and communications manufacturing and support. "Success owes more to marketing, manufacturing, management and maintenance than it does to a world first in anything," Dr. Caelli remarks. He says Australia must start manufacturing machines similar to the Apple and IBM microcomputers. "The point is that currently successful big technology such as Apple are not in the 'scientific breakthrough' league."

Oliver J. Nilsen (Australia) Ltd.

A diverse electrical equipment manufacturer, Oliver J. Nilsen (Australia) Ltd., commonly known as Nilsen, has taken a major plunge into a novel area: developing, manufacturing and marketing an advanced ceramic product, which is known as partially stabilized zirconia (PSZ).

The PSZ process was originally developed and patented by Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO). Nilsen's associated company, Nilsen Sintered Products, has the exclusive world license for the PSZ process and has been working with CSIRO on the development of PSZ products since 1975. PSZ is likely to have widespread applications in the manufacturing of extrusion dies and internal combustion engine components—both of which will benefit from PSZ's toughness, wear resistance and rapid heat dissipation properties.

Interest in PSZ has already come from a major diesel engine manufacturer in the U.S., and Nilsen Sintered Products has set up a branch office in the U.S. to further the marketing of PSZ products. Further PSZ development is under way in Australia, and CRA Ltd., a major Australian mining and manufacturing group, has acquired a 50 percent stake in Nilsen Sintered Products for an initial $2 million investment and the promise of an additional $5 million over the next five years.

Nilsen has annual sales of more than $60 million and spends more than 10 percent of its income from PSZ on further research and development of the product. The Nilsen group's managing director, Mr. John Nilsen, believes the development of PSZ products will generate more export revenues for the company, which has traditionally relied on its domestic market for the majority of its revenues.

"If we succeed with PSZ, about 75 percent of our sales of the product will be overseas," Mr. Nilsen says "The biggest potential users are in the U.S." Mr. Nilsen recognizes that Australia's distance from major world markets tends to be a problem with having a product on the leading edge of technology in a particular market.

The company's other products include electrical switchgear, porcelain electrical fuses and electric motor control equipment. It also operates a highly successful commercial radio station. "The future lies in the development of new materials and technologies that have potential," Mr. Nilsen says. "We're involved in PSZ because it gives us an edge in a particular market. We like the idea of having a few high fliers in the company, knowing that not all of them will work, but we are confident about PSZ."

F. H. Faulding & Co. Ltd.

For one of Australia's oldest drug companies, opening up an export market was seen as the only way the company could continue to grow. To compete successfully in the pharmaceutical business on an international scale is a daunting prospect, but F. H.

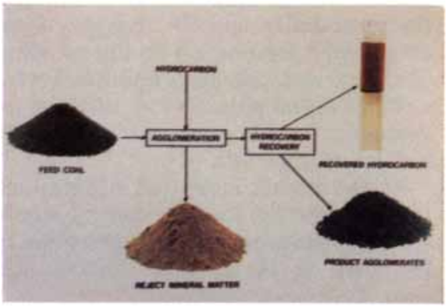
(continued on page A26)

A Revolutionary Coal Transport System

The 30 ton per hour pilot plant at Stockton Borehole Colliery, Taralba, New South Wales.

BP Australia, known corporately as "The Quiet Achiever," has diversified its interests over the past few years and is now heavily involved in the development of Australia's vast resources. Australia has some of the world's largest deposits of coal and BP Australia, working with other companies, is dedicated to developing those coal resources. The company is therefore vitally concerned with developing new, more efficient techniques to overcome the many problems that arise in this challenging field.

One of the most exciting joint ventures the company is undertaking is the transportation of coal by pipeline. Known as IPTACCS (integrated pipeline transportation and coal-cleaning system), the scheme was started as a means of lowering the costs associated with the recovery of coal from the New South Wales and Queensland coalfields and transporting it to steelworks, power generation stations and potential coal conversion plants.



The IPTACCS process.

In the IPTACCS process, coal fines are mixed with oil and water. The oil preferentially wets the coal fraction, causing it to flocculate. Turbulence compacts the flocs into agglomerate pellets about four millimeters in diameter. Although different methods of agitation can be used to effect the growth of the pellets, the IPTACCS joint-venture partners found the action of pumping the slurry through a pipeline to be the most satisfactory method. At the end of the pipeline, the oil is removed and the pellets are dried. In removing the liquids, it is possible to either retain or destroy the pellet structure. Hence suitable products can be made available for export, or prepared as pulverized fuel for power station boilers and industrial furnaces and kilns.

BP Australia and the Broken Hill Proprietary Company worked independently on various aspects of this technology until 1979, when they decided to pool their experience and technology. BP's area of interest was in the oil agglomeration process as a better method of cleaning coal. The company's researchers concentrated on perfecting the removal of the oil from the end product for recycling.

One of the major advantages of the IPTACCS technology has been its successful application in the recovery of coal fines that otherwise would have been lost as part of the normal coal-washing process. A pilot plant in New South Wales is operating at 30 metric tons per hour and has demonstrated the viability of the IPTACCS process. Because of the high coal matter recovery achievable using oil agglomeration, IPTACCS technology appears useful as a means of beneficiating steaming and coking coals.

Major potential outlets are long-distance coal transportation and beneficiation, coal recovery from rejects in existing washeries, treatment of coal fines when other techniques are not viable and improving specification yields from coals with poor washing characteristics.

The Development of Pharmaceutical Biotechnology in Australia

by Neville McCarthy,
Director of the Commonwealth Serum Laboratories

The Australian pharmaceutical industry was accurately described in 1979 in the Ralph report, the "Pharmaceutical Manufacturing Industry Inquiry," as predominantly engaged in the marketing of products formulated and packaged from imported active ingredients.

This general description, although accurate, has its notable exception, the government-owned pharmaceutical manufacturer Commonwealth Serum Laboratories (CSL), founded in 1916. Initiated during World War I as a manufacturer of vaccines and antisera, CSL's biological product and national-interest missions have been key aspects of its operation.

The establishment of the Commonwealth Serum Laboratories Commission in 1961 as a statutory authority allowed for these dual tasks under legislation requiring that commercial operations be dominant and meet post-tax dividend criteria from trading profits and that noncommercial operations be undertaken only to the extent of allotted government budget funds. Today, with an annual total revenue approaching $60 million, the commercial operations are responsible for more than 90 percent of total activity.

The worldwide pharmaceutical industry is known for its research and development investment, but apart from several short-lived ventures into R&D, the Australian subsidiaries of the multinational companies have not demonstrated this facet of their strength in Australia. Indeed, for an industry dependent on world markets and a critical mass sufficient to support many key activities, small national activities cannot be expected to be a miniature version of the full spread of parent company functions.

Sterilization unit.

What, then, is the place for an Australian enterprise in this global industry, particularly when our nation has only 15 million people and is still largely remote from the rest of the developed world? The finding and exploiting of market niches built on the strengths available is the path to be pursued, but the path is often poorly marked, often only discernible because others have already trodden it, and it frequently leads to dead ends unless one can find the key to pass through to the main routes of world marketing and distribution.

Basic research relevant to biological medicine, however, is strong in the public sector of Australia, with a broad spread across the related fields of immunology, microbiology and molecular biology, through the diverse divisions of the Commonwealth Scientific and Industrial Research Organization (CSIRO), the state-funded universities and concentrated effort in a few institutions of world reputation.

At the applied level in biological medicine there has been a notable R&D effort at CSL, resulting in the large range of products they have made over the past 67 years. Many of the major advances of this century in therapeutics have been among the milestones of CSL's growth, for example, industrial-scale manufacture of insulin in 1923, diphtheria toxoid in 1927, snake antivenoms in 1929, tetanus toxoid in 1938, penicillin in 1944, human serum fractionation (Cohn) in 1952, Salk polio vaccine in 1956, subunit influenza vaccine in 1966 and interferon in 1981.

Biotechnology is therefore not new in Australia, but today there is widespread interest and optimism in the "new biotechnology" because it is seen that spectacular new advances and fortunes are the "glittering prizes" for those with the courage to enter the field.

Will the economies of biological product manufacture and marketing somehow be different in the next 20 years than they have been in the past 20 years? A belief that they will be is probably at variance with the attitude of the established industry, which over the past two decades has regarded biological product manufacture as the low-profitability, high-risk sector of the pharmaceutical industry. Indeed, it can still be said today that the few multinational pharmaceutical companies that have retained a strong presence in vaccines do not have the public visibility and media glamour enjoyed, if not promoted, by the venture-capital entrepreneurs as they proclaim their pursuits.

Downstream high-capacity protein purification.

Nonetheless, it is from the broadly based companies with deep financial pockets, manufacturing and marketing skills and extensive international distribution networks that the innovative pharmaceutical products from the new biotechnology are most likely to come during the remainder of this century. But as is the case with all commercial enterprise, there will be the exceptions as an occasional entrepreneur succeeds and grows to multinational viability and the industry loses established members who fail to adapt to changing technology and markets.

What is the situation in 1983 for investment in pharmaceutical biotechnology in Australia?

At the lowest level (but not implied to be unworthy) the argument I would strongly advance is that a presence in the field is of national importance.

This argument parallels that for support of basic or pure research in any field, which is an argument acceptable to governments (and therefore to the people) in most countries, as evidenced by the systems they develop for grants to train and employ researchers.

Through hands-on experience and a working presence the identification and exploration of prospects worthy of possible commercial exploitation can be best undertaken.

Some would argue that this task should be left to the market system, and so it would be if the market system worked freely or were allowed to so work. Most nations, however, do not allow the free-market forces unfettered play in pharmaceuticals (if in any fields), and there are particular justifications for small nations making individual decisions on the grounds of protecting health and employment and buffering against complete dependence on overseas technology and the complex logistics of international supply networks.

Thus in Australia, with its small population and largely remote situation in the Southern Hemisphere, decisions can be appropriately made to support particular areas of technology and the manufacturing industry. The next question then has to deal with the identification of areas of support. The present discussion confines itself to pharmaceuticals, and so the answer need only relate to aspects of that industry.

The spectrum of drugs or agents regarded as lifesaving is now extremely broad; a simple list such as could have been compiled immediately following World War II is no longer possible. Concentration on treating indigenous diseases of particular national importance could provide another approach, but such diseases amenable to control or calling for a local thrust more readily come to mind for third world countries than for developed ones.

Biological production control instrumentation.

Hence, as in the approach to any marketing problem, the focus comes to be on product differentiation, that is, in what way an advantage can be exploited or even developed if it does not already exist.

This is not the place to debate whether a society can develop such situations through deliberate science programs or whether they emerge serendipitously, and then perhaps as a function of the quality or quantity of research. Existing skills and resources together with emergent new disciplines provide the logical base for a national effort if there is to be one, even though it may only be a microcosm of the efforts under way and well advanced in the populous developed countries of the Northern Hemisphere.

In Australia the base for pharmaceuticals is narrow, and in the research disciplines fundamental to products of biotechnology and in development and manufacture, there are few organizations with the appropriate depth of experience and demonstrated ability. The question of the choice of therapeutic area for human and veterinary product development therefore goes back to the technology and scientific strength married to the development and exploitation capability demonstrated by established manufacturing and marketing skills.

The question of "what to do?" is thus tied to the next one of "how to do it?"

Although public funding to support unprogrammed research in the biological sciences must continue with some encouragement to "areas of need" as identified by the National Health and Medical Research Council over the past decade, there has emerged an allocation of public funding by government for biotechnology programs leading to product development. This orientation toward industrial application introduces the need for protection of intellectual property and technical know-how, which previously has been of little concern to academics. At the same time industry and academia have very limited experience in working together. The situation is more complex when public funds bring the private sector and academic research institutes together, but there is no alternative to such arrangements, particularly when basic research abilities are diffusely spread across the nation in places such as the universities and CSIRO divisions.

Typical fermentation plant.

We should note, however, that publicly funded developments in other countries have taken advantage of existing national organizations in which the government has equity, or that such developments occurred through the establishment of government presence in joint-venture arrangements with major private-sector institutions. Through CSL, with its strengths at critical levels of product developmental steps, there is a unique opportunity for both public- and private-sector research innovations and ideas to find the path to successful commercial exploitation.

I would argue that an already-existing R&D-based biological product manufacturer with a strong national marketing presence and an international linkage through vigorous export performance and licensing arrangements can be the prime focus of pharmaceutical biotechnology in Australia. This is particularly so when government policy is supportive of national effort in this area. The Commonwealth Serum Laboratories meets all the criteria to be that prime focus.

Faulding and Co. Ltd. of Adelaide in South Australia has managed it.

F. H. Faulding has developed a method of packaging the widely used antibiotic erythromycin so that the drug is released some time after the capsule containing it is ingested. The benefit of this is in the treatment of conditions in the low region of the intestinal tract. In the manufacturing process erythromycin is coated with a substance before being encapsulated.

What distinguishes F. H. Faulding from many other Australian companies that have successfully tackled the export market is that the company deliberately set its sights on the overseas market for its new product rather than the domestic Australian market. According to Mr. Brian Hough-Davies, the group general manager of Faulding's product divisions, the local market was becoming very crowded. "Most of the major international companies had set up subsidiary operations in Australia, so we thought the only place for us to go was overseas," he says.

The company searched hard for a niche in the pharmaceutical market and began working on the controlled-release process. It was only interested in applying the process to existing drugs because it could not afford the time or the resources required to put any novel preparation through extensive testing procedures.

"Our target drugs have to be existing drugs on which all the toxicity tests have been carried out," Mr. Hough-Davies says. "The next requirement is that there must be something difficult about taking the drug—for example, some drugs leave you feelng sick because they are released in the stomach."

F. H. Faulding has signed a distributor in the U.S. and is hopeful this company will take up an option to manufacture erythromycin products using the Australian company's technology. F. H. Faulding & Co. is continuing to spend about 10 percent of its erythromycin-related revenues of $6 million a year on research and development.

Australasian Training
Aids Pty. Ltd.

About 600 kilometers south of Sydney on the banks of the river Murray is the bustling country town of Albury, New South Wales. Albury is the home of one of Australia's most unusual companies, Australasian Training Aids Pty. Ltd. (ATA). It was formed in 1957 by its current managing director, Mr. Lyndsay Knight, and has established itself at the forefront of global technology in the complex field of military target training systems.

In the past five years, ATA has completed major contracts worth more than $60 million. It has done this in conjunction with wholly owned subsidiaries in the U.S. and Britain. It is currently negotiating new sales with a total value of several hundred million dollars for completion within the next five years.

Recent developments by the 200-strong work force include a gunnery training system, which provides immediate down-range feedback on the fall of shots. It displays this information on visual display terminals at the firing point and the range control tower.

In the U.S. ATA is currently negotiating for the installation of a semiautomatic infantry range system at Fort Jackson, South Carolina, and a computer-controlled infantry range system at Fort Polk, Louisiana.

The 64-meter Parkes radio telescope, New South Wales, was constructed by CSIRO in 1961.

Dunlop Olympic Group

Australia has more than its fair share of rough roads. Technology used to build tires for use on the generally smoother roads of North America and Europe is not necessarily applicable to Australian conditions. Olympic Tyre and Rubber Company, now part of the Dunlop Olympic Group, had adapted the steel-belted radial-ply tire for use on automobiles and took this one step further by developing a tubeless version for use on trucks and buses.

Advantages of the steel-belted radial tire over cross-ply and fabric-belted tires include better handling characteristics, improved fuel consumption and lower wear rates. Olympic put its newly developed tire through a rigorous testing program. This culminated in the company's entering a five-ton truck in an international rally. The truck was not among the winners of the London-to-Sydney Rally in 1977, but it was the only vehicle to finish on the same set of tires it had at the start.

Rheem Australia Ltd.

While Rheem Australia admits it is not the major supplier of solar collector panels in Australia—the honor goes to a Western Australian company—it has built up considerable expertise in associated technologies, including water heaters and steel containers.

The publicly owned company is investing $4 million in a new production line at its Sydney plant. This will have the capacity to produce 30,000 units annually of a solar collector panel developed in Australia. The new panel does not use copper but a special so-called nickel-black coating on steel or aluminum collector panels—both metals are cheaper than copper.

Rheem is Australia's largest manufacturer of electric and gas hot water systems. It first entered the solar water heater field in 1980 after a major research and development program. The company's solar-powered hot water systems are now used in commercial and domestic installations throughout Australia.

Wormald International Ltd.

Fire protection products play a crucial role in the protection of key sections of the Kennedy Space Center at Cape Canaveral, Florida. A computerized monitoring and surveillance center in the British city of Birmingham acts as a silent guard for a variety of industrial and commercial premises. In Brazil, the leading manufacturer of fire hoses is recognized for the high-quality agricultural fire hose it makes. The company behind these diverse pro-

(continued on page A30)

1985 Satellite Launch to Create New Communications Era

In 1985, in the four fiery minutes that it takes for a NASA space shuttle to blast into space, Australia will join the growing number of nations that use space-age communications systems. During that year, two communications satellites will be launched to provide the basis for a state of the art satellite network of world standard. A separate company, AUSSAT Pty. Ltd., has been formed by the Australian government to operate and control the system and also has the responsibility of marketing it and interfacing it with the existing terrestrial telecommunications system. The Australian government is the owner of all the issued capital in AUSSAT.

This planned domestic satellite system uses existing proven technology in an innovative way and brings to the country new skills in the operation of the satellite system. From this point of view, the satellite project is, and will continue to be, a positive force for the transfer of technology to Australia. Many highly experienced Australians have already been attracted back to Australia by the opportunity to apply their unique skills to this important national project.

The potential exists for all Australians to benefit from this exciting technological development, in both commercial and private activities. It is also likely to have long-term beneficial effects on the pattern of life in Australia. At the moment, the country is sometimes a prisoner of its own size, as major cities are spaced 1,000 kilometers or more apart and physical communications can be expensive and time consuming. Australia's existing terrestrial telecommunications network is acknowledged to be of world standard but its operator, Telecom Australia, has been unable to provide its full range of services throughout the nation because of the distance-related nature of its costs.

The upgrading of the national telecommunications network by the addition of satellite-based technology will open up the possibility of placing major service organizations outside the cities, since they will have access to all modern forms of telecommunications required for such businesses. All Australians will be able to view significant national events and to enjoy new means of relaxation and entertainment.

In addition to easing pressure on the cities, the satellite system can give all Australians access to common services in a wide variety of areas. For example, there must frequently be a core population level before a new service can be introduced into a community. The use of satellite technology will ensure that if the service can be offered via telecommunications, then it can be delivered anywhere in Australia.

The satellite is designed specifically for Australia's needs; it has a complex antenna system to shape a national beam and four regional beams that each cover about a quarter of the continent. The design will permit the introduction of a quasi-direct broadcast service, so that residents of remote areas can receive television and radio programs directly from the satellite. Another small beam will cover Papua New Guinea to provide capacity for domestic telecommunications and television relay purposes.

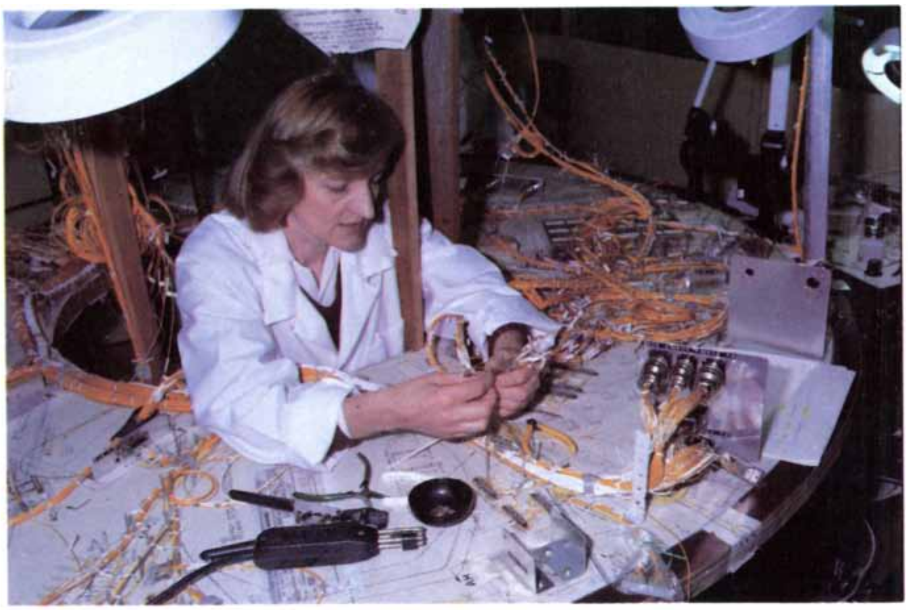
Major earth stations currently under construction by AUSSAT in the eight capital cities of Australia will provide gateway access for users of the system.

An achieved objective of the Australian satellite program has been to provide a substantial degree of Australian content. The program has also led to the creation of high-technology employment opportunities in areas such as satellite system design and operation and the design and construction of earth stations. The widespread use of small earth dishes in remote areas of Australia is likely to open up export market opportunities for Australian companies that have perfected the technology involved in manufacturing small earth stations.

As part of the technology transfer program, Hughes Communications International has given contracts to Australian companies for some elements of the satellite hardware and for the development and programming of the satellite control and monitoring systems. The system design and overall technical specifications for the satellites were prepared by Australian engineers with support from overseas consultants. An engineering and technical group from AUSSAT is currently working with the satellite manufacturer in the U.S. to ensure that there will be a maximum transfer of satellite technology to Australia and that AUSSAT is well-equipped to operate and control the satellites after launch.

Because the transfer to Australia of this advanced technology started early in the Australian communications satellite program, less reliance on overseas engineering assistance will be necessary during the design, tendering and construction phases of later generations of satellites. AUSSAT estimates that current and future contracts will create more than 1,000 employment positions in Australia.

An Australian technician assembling components for AUSSAT's satellites. Sydney, NSW.



A Premier State

New South Wales calls itself the Premier State with considerable justification. Its capital, Sydney, is one of the world's great cities and is certainly the best-known Australian city. It is the site where the cornerstone of European settlement in Australia was laid in 1788.

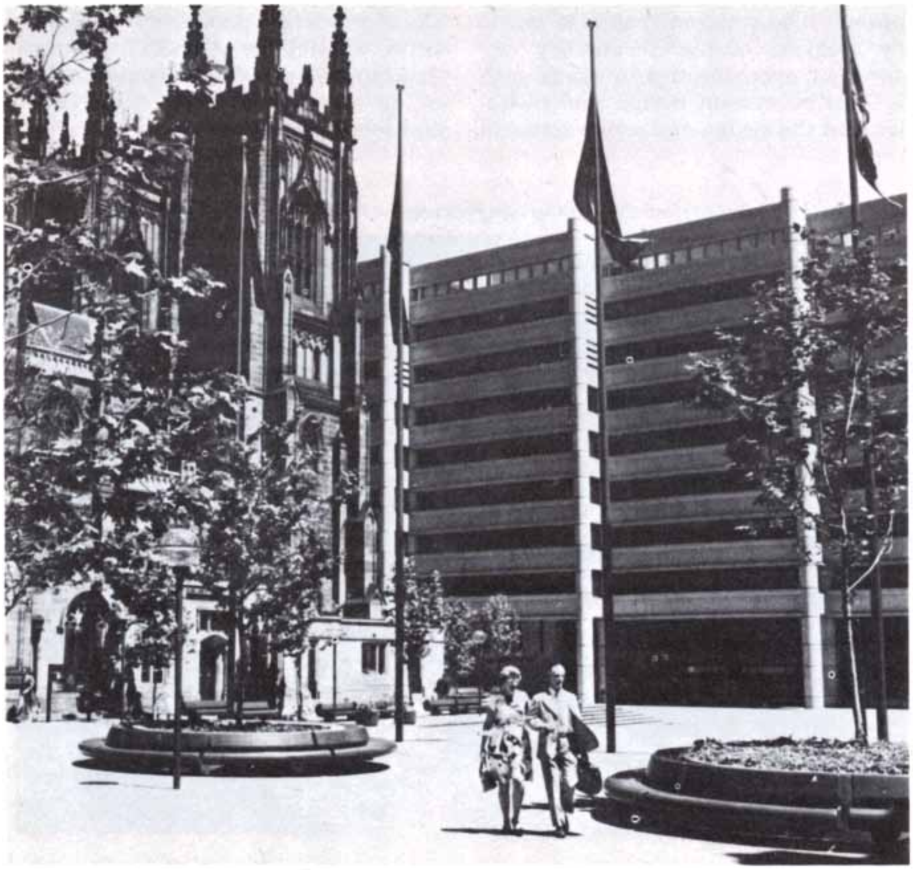
Today Sydney has become the hub of knowledge—of intensive industries such as finance, business services, medical technology, communications, electronics, computer software and biotechnology. Sydney has also attracted industries requiring external consulting, contract services, the facilities of universities, technical or private research or advisory organizations, all of which are readily available.

Sydney is the major center in Australia for the computing and information technology industry, with some 350 companies directly involved in this area. Of these 26 are Australian subsidiaries of major international computer hardware, software and service companies.

The Sydney suburbs of North Ryde and nearby West Lane Cove have in the past 10 years attracted an amazing variety of advanced technology companies including those in information processing, computing, pharmaceuticals, electronics, instrumentation and associated services. The North Ryde advanced-technology zone also houses the major complex of Australia's largest indigenous electronics company, AWA (Amalgamated Wireless Australasia Ltd.).

The Nucleus Ltd. group based in Lane Cove is an example of a successful advanced-technology company operating in Sydney. The group is at the forefront of biomedical technology with developments in ultrasonic body scanning and telemetrically reprogrammable cardiac pacemakers, not to mention the bionic ear for the profoundly deaf.

The international life style of Sydney has been greatly enhanced by the introduction of many different cultures brought here by new settlers from all over the world. In its stunning Opera House, the city also boasts a world-famous cultural center. Housing costs in the Sydney region are comparable with those in other major cities in the world. The cost of office and manufacturing space in Sydney, however, is probably lower than in any other major Western-style city in the world.



Sydney, one of the world's great cities, offers a contrast of the old and the new. In the photo (above) a new pedestrian plaza is juxtapositioned with St. Andrews Cathedral.

The head office of Australia's international airline, Qantas, is in Sydney, as are the regional offices of other major international airlines, making the city's airport the most significant in Australia in terms of overseas services.

Sydney's historical significance and its magnificent setting on the shores of a very large deep-water harbor have meant the city has long been the natural gateway for Australian imports and exports.

Although geographically large, New South Wales is a highly urbanized society with a population of more than five million. The state has a diverse economy. It produces, by value, about 30 percent of Australia's agricultural output and about 40 percent of its manufactured goods.

The major industrial port cities of Newcastle and Wollongong flank Sydney to the north and south. These three cities account for 80 percent of the population of New South Wales, and the majority of Australia's advanced-technology industries are found in them. As regards intellectual raw material, the Sydney-Newcastle-Wollongong area has five universities together with 15 colleges of advanced education and 80 technical colleges.

With a solidly based manufacturing infrastructure and the largest local market in Australia, New South Wales is encouraging the introduction of advanced technology into its existing industrial base as well as fostering new advanced-technology industries.

The state government has taken a number of initiatives to foster the development of industries based on advanced technology. One such initiative was the establishment of the Advanced Technology Centre within the Department of Industrial Development and Decentralisation. The center is a focal point in New South Wales through which government assistance to promote the increased use of advanced technology in industry can be channeled, and it is a place where people can go for both assistance and advice, financial and otherwise, about advanced technology: its introduction, development, commercialization and marketing.

With the technical and business infrastructure in place and positive support from its government, New South Wales is perfectly positioned to provide the potential for advanced-technology industries to capitalize on the vast and rapidly developing markets of the Pacific Basin.

jects is one of Australia's few multinational companies.

Wormald International Ltd. has its head office in Sydney and operates subsidiaries in more than 20 other countries. It is one of the few leading special-risk fire protection companies in the world. It has diversified into other forms of security and now claims to be uniquely placed to provide integrated fire and security installations in large commercial, industrial and governmental complexes. This integration involves extensive use of advanced computerized systems.

Wormald has shown that it is prepared to diversify in the quest for new products. Earlier this year it announced a joint-venture development company with the federal-government-backed Australian Industry Development Corporation. The new company, Australian Biomedical Corporation Ltd., will develop and market Australian biomedical inventions for world markets. Although Wormald's senior management admits the new company is at a very early stage of development, they are confident it has considerable potential. Australian Biomedical Corp. has already attracted a number of inventors and institutes, which have agreed in principle to offer marketing rights covering about 30 products to the corporation. The initial range of products includes electronic patient-care equipment, medical and surgical equipment and home-health-care products.

Biotechnology Australia Pty. Ltd.

Biotechnology Australia Pty. Ltd. is one of Australia's largest biotechnology companies and among the few with the capacity to manufacture products on a large-scale commercial basis. It was first set up in 1978 by a Sydney researcher, Dr. David MacLennan, with the aim of establishing an Australian ethanol based on genetic engineering techniques. But funds were short, and the company was taken over by CRA Ltd. in 1981, which now holds 70 percent equity in the company. Biotechnology Australia now has a staff of 65, including 48 graduates housed in a modern laboratory complex in the Sydney suburb of Roseville. The company is concentrating on mining and agricultural products, in particular advanced animal vaccines, animal growth hormones, novel methods of biologically controlling pests and ore leaching.

In the longer term it also plans to work on human health products, particularly vaccines, such as the malaria vaccine. Two products—vaccines for scours in pigs and fleece rot in sheep—are nearing commercial production. The company hopes to have a number of products on the market within the next five years and currently has a portfolio of 12 different projects under way. It does not expect to make a profit for six to eight years.

Australian Monoclonal Development

Australian Monoclonal Development was established in March, 1982, in Artarmon, Sydney, by Dr. Alessandra Pucci to produce monoclonal antibodies for veterinary and pathology diagnostic laboratories, research groups and pharmaceutical companies. It has a staff of 25 and has 35 monoclonal antibodies on the market in three different categories: research use, immunological and endocrinological diagnostic kits and veterinary diagnostic kits for diseases such as brucellosis as well as for meat identification.

The company has a unique structure, with six consultants from leading research institutes and universities around Australia. The commercial monoclonals are mostly the result of basic research by these consultants and other Australian research groups.

Nucleus Ltd.

Nucleus Ltd. based in Sydney, is Australia's largest biomedical equipment company, with a turnover of $70 million a year and profits of $166,000 in the half-year of June, 1983. It was set up in 1966 and since then has grown steadily, now employing more than 500 people in Australia and 250 overseas. Nucleus Ltd. is divided into four main operating companies: Ausonics Pty. Ltd., which develops, manufactures and markets ultrasound diagnostic imaging equipment; Domedicia, which manufactures hemodialysis equipment and imports and distributes renal dialysis equipment; Medtel, which manufactures cardiac monitors, and Telectronics, which designs and manufactures electronic pacemakers and bone growth stimulators. The company went public in 1980, raising $5.6 million in capital. It exports 80 percent of its products overseas and is regarded as Australia's most successful high-technology company. It spends approximately 8 percent of sales on R & D Sales topped $37 million in the six months to June, 1983.

D. D. Webster Electronics

D. D. Webster Electronics of Melbourne, which was established in 1970, is now one of Australia's leading computer manufacturers. In 1977 the company launched its minicomputer

Department of Science, Antarctic Division

An Antarctic research team preparing its helicopter for an ice-edge flight.

range: Spectrum Eleven, which is compatible with a wide range of interfaces and applications packages covering accounting, financing, manufacturing, engineering, word processing, electronic spread sheets and educational applications. Memory ranges from 32 kb to 550 kb and the company is competing well against overseas competitors.

Quentron Pty. Ltd.

Quentron Pty. Ltd. is Australia's largest laser company. Based in Adelaide, it designs and manufactures a range of lasers, precision optics and optical systems for research, industry and military applications. There are three companies in the group: Quentron Pty. Ltd., Quentron Optics Pty. Ltd. and Quentron Electronics Pty. Ltd., and the shares are held by the Dobrinine family and staff. Quentron was set up in 1969 and has been expanding rapidly since. In the past few years expansion has been at the rate of 30 percent a year, with sales reaching $3.1 million in 1981–82. Imports of lasers make up two-thirds of the sales; the remainder are designed and manufactured by the company in Adelaide.

Since 1974 Quentron has plowed most of its profits back into research and the development of new laser products. One such product is a high-energy-pulse metal-vapor laser for cancer treatment. Others include a laser patient alignment system for radiotherapy, a laser-based mortar simulator for training soldiers and a laser lecture pointer, a unique multiple-optical-fiber laser-light delivery system for cancer treatment and high-energy-pulse lasers for research.

Digital Electronics Pty. Ltd.

Digital Electronics Pty. Ltd. of Sydney, founded in 1970, markets a range of plug-in boards, electronic laboratory trainers and mini- and micro-business computers and has developed special devices for weather and wool measuring. It has a staff of 65 and a turnover of $6.5 million a year.

Solahart

Solahart of Perth in Western Australia is Australia's largest manufacturer of solar water heaters. It has dominated the market for many years, having about 80 percent of the domestic market and over 90 percent of the export market.

Much of Solahart's success can be attributed to the entrepreneurial drive of its managing director, Wayne Reed, who masterminded a takeover of the company by its employees in 1973. The reorganized company was so successful that in 1979 it attracted the attention of Shell, which bought 50 percent of the shares but has no direct involvement in the running of the company.

Solahart is very reluctant to give any information on sales and profits, but the value of its sales are estimated to be about $20 million.

Machine Dynamics Pty. Ltd.

Machine Dynamics Pty. Ltd. of Melbourne is Australia's only robot manufacturer and has developed a number of pick-and-place and point-to-point robotic devices. The company was established in 1968; it has grown slowly and now employs 17 people.

Machinery Dynamics has sold over 50 of its first-generation robots, which include machines for spot-welding, loading parts onto a conveyor and unloading machine tools. It is estimated that only 200 robots are as yet in use in Australian industry, and the company faces stiff competition from overseas products. Managing director, Len Whelan believes demand for robots in Australia will grow slowly and will depend on the creation of a demand for flexible automation systems based on small-batch production. He also feels that overseas imports of robots should be limited by quotas if a robot industry is to succeed in Australia.

Varian Techtron

Varian Techtron of Melbourne is the world's second-largest manufacturer of atomic absorption spectrophotometers and a major manufacturer of UV-visible spectrophotometers. It began in the 1950's as Varian Associates but was taken over by the U.S. company Techtron Appliances in 1967. It now employs over 300 people, and sales exceeded $25 million last year.

AUSSAT

Australia will be one of the last major nations to establish its own domestic satellite communications system. The satellite itself is due for launching by the space shuttle in 1985. By then a network of earth stations will have been established together with telemetry and tracking stations in Sydney and Perth. The system will be operated by a partially government-owned company called AUSSAT.

The move to a domestic communications satellite system will certainly maintain within Australia a growing body of talent in space technology. Future demands for meteorological surveillance and resources development and defense-related issues are likely to be met through the use of available commercial systems or specially constructed systems suited to Australia's largest size and extensive resources.

Advanced Satellite and Computer-Assisted Exploration Technology

by Murray Longman

Identification of the linear features known as lineaments is an extremely subjective science, but one that has been practiced for decades in many scientific disciplines, including astronomy and earth science. Lineament features were first recognized using relatively crude telescopes and the lineament map known as the Martian canal system was produced.

The same principles have been applied to the study of the earth, based on pattern recognition of certain features such as the parallellism of continental margins, the broad rift-valley structure of South Africa and the distribution of submarine ridges. With the development of aerial photography, these linear features became visible in many patterns, and the introduction of lineament analysis as a science in its own right commenced. In addition to the great variety of smaller features, lineaments of continental size were also visible on photomosaics and topographical maps produced from aerial photography.

A new dimension of lineament analysis was introduced when space photographs from the Apollo missions provided an opportunity to study the earth from space. With the launching of the first Landsat satellite in 1972, analysis of lineaments on both a continental and a local scale could be undertaken on a routine basis.

Detailed geological mapping in mining areas throughout the world has revealed a strong correlation between mineralization and structure. With the introduction of this new method of analyzing data many researchers, particularly O'Driscoll in Australia, quickly showed that major mineral deposits were at the intersection of regional lineament zones. O'Driscoll's work in Australia located Broken Hill, Kalgoorlie and Kazbalda, which are major base-metal-, gold- and nickel-producing areas at the intersection of regional lineament zones.

During 1976 the senior partner of M. J. Longman and Associates discovered that geological and structural maps could be produced directly from computer-compatible tapes supplied by NASA rather than by visual interpretation of Landsat images. After considerable research, computer programs were developed to produce geological maps. Further research led to the development of software that could produce lineament maps comparable to those produced by conventional visual analysis of Landsat images.

However, to produce maps comparable in quality to visually interpreted maps, approximately 10 percent of the data available on the Landsat data tape is used. These computer programs can utilize a much broader data base than was previously available.

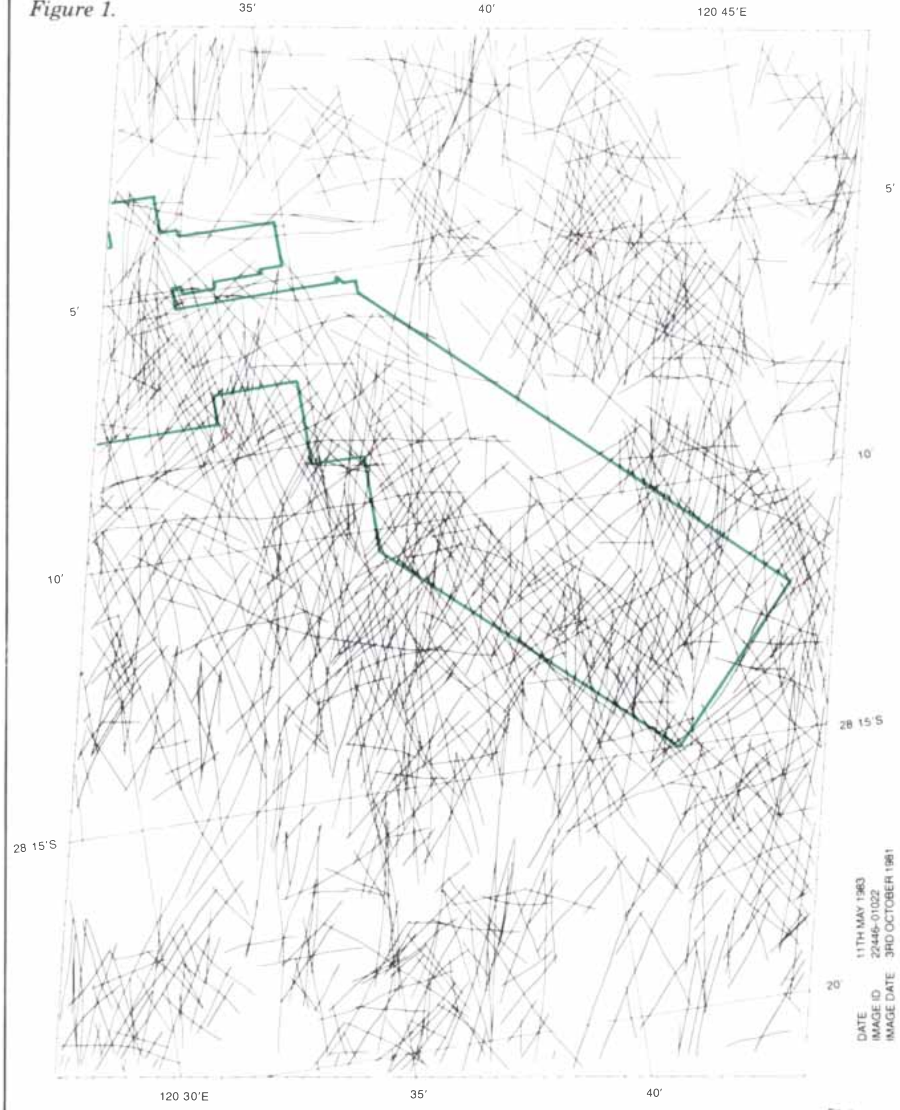
The principle advantages of computer processing other than the capacity to utilize the data core fully is that it removes from interpretations the individual bias present in all visual ones. Because the data is now in digital form, statistical analysis of complex data transformations is possible.

Initial work with the programs over known mining areas revealed that with the increased sensitivity made available by lineament analysis, individual mineral deposits could be outlined and the characteristics of the lineaments associated with each deposit determined. With this approach, a portfolio of lineament characteristics associated with various mineralization types has been assembled and these characteristics applied to various exploration targets.

For exploration this approach is an excellent one, because the Landsat data provides regional coverage of any continent and the lineament analysis locates areas that have a greater potential for mineralization. These areas are then studied with conventional geological means to provide an additional control to the system and to assess the distribution and grade of the mineralization.

The effectiveness of the lineament exploration system is illustrated by a case study of an area near Lawlers in

Figure 1.



LANDSAT–COMPUTER INTERPRETED LINEAMENTS
GREAT EASTERN EXTENDED, LAWLERS DISTRICT. WESTERN AUSTRALIA

M.J. Longman

the East Murchison Goldfield of Western Australia. The case study is of an area being explored by a junior Australian exploration company, Forsayth Oil & Gas NL. A regional analysis of this goldfield using specially enhanced Landsat imagery, under the supervision of M. J. Longman and Associates, indicated from visual analysis that within the Lawlers area, approximately 320 kilometers north of Kalgoorlie in Western Australia, an anomalous situation appeared to be present.

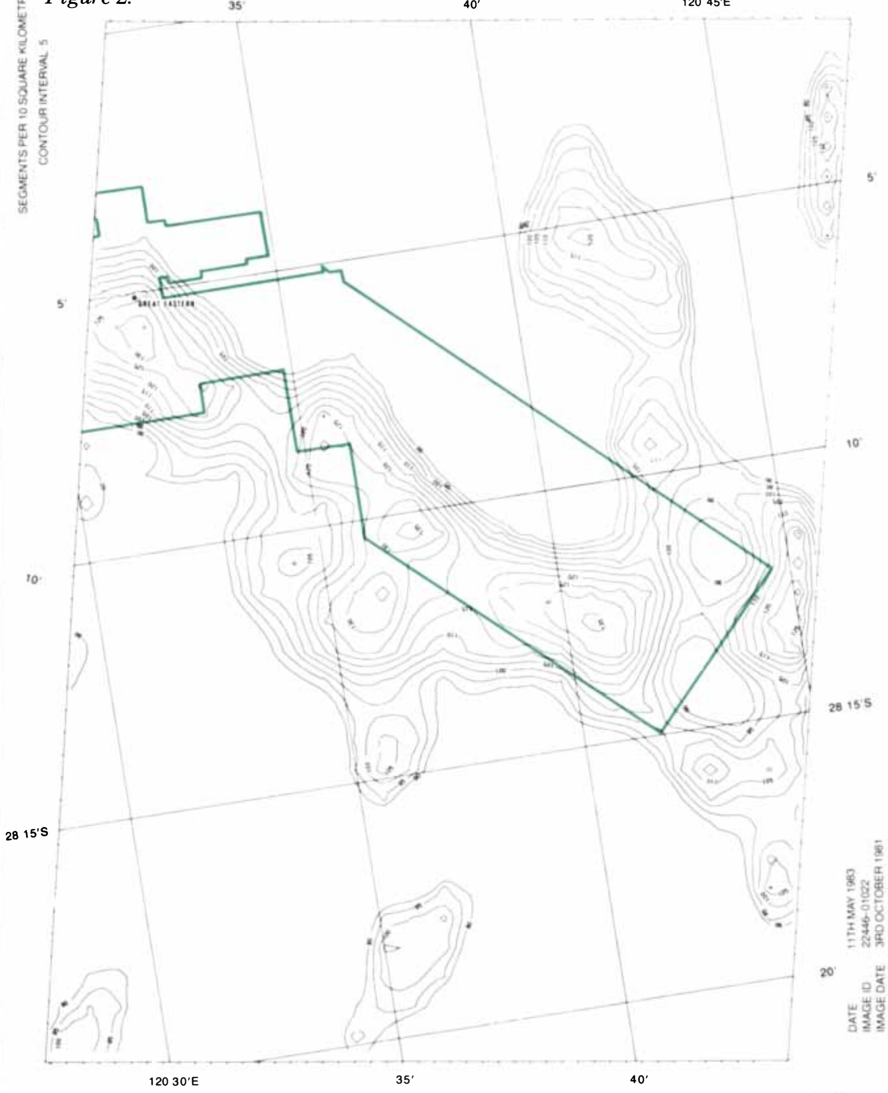
A lineament analysis of this area, using the computerized techniques developed by Murray Longman, indicated that within a 2,000-square-kilometer region three areas totaling approximately 30 square kilometers were favorable for the accumulation of structurally controlled mineralization.

The Landsat image produced by satellite from approximately 600 miles above the earth's surface shows well-defined lineaments and faults that are visible to the unaided human eye.

Figure 1 shows the lineaments within the Lawlers region as determined by computer analysis of the data recorded on the NASA Landsat tape. Each lineament represents a fracture of the earth, and mineralization may occur in highly fractured areas. A similar configuration might be obtained using visual analysis of the photographic image. This method, however, requires trained interpreters working many hours, and the results are highly subjective. The computer analysis has obvious advantages: unbiased interpretation, timely analysis and processing of lineaments that cannot be detected by the unaided human eye.

Figure 2 illustrates a further extension of the computer process in that the lineaments shown in Figure 1 have been contoured according to fracture density for greater ease of interpretation. The target areas are clearly visible and occur where the fracture density is highest.

Figure 2.



LANDSAT–COMPUTER INTERPRETED LINEAMENTS
GREAT EASTERN EXTENDED, LAWLERS DISTRICT. WESTERN AUSTRALIA

M.J. Longman

When the target areas had been located, Western Australian Mines Department records were checked against the target areas to determine whether the prospective ground was available for pegging. After it was established that tenements were available, exploration and prospecting license applications were made and ground geological investigations commenced.

In the southern area three extensive zones, up to three kilometers in length, were mapped with gossans, iron-rich chemical deposits indicating base-metal and gold mineralization at depth. Initial shallow drilling has intersected a zone up to nine meters wide of base-metal mineralization, and exploration is continuing with an extensive drilling program. In the central area a stockwork of intersecting quartz and granite porphyry dikes has been located. Investigations are to be continued at a later stage.

Adjacent to the northern anomaly is the Great Eastern Mine. This mine, which was operated by Bewick Moreing with Mr. Herbert Hoover in charge, closed in 1903 after producing 4,151 kilograms of gold from 222,000 tons of ore, an average grade of 18.7 grams per ton. Hoover, a mining engineer, went on to become the 31st President of the United States. Due to the significance of the lineament anomaly, the property was acquired by Forsayth Oil & Gas NL and the available drill-hole data reassessed to determine the true potential of the area. Preliminary calculations indicate an in-ground resource to a depth of 30 meters of 340,000 metric tons at a grade of 4.00 grams of gold per ton. The potential for at least one million tons of similar grade ore exists in the area.

The above example illustrates that by basing an exploration program on the lineament analysis system, rapid and accurate assessment of exploration targets can be undertaken. In this case, the elapsed time from concept to discovery was five months, resulting in considerable cost savings to Forsayth.

A worldwide analysis of major mineral deposits has indicated that the Landsat lineament characteristics of these deposits can be readily developed and the method can be used to assist conventional geological and geophysical exploration methods.

The Future: Traditional expertise offers promise as declining productivity urges diversification

Since coming to power last March the Labor government has adopted a strong new policy boosting industrial innovation and development. The intention is to stimulate the development of new high-technology industries and to encourage the introduction of new technology to improve traditional industries. Such measures are seen by the government as a major way of boosting economic growth and lifting the productivity of Australia's declining manufacturing sector.

As the Prime Minister, Mr. Bob Hawke, said at the opening of Australia's first national technology conference in September: "The effectiveness with which we use new technology will determine whether we can reverse the long downward slide in our living standards relative to those of other countries—from one of the highest per capita incomes in the world 30 years ago to one of the lowest of the industrial countries."

This rapid decline in productivity, employment, exports and economic growth is one of the major concerns of the government. Between 1962 and 1980 only four out of 24 OECD countries had a GDP growth rate per capita lower than Australia, and Australia's ranking in terms of GDP per capita fell from ninth to twelfth.

The decline is particularly apparent in the manufacturing sector, where Australia's percentage share of GDP has slipped from 28.8 percent in 1963 to 21 percent in 1979 and is still falling. This is also reflected in the very low ratio of exports to imports relative to other OECD countries and the particularly low level of exports of technology-intensive products.

In 1980 Australia exported US$81 per capita in high-technology goods compared with US$2,584 for Switzerland, US$1,378 for the Netherlands (countries with smaller populations than that of Australia) and an average US$468 for the OECD. This is again reflected in the extremely low level of research and development carried out by private industry, which when measured as a percentage of GDP, is one of the lowest among OECD countries.

The new government is determined to reverse these trends and despite a tight budget for 1983–84 has increased spending on industrial research and development by 36 percent. It has also more than tripled funding for technology programs and introduced a new tax incentive scheme to boost the venture-capital market in Australia.

Funding for the Industrial Research and Development Incentives Scheme (IR&DI) has been increased to $71.6 million, compared with $52.8 million last year. The scheme has been expanded to encompasses computer software research and includes $2.5 million specially earmarked for biotechnology.

The increased funding for technology and innovation programs will make possible a number of new initiatives, including analysis of market opportunities for specific new technologies, the establishment of a number of new technology demonstration projects in new and traditional industry and the establishment of a coordination unit to improve information flow to industry.

The government has identified a number of particular high-technology areas for specific support, including biotechnology, energy, information, materials, communications and biomedicine. Within these areas the government is concentrating aid on technologies that can achieve productivity increases across a wide range of industries, such as computer-aided manufacturing; that have good export potential; that are fundamental to the next generation of technological change, such as microelectronics and biotechnology, and that are based on particular national research strengths.

However, many see these measures as inadequate without the provision of a number of other support systems and in particular the provision of risk capital to get innovations from the research and development stage through to commercialization.

Agricultural Research

Australia is now the world's largest exporter of wool and meat and the third largest exporter of wheat—all against the seemingly insuperable problems of a harsh climate, poor soils, extensive droughts and vast distances from export markets.

Pasture improvement provides the most striking example of how research by CSIRO, state departments and universities has transformed a country with the most limited natural resources into one of the world's major agricultural nations.

Introduction of legumes, subterranean clover and lucerne in the temperate regions and stylos, lucerne and sirato in the tropical regions has boosted soil fertility and enabled highly productive grass species to grow in combination with the legumes. This has made possible a fivefold increase in stocking rates, up to 100 percent increases in meat production as well as the sowing to pasture of an extra 20 million hectares over the past two decades.

Australian researchers now have a world reputation for their expertise in nitrogen fixation and *Rhizobium* bacteria and have built up a collection of about 4,000 grasses and 12,000 legumes and associated bacteria to help in research programs. New pasture species and management systems and crop improvement have contributed significantly to the production increases.

Improved livestock breeding has also played a major role in boosting production. A number of new breeds, such as the Belmont Red and the Australian Milking Zebu, have been produced for the tropical regions. *Bos indicus* blood has been gradually infused into breeding programs to produce crosses that combine the meat production qualities of the British breeds with the ability to tolerate heat, resist parasites and survive on low-quality feed of the zebu and brahman.

Selective breeding programs to improve the productivity of sheep, cattle and pigs under Australian conditions have also proved very effective, along with world-class work on twinning in sheep and improved wool production.

Australian researchers have built up a world reputation in ruminant nutrition, digestive physiology and wool biology, and work is under way on new chemical shearing methods for sheep. A unique robot sheep-shearing system is under development with support from the Australian Wool Corporation. It involves the production of an advanced robot system with a type of artificial intelligence that enables it to learn and remember from one sheep to the next and to adjust the cutters continually to keep them moving just above the skin.

In the past three years of research

(continued on page A38)

Agricultural Research for the Developing World

Australia is one of the few developed nations that has agricultural environments and problems that relate closely to those found in many third world countries. This comparative advantage in assisting developing countries with their agricultural research problems has been exploited by the Australian government in the formation of the Australian Centre for International Agricultural Research (ACIAR).

The new center is a statutory authority, formed in June, 1982, in response to the need for an independent professional scientific body to initiate a new approach to the management of Australian aid in the broad field of agricultural research. This approach involves the establishment of a partnership between Australian research scientists and those from developing countries to identify and help solve agricultural problems.

Since its establishment, the center has developed a number of collaborative research projects with developing countries. The projects are based on priorities established after intensive consultations, including advice from the center's international-policy advisory council, half of whose membership is from developing countries. These priorities represent problems that are common to several countries and with which Australia has special research competence.

ACIAR is developing its projects for a number of program areas, including land use, soil and water management, plant improvement, nutrition and preventative health, animal and fish production, animal health, postharvest technology, farming systems, forestry, socioeconomics and communications.

The center acts essentially as an entrepreneur, investigating joint opportunities, planning and developing research programs, organizing and contracting for technical resources and handling funding. Once the agreed-to projects have been defined, they are contracted to Australian research establishments, which manage the research activities, usually undertaken both in the developing countries and in Australia.

Much of the effort is being focused on Australia's nearest neighbors in Southeast Asia, particularly the ASEAN and South Pacific countries and Papua New Guinea. The center is also considering joint projects in Africa, South Asia and China.

ACIAR has a budget of $25 million for its first three years of operation. Among the early projects receiving ACIAR support are research on improvement of the pigeon pea, a drought-tolerant legume widely grown in India, parts of Africa and Southeast Asia; an identification system for viral diseases of legumes, and long-term storage of grain in the Tropics. Provision of fuel wood for the billion or more people in developing countries who depend entirely on wood for cooking and heating is another urgent problem. Australian fast-growing hardwoods such as eucalyptus, acacia and casuarina are ideal for this purpose, and a project has been developed to provide seed and silvicultural technology to help identify suitable species for collaborative projects in Africa and Southeast Asia.

Projects in the farming systems program will be especially concerned with identifying constraints facing resource-poor farmers in order to assist in designing technology options and agricultural policies that can alleviate them. One such project already under way examines the constraints to adoption of new rice-based technologies in the less reliably irrigated regions of Sri Lanka and the Philippines. Other initiatives could include investigations into the use of the vast untapped estuarine resources of Southeast Asia, which cover an estimated six milion hectares; the use of tropical crop and forage legumes in developing lea farming systems for the Tropics, and research into increased ruminant productivity.

Other projects in the fields of animal and plant nutrition, communications and socioeconomics are also in progress. In addition, ACIAR will endeavor to sponsor research on a range of commodities and to ensure that the special needs of resource-poor farmers are catered to in the composition of its research portfolio.

A major justification for ACIAR is the high payoff of agricultural research in developing countries. The benefits include increased agricultural production at lower costs, generation of attractive returns on investments and enhanced equity. Annual rates of return of up to 60 percent are common and indicate that further investments are needed to exploit opportunities for increased production.

The outputs of ACIAR's programs will be new knowledge, new technology and increased capacity of the developing countries to undertake research, as well as the benefits from direct interaction between scientists from Australia and the developing countries. By utilizing these scientists' special skills, ACIAR presents Australia with a unique opportunity to contribute to the enhancement of agricultural production in the developing world.

ACIAR specializes in agricultural development geared to the developing countries

An Advanced Electronics Pace-Setter for 70 Years

Government recognition of "sunrise" industries is certainly endorsed by Australia's largest indigenous electronics company, Amalgamated Wireless (Australasia) Ltd. (AWA). AWA believes it has been a sunrise organization since 1913, when it was formed.

AWA designs, manufactures and maintains an extensive range of electronic systems including fully integrated telecommunications equipment and components, air navigational aids, data processing systems and advanced defense-related hardware. The company operates extensive services for the maintenance of radio and navigational equipment, and it is well known as a supplier of high-quality domestic electronic equipment. AWA is also the operator of a commercial television station in Brisbane, the capital of the state of Queensland, and of seven radio stations in eastern Australia.

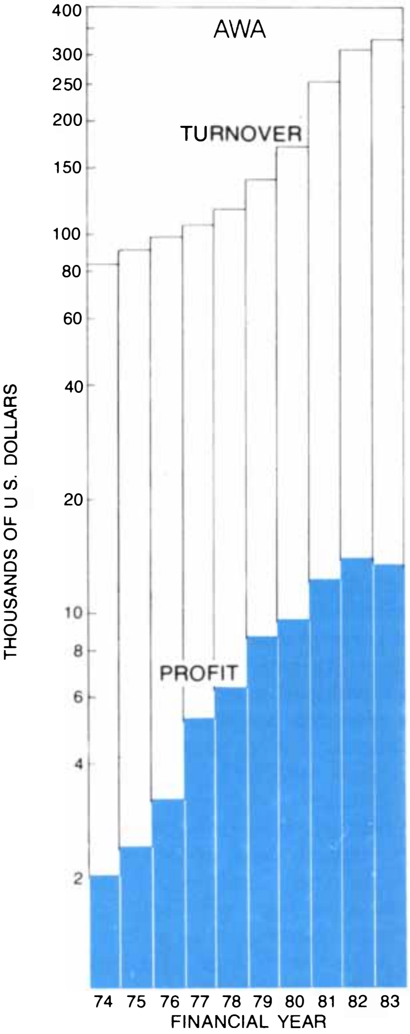
The successful marketing of its widespread portfolio of electronics-based systems and services is backed by extensive manufacturing facilities largely centered in its major divisions at Ashfield, North Ryde, Rydalmere and Leichhardt in Sydney. The total area used by the company covers 140,000 square meters, of which factory areas occupy 110,000 square meters. Its current activity in high technology is significant and diverse, from microelectronics to satellite monitoring systems. The list includes HF and microwave radio, over-the-horizon radar, computer systems, sonobuoys, data terminals and even betting systems.

AWA is a publicly owned company employing more than 5,000 people, including 300 professional engineers and scientists. Over 85 percent of AWA's stockholders are Australian, and with annual sales approaching US$350 million and assets in excess of US$200 million, the company can point to a decade of steady growth.

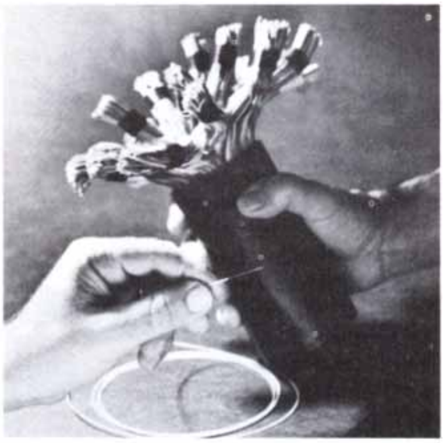
Its major customers include Australia's national telecommunications operator, Telecom Australia; Australian Defence and other federal and state government departments; major airlines, and commercial, mining, oil and shipping companies.

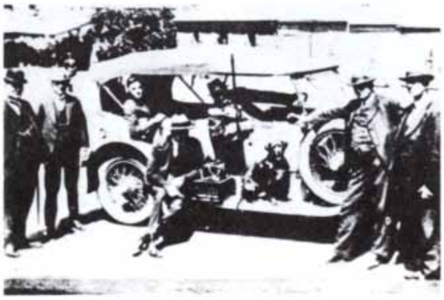
Throughout its history, AWA has had a reputation for quality, innovation and invention. The company operates extensive research and development facilities. The research laboratory is at its major high-technology manufacturing complex in the Sydney suburb of North Ryde. Underlining the importance attached to this laboratory is the fact that the chief scientist who manages this activity reports directly to the chairman of the board of directors of AWA. Manufacture complies with the demanding quality standards for defense and professional customers, meeting the international standards of MIL, DIN and IEC in addition to local Australian requirements.

Since 1972 AWA has been researching the manufacture of optical fibers and optical-fiber communications. Earlier this year the company successfully commissioned a digital internal communications system on H.M.S. *Yarra*, a destroyer of the Royal Australian Navy. The system uses optical-fiber cables to connect digital voice



(Below right) Mobile radio telephones were "leading edge" technology in the 1920's. This picture shows the first such system supplied by AWA. (Above right) Optical Fiber manufactured at AWA carries more signals than the 800 pair cable shown in the background. (Below) AWA designs, develops and manufactures data terminals at its North Ryde plant. This VTE-8 model is typical of current products.





telephones to a digital switch. The equipment was designed and developed in AWA's research laboratories and manufactured by the company. AWA is exploring the opportunities for such systems in defense and commercial outlets.

In other areas, AWA is working on the reduction of the bandwidths required for quality television and telephone transmission using digital signal processing and optical fiber as well as cables. Development is being carried out on local networks using optical fibers to interconnect voice and data terminals.

Fibers are also being used in underwater towed array investigations for the Australian government. The Barra Sonobuoy Project for the Australian and British governments reflects the skill and capability of Australian designers and manufacturers. This product was developed and is produced at AWA's North Ryde plant.

Australia's remoteness from the rest of the world has meant that AWA has had to develop a high degree of self-sufficiency and still remain in the forefront of technological development.

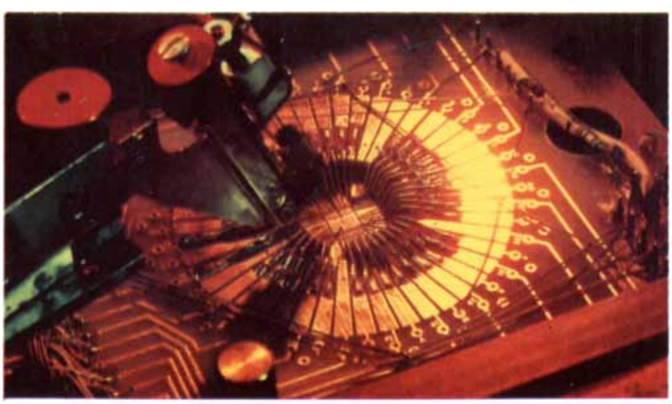
AWA has done this as a public company dependent on investor support.

The Barra sonobuoy is claimed as the world's most sophisticated passive directional sonobuoy design in current production. It was developed under Australian defense funding and is now being manufactured at AWA's North Ryde factory.

AWA's microelectronics facility develops and builds custom integrated circuits in a state of the art plant that offers five micron technology matching international standards. At the Ashfield division multiple manufacture of telephone instruments is a major activity; the Leichhardt-based marine and aviation business on the other hand, is service-oriented, offering both service and agency sales from depots at Australia's major sea- and airports.

The company's traditional markets have been in Australia. From this base a developing export business supplies AWA products worldwide. AWA is one of the major custodians of Australian expertise in advanced electronics technology and thus of strategic importance to the Australian nation.

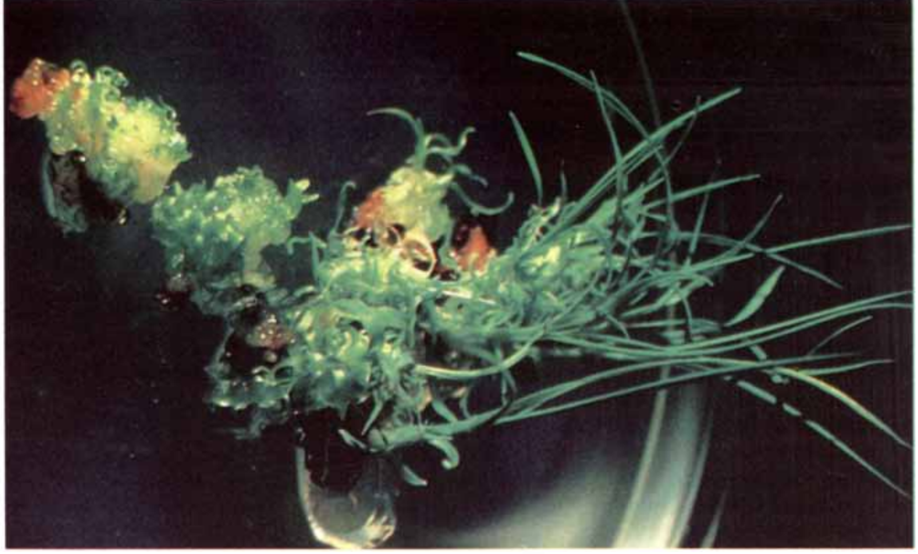
(Below) Specialized data systems such as this slaughter house installation are designed and supplied by AWA. This system includes data terminals capable of accepting steam cleaning. (Bottom) This Doppler V.H.F. Omni Range (DVOR) Navigation beacon assists aircraft landing at Nepal's Kathmandu airport. The system was designed, supplied and installed by AWA's Airways Systems Group. (Below) A state of the art frequency synthesized 900 MHz radio for 60 channel telecommunications. (Bottom) Testing an integrated circuit at AWA Microelectronics manufacturing facility.



more than 200 sheep have been shorn and only about a dozen cuts have occurred. The corporation is to decide shortly whether the research should proceed to commercial production.

Biological control of plant and animal pests is another major area of Australian expertise. Much work is now concentrated on integrated systems of chemical and biological control and on computerized management systems such as the Siratac system for cotton. This is operated from an on-farm computer, which is fed information on pest numbers, crop maturity and weather conditions and provides farmers with daily advice on spraying or other management options.

Much research has also gone into ways of coping with drought and into new farming systems, such as minimum tillage improved ways of handling crops after harvest, and into new crop species for the Tropics, such as kenaf, cassava and the pigeon pea.



Research into rapid breeding techniques conducted by the Division of Plant Industry.

Synroc

A totally new concept for the safe disposal of high-level nuclear waste has been developed by a team of scientists from the Australian National University. It is based on a highly stable synthetic rock known as Synroc, which actually incorporates the radioactive elements into its crystalline structure, immobilizing them for thousands of years.

The synthetic rock consists of three naturally occurring minerals—zirconlite, perovskite and hollandite—made from oxides of five simple elements—titanium, zirconium, calcium, barium and aluminum. These minerals can take into their crystal structure and immobilize almost all of the 50 or so radioactive elements present in high-level nuclear waste.

Synroc is the result of more than four years of work by a research team led by Professor Ted Ringwood, head of the university's Research School of Earth Sciences. Tests carried out in Australia and overseas have shown that Synroc is much more stable than borosilicate glass under high temperatures and pressures.

A demonstration plant under construction at Australian Atomic Energy Commission headquarters at Lucas Heights near Sydney should be completed within the next 18 months. It will be capable of handling 20 kilograms of Synroc an hour, containing four kilograms of waste and would be able to solidify all the wastes from 25 1,000-megawatt nuclear power stations—about the number needed to generate all of Australia's electricity needs.

The AAEC has been responsible for scaling up the Synroc process since 1980 and developed a new technique called "in can" hot-pressing. This made possible production of 30-kilogram Synroc blocks from a 20 percent mix of simulated waste and the Synroc chemicals hot-pressed into stainless steel tubes or canisters.

The technique has been further refined and a new process of uniaxial hot-pressing in bellows-shaped containers developed. The powdered chemicals are transformed to Synroc at extremely high temperatures and pressures (30 Mpa and 1,150 degrees C.) and at the same time the bellows are compressed and collapse uniformly, remaining tightly sealed around the Synroc pancake. These containers can then be sealed and buried up to four kilometers underground in areas of impermeable rock. Borosilicate glass cannot be buried at these depths because it is unable to withstand high temperatures and is susceptible to leaching by groundwater over 100 degrees C.

The Antarctic

Australia's involvement in the Antarctic stretches back to almost the turn of the century, when many Australians took part in the British expeditions led by Scott and Shackleton. The first Australian expedition, mounted by Sir Douglas Mawson between 1911 and 1914, was responsible for establishing a number of bases as well as mapping over 1,000 kilometers of coastline. Many more expeditions followed, and laid the basis for Australia's present claim to sovereignty over almost half of the Antarctic continent—a claim that is contested by many countries. It came into force in 1936 when Australia took over responsibility for a six million square kilometer area stretching south of 60 degrees south between 45 and 160 degrees east longitude, now known as Australian Antarctic Territory.

Mainland stations were set up at Mawson in 1954, Davis in 1957 and Casey in 1969 and all of them have been continuously manned. A $60 million rebuilding program, to be completed by 1990, is now under way.

Each year about 100 men (and, more recently, women) live at the stations over the winter to carry out research in many fields, including glaciology, atmospheric physics, biology, geology, meteorology and medicine. In the summer they are joined by parties of expeditioners who carry out intensive research programs, returning to Australia before the winter pack ice closes in.

Currently, the stations can only be approached during the summer by three chartered supply ships, meaning that the winter expeditioners can be isolated for periods of up to 12 months. This inaccessibility has significantly reduced the effectiveness of Australia's research programs over the years, and there are now plans to establish a new intercontinental air link between Australia and Antarctica. This would involve the construction of three runways: a compacted snow runway at Casey station, a rock runway at Davis and an ice strip at Mawson, as well as the construction of a new Antarctic research vessel. The whole transport system is estimated to cost well over $60 million.

Overall, funding for the Antarctic increased only marginally in this year's budget. Many believe that Australia's present level of activity in the region is much too low to maintain its claim to such a large area of the continent. In fact, the chairman of the government's Antarctic Research Policy Advisory Committee (ARPAC), Professor David Caro, has called a number of times for substantially increased fund-

(continued on page A40)

Rapid Strides in VLSI Chip Design

by Dr. J. Craig Mudge,
VLSI Program, Adelaide

The further microminiaturization of semiconductor technology makes possible the fabrication of very complex functions, including application-specific computers, on a single chip. Mechanisms recently developed at CSIRO for the clear separation of design and fabrication have put Australia's chip-design capability ahead of that of most nations.

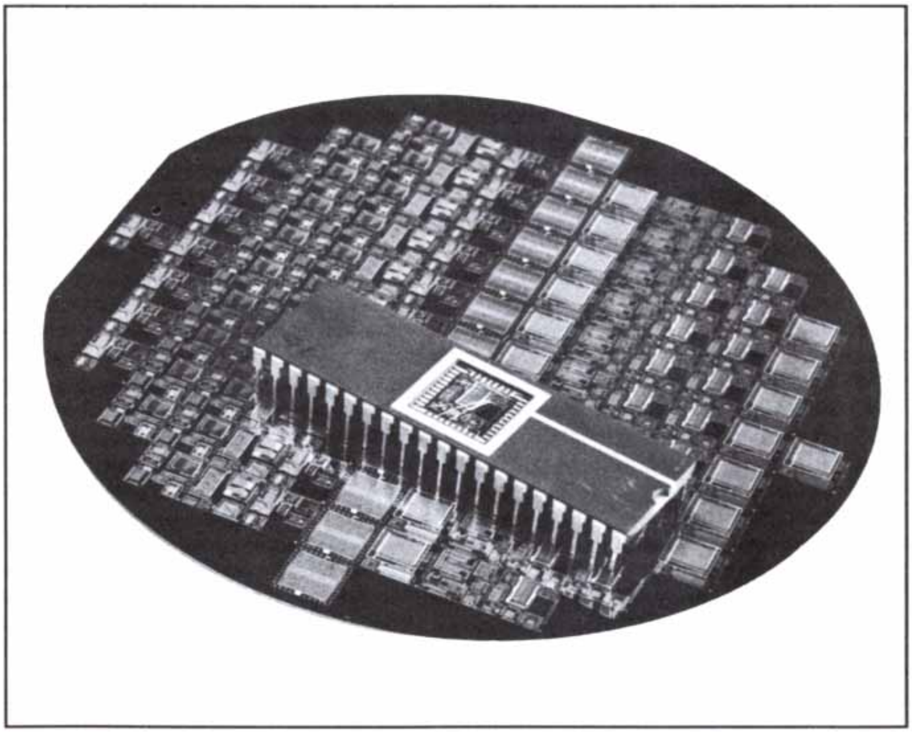
The emphasis on design over fabrication is the major element of a strategy, devised by CSIRO's VLSI program, aimed at achieving a steep increase in Australia's technological capability. The emphasis on design was chosen for two reasons. First, the design of very complex chips is a research opportunity in itself—reducing design time to a reasonable level, that is, a few man-years, is a widely sought goal. Second, a nation with a small population has a better chance of making an impact on international technology if it concentrates on the brain-intensive aspects of VLSI technology rather than on the capital-intensive ones.

In just under two years, Australia has moved from the position of having no user-designed chip capability to the point where many designers have been trained and nearly 200 different chips for application-specific functions have been conceived and fabricated. Although most of these early designs were done by university staff and students, about a dozen chips from innovative companies and research institutions are going into production, embedded in systems primarily for export. Applications have included a borehole logging probe for mineral exploration, a robotics vision processor, a cochlea implant, several telecommunications circuits, signal processing for radio-telescope data, computer-graphics subsystems and some novel analog circuits.

This rapid propagation of knowledge is being used as a model by many European and Southeast Asian countries wishing to develop a similar, widespread chip-design capability. The seeding exercise began in February, 1982, with a three-week instructors' workshop in Adelaide. Most of the participants were university lecturers, who returned to their institutions and established chip-design courses based on the pioneering work of Mead and Conway. The backbone of the rapid propagation is AUSMPC, Australia's Multi-Project Chip Implementation System. Through its cost-sharing technique and its clear interface with fabrication, AUSMPC provides a rapid chip-prototyping service for designers from each state in Australia. CSIRO coordinates AUSMPC over a national electronic-mail communications network. Four runs per year, containing about 30 independent designs each, are now being conducted. Some of the steps of fabrication, in particular electron-beam mask manufacture and wafer processing, are subcontracted to firms in the U.S. Typical turnaround from the submission of design data to the receipt of packaged, documented chips is 10 weeks.

Having fulfilled its technology transfer role by early 1983, the VLSI program turned to its major research: it aims to design chips with 10 to 20 times the density of AUSMPC designs. To achieve these densities, which are in excess of 100,000 transistors per chip, two advances were necessary. First, a new design methodology and associated computer-aided design programs had to be developed. Second, the existing design/fabrication interface, adequate for a five-micron process, had to be carried a step further to interface with an advanced two-micron process. Through its research collaboration with VLSI Technology, Inc., in San Jose, Calif., CSIRO has developed a standard interface with a double-metal, two-micron NMOS process. Two test runs have been completed. CSIRO's design capability is being measured by a signal-processing application; a single-chip word recognizer, using dynamic-time warping for reference comparison, is nearing completion.

CSIRO's research capability in VLSI architecture and its design technology are resources for industrial firms embedding chips of their own design in advanced electronic equipment. Scientists and engineers have access to chip-design courses in their local universities, as well as to three commercially available short courses. Moreover, the shared-cost fabrication technique has been emulated by three universities, including the Royal Melbourne Institute of Technology, which developed the world's first multiproject gate-array implementation system.



A user-designed chip and the four-inch wafer from which it was packaged. With this technique, from 30 to 50 Australian designs share an expensive fabrication run.

ing and levels of research in the region. In August he warned that if Australia is to remain a major Antarctic power, it must provide adequate transport and sufficient funds for a first-rate research program, saying, "If Australia cannot afford this level of activity, it would be more honest to move out or at least to reduce the number of stations to be supported."

Australia's future activities in the region now much depend on the importance the government decides to place on its claims to the region, given the potentially rich mineral and marine resources. This is likely to be a decision that will be made outside the science and technology portfolio and within the government's overall foreign affairs policy.

Astronomy

Australia has been at the forefront of international optical and radio astronomy for almost 40 years, due in part to its unique position to observe the southern skies but also because of the wealth of major astronomical instruments that have sprung up in the country since the 1950's.

From 1945 to 1960 a team of CSIRO scientists was responsible for a string of new discoveries, which culminated in the construction of the 64-meter Parkes radio telescope in New South Wales by CSIRO in 1961. It was one of the most advanced radio telescopes in the world, and it was used to make many major discoveries, including the detection of the first quasar.

The Parkes telescope has now been superseded by a number of more sophisticated instruments in the Northern Hemisphere, and this led the federal government last year to fund the construction of a new $30-million synthesis telescope. It will link major telescopes across the continent to form one of the most sensitive high-resolution radio telescopes in the world.

Six new 22-meter antennas are to be built, five of them in a 6-kilometer line at Culgoora Observatory in New South Wales and the sixth at Siding Spring Observatory. These will be linked with the existing Parkes telescope, and operating together they will form a single telescope with a dish diameter equivalent to 300 kilometers.

Further links with dishes at Carnarvon in Western Australia, Alice Springs, Tidbinbilla near Canberra, Sydney University and Hobart will extend the telescope to a dish diameter of 3,000 kilometers, spanning the entire continent. It will be known as the Australia Telescope and begin operation in the bicentennial year, 1988.

Optical astronomy has also flourished with the construction of a number of new instruments in the 1960's and 1970's, in particular the 3.9-meter Anglo-Australian telescope, which is jointly operated by Britain and Australia. This telescope has dominated optical astronomy since its first use in 1974; it has been responsible for many new discoveries, including the Vela Pulsar in 1977 and most recently the identification of the fastest spinning pulsar known in the universe, PSR 1937-21.

The Anglo-Australian telescope is at the Siding Spring Observatory near Coonabrabran in New South Wales, which is the major optical astronomy site in the Southern Hemisphere and the home for other instruments. These include the U.K. Schmidt telescope, which has played a significant role in mapping the southern skies.

A new 2.3-meter optical telescope is nearing completion at the observatory, which has been built to a totally new design devised by scientists from the Mount Stromlo Observatory near Canberra. It combines three major technological advances—a rotating building, a large, thin mirror and an altazimuth mounting—into a single telescope, cutting the cost to a tenth of that of a conventional instrument. This makes the telescope extremely versatile and will make possible its use 24 hours a day, operating in the infrared during the daytime.

In all, Australia has about 40 optical and radio telescopes, and in addition a number of X-ray, infrared and cosmic-ray instruments.

The country is also moving into the field of space astronomy and is a partner with Canada and the U.S. in the Starlab project: a joint initiative to launch a space telescope aboard the space shuttle in 1990. Australia is responsible for designing the central element of the telescope—the instrument package, which contains a unique ultralarge-format photon-counting array developed by scientists from Mount Stromlo Observatory.

Medical Research

The pioneering in vitro fertilization work by Professor Carl Wood and his team from Monash University in Melbourne has highlighted worldwide the excellence of Australian medical research. The work has led to the birth of the first test tube baby, the first test tube twins and triplets and the implantation of the first frozen embryo earlier this year.

Major advances have included work by Sir Macfarlane Burnet on viral vaccines, the establishment of the effects of thalidomide by Sir William McBride and pioneering work by Sir Norman Gregg on the effects of maternal Ru-

(continued on page A44)

AUSTRALIAN RESEARCH ORGANIZATIONS

CSIRO Headquarters, PO Box 225, Dickson, ACT 2602, tel: (02) 484211
CSL 45 Poplar Rd, Parkville, VIC 3052, tel: (03) 3891911
AMDEL Head Office, Flemington, Frewville, SA 5063, tel: (08) 791662
AERONAUTICAL RESEARCH LABORATORIES PO Box 4331, Melbourne, VIC 3001, tel: (03) 6477511
DEPARTMENT OF SCIENCE AND TECHNOLOGY PO Box 65, Belconnen, ACT 2616, tel: (062) 641155
ANTARCTIC DIVISION Channel Highway, Kingston, TAS 7150, tel: (022) 290209
AUSTRALIAN ACADEMY OF SCIENCE PO Box 783, Canberra, ACT 2601, tel: (062) 486011
AUSTRALIAN ACADEMY OF TECHNOLOGICAL SCIENCES 191 Royal Parade, Parkville, VIC 3052, tel: (03) 3473166
AUSTRALIAN ATOMIC ENERGY COMMISSION Lucas Heights Research Laboratories, Private Mail Bag, Sutherland, NSW 2232, tel: (02) 5430111
AUSTRALIAN MINERAL INDUSTRIES RESEARCH ASSOCIATION 11th Floor, 63 Exhibition St, Melbourne, VIC 3000
AUSTRALIAN SCIENCE AND TECHNOLOGY COUNCIL PO Box 52, Canberra, ACT 2600, tel: (062) 724655
BUREAU OF METEOROLOGY 150 Lonsdale St, Melbourne, VIC 3000, tel: (03) 6620311
BUREAU OF MINERAL RESOURCES PO Box 378, Canberra, ACT 2601, tel: (062) 499111
HOWARD FLOREY INSTITUTE University of Melbourne, Parkville, VIC 3052, tel: (03) 3451844
MATERIALS RESEARCH LABORATORIES PO Box 50, Ascot Vale, VIC 3032, tel: (03) 317222
NSW SCIENCE AND TECHNOLOGY COUNCIL 151 Macquarie St, Sydney, NSW 2000, tel: (02) 273978
SOLAR ENERGY RESEARCH INSTITUTE OF WA GPO Box R1283, Perth, WA 6001, tel: (09) 3264157
WALTER AND ELIZA HALL INSTITUTE Post Office, Royal Melbourne Hospital, VIC 3050, tel: (03) 3471511
AUSTRALIAN SCIENTIFIC INDUSTRIES ASSOCIATION c/o Regional Administrative Office, 9 Queens Rd, Melbourne, VIC 3004, tel: (03) 2687111
AUSTRALIAN INSTITUTE OF MARINE SCIENCE Cape Ferguson, PMB No. 3, Townsville MSO, QLD 4810, tel: (077) 789211
AUSTRALIAN COAL INDUSTRY RESEARCH LABORATORIES 14 Delhi Rd, North Ryde, NSW 2113, tel: (02) 8875341
VICTORIAN SOLAR ENERGY COUNCIL 151 Flinders St, Melbourne, VIC 3000, tel: (03) 6544533

A Major R&D Spender

ICI Australia has operated one of the largest private-enterprise research organizations in Australia since the early 1950's. Its list of achievements is notable, and some of the country's most prestigious awards have been received by ICI scientists.

While working as part of the worldwide ICI group, ICI Australia has carried out significant research and development to solve problems posed by the particular markets and environment in Australia. In a number of cases the results of this research have been licensed overseas.

In the early 1960's the manufacture of the two basic industrial plastics, polyvinyl chloride and polyethylene, were integrated around a common feedstock (ethylene), and an Australian petrochemical industry became a reality. The key was the oxychlorination route to vinyl chloride. While this was simultaneously developed throughout the world, as a result of local research ICI Australia was the first in the Southern Hemisphere to have a plant on stream.

At about the same time, the flame ionization detector was invented and subsequently licensed to 37 companies worldwide. This device launched a new era in the detection and identification of minute quantities of material, and probably no other instrument has made as great a contribution to environmental protection. The Royal Australian Chemical Institute recently organized a commemorative symposium to celebrate the discovery.

The company has strongly supported local agriculture by manufacturing fertilizers and since 1954 has actively developed animal-health and plant-protection products. When the anthelmintic tetramisole was discovered in Belgium, ICI Australia could have contented itself with obtaining selling rights. Instead, in competition with the chemical giants of the world, a novel economical manufacturing process was developed and licensed overseas. When the improved drug levamisole replaced the earlier one, ICI Australia again developed the best commercial route and continues to receive significant royalties in addition to the advantages of local manufacture. Levamisole has been a major aid to farmers faced with the difficult problem of controlling intestinal worms in sheep and cattle.

Local needs have always been of prime importance in setting research targets. When the cattle tick problem reached epidemic proportions in Queensland in 1971, ICI Australia had already been working on the problem for seven years. Two years later Promacyl, probably the first biologically effective chemical entirely discovered and developed in Australia, was marketed. It has provided excellent tick control since 1973, in contrast to the short life of most overseas-developed tickicides. A new-generation tickicide with enhanced properties, Cyhalothrin, followed in 1982 and is destined for a wide spectrum of markets.

When the health hazards of vinyl chloride in PVC manufacture were recognized, ICI Australia developed its own plant modifications, including the inhibition of the buildup of deposits in reactors and steam-stripping product purification, rather than face the time lag in following overseas developments. Major contributions to other Australian industries include the development of safe, water-resistant, slurry explosives for underground blasting with minimal shock waves, and high-boiling brake fluids to provide cooler and safer braking for Australia's motor vehicles.

The ICI Australia group has been in the forefront of many of the developments that helped raise living standards and made significant advances in the areas of medical care, food production, textiles, paints and building materials. Its range of products includes industrial chemicals, commercial explosives, fertilizers, agricultural chemicals, pharmaceuticals, plastics, synthetic fibers and paints. A subsidiary, Dulux Australia, developed a unique paint ingredient called Spindrift, which made much higher opacity possible. The biggest paint manufacturer in the U.S., Sherwin Williams Company, has signed an agreement for rights to use the Dulux technology, Tioxide International has taken out a license to develop the U.K. and Western European markets and other agreements are under way in Japan.

Industry and academic collaboration has long been encouraged. Many leading Australian academic scientists had their early career in ICI Australia research. A major collaboration with CSIRO led to the commercial development of an imaginative ion exchange process for the desalination of brackish and effluent water.

All major developments require a variety of skills. ICI Australia's scientists and engineers work in the context of a diverse organization of 11,000 people. Although a majority ownership lies with ICI PLC, the largest chemical manufacturing company in the U.K., a substantial percentage of stock is owned by Australian interests. While benefiting from the resources and experience of the worldwide ICI group, the Australian company has demonstrated both a commitment to local needs and problems and the capacity for world-level research and development. Its association with the ICI group means that processes and products flowing from this research can be marketed effectively in many countries.

The company is justifiably proud of its past achievements and intends to maintain its reputation in the research and development of new products and processes. There is a continuing review and reallocation of resources. The company moved a few years ago to provide contract services in computing and chemical engineering to meet the needs of other industries, education and government. More recently it set up ICITACS to make available even more widely its technical, analytical and computing expertise at three levels: consultation and advice; design, construction and production, and routine scientific and technical services. Already ICITACS has available compact interfaces for real-time monitoring and devices for control of data acquisition by computers.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION:

Strength Through Understanding

Australia is geographically far from the mainstream of current military conflicts. However, the island continent, ringed as it is by long stretches of sparsely populated coastline, still needs to maintain a modern defense force that uses technology to the best advantage.

The Defence Science and Technology Organisation (DSTO) advises the Australian Defence Force (ADF) on technology that is cost-effective for the Australian environment. DSTO also adapts and develops new equipment specifically for Australian conditions.

Over 1,000 research scientists and engineers at DSTO's 10 establishments, spread across Australia, work in a wide range of fields from aerodynamics to electronic warfare, from the design of sonar systems to the development of new materials.

The Thrust of Defense R&D

DSTO's mission-oriented R&D provides the knowledge base for solving hardware problems, for advising on the long-term development of the ADF and for scientific support and guidance in defense policy. The type of work ranges from consulting tasks for the ADF and industry to much larger projects such as the Jindalee over-the-horizon radar system, which can detect aircraft and ships thousands of kilometers from the transmitter.

A strong technology base is maintained to evaluate the implications of new weaponry, as well as to develop effective solutions to special and immediate problems and provide the expertise to handle future problems. DSTO needs to identify in advance new technologies that will be particularly useful for Australia in multiplying defense force effectiveness; for example, wide-area surveillance and electronic warfare equipment, "smart" weapons and remotely piloted vehicles. It is then ready to advise on the purchase of new equipment (such as the recent selection of the F/A-18 to replace the Royal Australian Air Force's aging Mirages) and how the ADF can make the best use of the equipment.

Industrial Involvement

Because of Australia's small population and limited industrial R&D and manufacturing base, much of the equipment used by the ADF is bought overseas. Some of this, however, must undergo substantial alteration to become suitable for the Australian environment, or alternative equipment must be developed locally. Increasing emphasis is also being placed on self-reliance so that maintenance and modifications can be made efficiently by local industry and the life of expensive equipment can be extended.

Military equipment developed specifically for Australian needs includes the Mulloka high-frequency shipborne sonar, which is designed to be effective in the warmer waters around much of the continent, and the Jindalee over-the-horizon radar. Other Australian-developed equipment has found worldwide application, for example, the Ikara long-range antisubmarine weapon system, which has been in service with the Australian, British and Brazilian navies for over two decades, and the Barra air-dropped sonobuoy, which is used by both the Royal [British] Air Force and the Royal Australian Air Force.

The Ikara, Barra and Mulloka projects and the Karinga cluster bomb have led to $50 million in development contracts in industry and $320 million in production contracts. Over $140 million worth has been exported.

Industry is involved as early as possible in developments likely to lead to volume production, and DSTO devotes significant resources to helping transfer R&D into industrial production.

Moreover, DSTO provides a wide range of consultancies and technical support to help industry gain offset contracts. For instance, a Queensland firm, assisted by DSTO, has developed an electroslag welding technique that has enabled it to successfully compete for the fabrication of the external A-frame propeller shaft bearing for patrol frigates bought from the U.S. The component is now being provided to the U.S. Navy shipbuilding program.

Civilian Applications

Defense research has had a number of civilian applications. Oceanographic studies have helped in navigation, fisheries and meteorology. Structural-fatigue research has assisted in the development of monitoring equipment used on both military and civilian aircraft.

DSTO expertise has helped in the prototype Interscan microwave landing system and in the construction of cochlea implants for a bionic ear developed by the National Acoustics Laboratories. DSTO constructed a high-level-impulse noise generator that has aided in studies of hearing loss of both military gunners and industrial construction workers. Other examples include a letter-bomb detection system used at the Melbourne Commonwealth Heads of Government Meeting and the Brisbane Commonwealth Games and a cable cutter that saves low-flying agricultural aircraft from crashing if they hit power lines.

When defense priorities permit, DSTO facilities can be made available for outside research. For instance, the organization's wind tunnels have been used in the design of buildings and ship funnels and for studying airport wind patterns.

Scientific Contacts at Home and Abroad

As a matter of policy, contacts between defense scientists and researchers in industry, government and universities are being increased. In the past five years the number of research contracts with tertiary institutions has greatly expanded. They now include work on signal processing, blue-green lasers, aircraft gust loads, strength degradation of brittle ceramics and the growth of single crystals for lasers.

DSTO values its strong cooperative links with allied nations, which enable it to keep in touch with overseas developments and to avoid duplication of effort. In addition to regular exchanges of staff and information through a network of formal arrangements, there are also joint research projects and programs.

The Way Ahead

DSTO will continue to update its skills and its technology base to enable it to solve pressing problems for the Australian Defence Force and the defense industry, and to provide them with soundly based advice. Three solutions that have met the challenges of the Australian strategic, military and physical environments are described opposite.

Examples of DSTO's Work

DSTO's tasks range from solving several thousand small problems for the Australian Defence Force and the defense industry each year to large indigenous projects such as Jindalee. The following three examples have been chosen as being representative of the larger tasks.

Camouflage

The distinctive greens, blues and yellows of the Australian bush require

The photo at left shows the Australian-made and -designed camouflage uniform, which blends effectively in both dry and wet Australian environments. The photo at right shows the use of a DSTO-developed crack patching kit on a Mirage wing. Patching cracks caused by fatigue or stress is a relatively simple idea that required some high-technology research to make it effective on high-performance aircraft.

matching camouflage. As a result of extensive research at the Materials Research Laboratories (MRL) in Melbourne on the colors, patterns and textures of Australian deserts, savannah, temperate forests, tropical jungles and snow country, essential characteristics of the landscape have been quantified. Using data on pattern shape and color correlations, a computer was used to derive suitable shapes and colors for a disruptive camouflage pattern.

Rubber technologists at MRL have also designed a directed molded PVC-nitrile rubber sole, pigmented to match the disruptive patterning of a combat uniform, and tested it with various tread patterns. Toe caps and boot uppers are similarly camouflaged. Reversible camouflage nets that match either wet or dry environments have also been designed. Extensive field trials have been conducted to ensure that the textiles and dyes used for uniforms and nets can withstand Australia's intense ultraviolet light and other harsh conditions.

Satellite imagery is used to classify terrain and vegetation so that appropriate camouflage can be selected for a given area. Scientists are also working at reducing the infrared signatures of vehicles and equipment and studying materials that will obscure sources of infrared on the battlefield in the same way that smoke obscures sources in the visible part of the spectrum.

Crack Patching

Metal fatigue and stress corrosion can lead to the development of cracks in aircraft components. Replacement of the affected parts is usually expensive and time consuming, and in situ repairs, using traditional methods, can cause further cracking and internal damage. However, scientists from the Aeronautical Research Laboratories in Melbourne have developed a new, highly cost-effective method of patching cracked components in the field quickly and efficiently. The technique is based on advanced fiber-reinforced plastics adhesively bonded to the cracked component and has been used successfully on a wide variety of Royal Australian Air Force machines, including the wings of the Hercules and Mirage aircraft, the fuselage of Orions and the wheels of Macchis.

The technique has a number of advantages over conventional ones. Adhesive bonding provides a very efficient load transfer from the cracked component and produces a sealed interface, which reduces the danger of internal leaking such as from a fuel tank and further corrosion. The advanced fiber-reinforced plastic also gives high directional stiffness, good resistance to cyclic loading and corrosion, excellent formability and low density.

Field techniques and kits have been developed that contain grit-blasting and anodizing units for surface treatment, and temperature controllers and hydraulic pressure indicators for use during the adhesive curing process. Specialized training programs are also run to teach Royal Australian Air Force technical personnel how to apply patches in the field.

Laser Bathymetry

Using existing conventional hydrographic techniques, it would take over 50 years to map Australia's coastal waters, which conceal a vast continental shelf with an area of over 2.2 million square kilometers, more than half of it under a depth of less than 50 meters.

DSTO has developed a unique laser airborne depth sounder (LADS) that will be used to chart Australia's shallow coastal waters. It will obtain data at least 10 times faster than present techniques using boats, and in areas inaccessible to boats.

LADS has been developed entirely within Australia, by scientists from the Electronics Research Laboratory in South Australia, and has now been taken up by a consortium of Australian companies led by Thorn-EMI Electronics (Australia) Pty. Ltd. Commercial production of the system is expected to begin in 1984. Very short (five nanosecond) pulses of green and infrared laser light are simultaneously beamed down to the water from an aircraft flying at a height of 500 meters and a speed of 70 meters per second. The infrared pulse is reflected from the surface and used as a surface reference, while the green pulse penetrates into the water and is reflected from the bottom and back to the aircraft. Water depth is then determined by measuring the time difference between the arrival at the aircraft of the reflected infrared and green pulses.

The green pulse is scanned back and forth across the water perpendicular to the track of the aircraft to form a swath 260 meters wide with a nominal spot spacing of 10 meters. Soundings are made at the rate of 168 per second and have been made in clear water to a depth of over 50 meters. Computer software and hardware have been developed to process within 24 hours the two million soundings a day that will be produced by the LADS system and to supply validated data to the Australian Hydrographic Data Base.

bella on the developing baby.

There are now 10 medical schools in Australia, plus the John Curtin School of Medical Research at the Australian National University, 14 private or hospital research institutes and half a dozen government health laboratories. Funding comes from the government, private sources and overseas research grants and in 1982–83 topped an estimated $90 million.

Compared with other areas of science and technology, medical research is relatively generously funded by the government: a total of almost $60 million was provided in 1982–83.

The major government-funded research center is the John Curtin School of Medical Research, which has a budget of over $10 million and a staff of more than 300. The largest of the private medical research institutes is the Walter and Eliza Hall Institute of Medical Research, set up in 1919 and now a world leader in immunology, cancer and molecular biology research. The institute has recently been responsible for isolating malaria antigens, which will be used in developing a malaria vaccine in Australia.

Other major private medical research institutes include the Howard Florey Institute of Experimental Physiology and Medicine, which specializes in endocrinology, hypertension and behavior and has done much pioneering work on the human birth hormone relaxin, and the Baker Medical Research Institute in Melbourne, which concentrates on cardiovascular research, atherosclerosis and hypertension.

In Sydney, independent research institutes, such as the Kanematsu Research Institute, are at the forefront of leukemia research, and others, such as the Garvan Institute of Medical Research, are pioneering work into metabolic disorders such as diabetes.

Major advances made by Australian medical researchers in recent years include the determination of the link between analgesics and kidney damage, detailing of the natural history of melanoma and the mechanism of the inheritance of malignant hyperthermia. Other outstanding work has included the discovery of rotavirus as the cause of infantile diarrhea in both developed and developing countries, the use of transfer factor to slow down the progression of multiple sclerosis and work on tissue rejection and the transplantation of pancreas islets in the treatment of infantile diabetes.

The use of ultrasound as an investigative tool in cancer and the development of a number of new ultrasonic techniques have been pioneered by the Ultrasonics Institute in Sydney.

Medical research, particularly as preventive medicine, is a major priority in Australia and government support is likely to continue at a relatively high level.

Biotechnology

Despite a strong research base, world class researchers and a number of major advances in biotechnology, Australia has lagged behind in the establishment of a commercial biotechnology industry.

The government, however, has identified biotechnology as a major area of high-technology development, and this year, for the first time, funds have been earmarked specifically for biotechnology research and industrial scale-up. A total of $12.2 million has been provided for a range of projects, including $1.5 million to support the new National Biotechnology Scheme, aimed at commercializing university and other research. This is seen as totally inadequate to meet the demand, but it is hoped it will act as a catalyst to get the industry off the ground.

So far, major activity has been in the field of monoclonal antibodies. A number of diagnostic kits for animal and human diseases should soon reach the market. A number of genetically engineered animal vaccines, including one for scours in pigs and another for fleece rot in sheep, are also nearing commercial production.

One of the most exciting projects is the plan to develop a malaria vaccine in Australia, following the successful isolation of antigens from one of the deadliest strains of malaria, *Plasmodium falciparum,* by a research team from the Walter and Eliza Hall Institute for Medical Research in Melbourne.

The institute; the government-owned Commonwealth Serum Laboratories; a private company, Biotechnology Australia Pty Ltd., and the Queensland Institute of Technology hope to develop the vaccine locally over the next five years. This represents a major step forward, as until now much of Australia's best work has had to go overseas for commercial development. For instance, most recently further work on the human pregnancy hormone relaxin, successfully isolated by a team from the Howard Florey Institute in Melbourne, was handed over to Genentech in the U.S.

Another major area of expertise in Australia is in plant genetics. Researchers at CSIRO's Division of Plant Industry have isolated a transposable element, or more colloquially, "jumping gene," in maize that is capable of moving from place to place on a chromosome or between chromosomes. It could be used to insert new desirable genes into the plant to combat disease or promote drought resistance.

In other work, the Australian National University's Centre for Recombinant DNA Research, headed by Dr. John Shine, is investigating the transfer of nitrogen fixation genes from legumes to nonlegume crops. The work is being supported with a $1.8 million grant from Agrigenhics Corporation of Boulder, Colorado. Already the team has managed to isolate a number of genes that enable bacteria to interact with the plant to form a root nodule and genes that make possible nitrogen fixation, thus putting the center at the forefront of world research in the area.

The center also recently has been responsible for cloning the human

CSIRO's Sirotem, a computerized direct digital borehole logging system and data processing center, can detect ore bodies covered by thick layers of conductive material.

CSIRO

Assembly-line production of the atomic absorption spectrophotometer.

CSIRO

blood pressure hormone renin for the first time. This work is being supported by California Biotechnology Inc. in the U.S. The aim is to eventually design an inhibitor of renin that can be used as a hypertensive agent for treatment of high blood pressure.

The center is also involved in a collaborative project with the University of New South Wales and the Garvan Institute of Medical Research in Sydney to produce commercial quantities of human growth hormone for the Australian market. Novel biological insecticides, animal hormones and vaccines, animal breeding, waste treatment and ore leaching are other areas of active research.

The future development of a biotechnology industry in Australia now depends on how effective new venture-capital tax incentives and increased research grants are in mobilizing the industry and freeing finance for commercial development. Currently, regulation of biotechnology research and development is purely voluntary. The federal government has established the Recombinant DNA Monitoring Committee, responsible to the Department of Science and Technology, which has laid down voluntary guidelines for laboratory and industrial research. These specify various levels of laboratory containment, reporting procedures, the setting up of laboratory safety committees and approval procedures aimed at eliminating or minimizing risk from the work. The guidelines are based on those laid down by the National Institutes of Health in the U.S.

Lawyers, theologians and a number of concerned scientists have raised community concern over the use of recombinant DNA techniques and have called for legislation to control the new technology. However, at present it looks as if regulation of the embryo industry will remain voluntary.

Solar Energy

Not surprisingly, much research has gone into the use of Australia's most abundant energy resource—sunshine. In the 1950's and 1960's Australia led the way in the development and use of solar collectors and has a flourishing domestic industry and is the world's largest exporter of solar systems.

Research is currently under way to improve both the basic flat plate collector and to develop improved solar water heating systems. Most recently Rheem Australia Ltd. decided to establish a pilot manufacturing plant to produce evacuated-tube collectors developed by Sydney University researchers. These will be aimed at both the industrial and the domestic market. A Japanese company, Nitto Kohki, has negotiated a licensing agreement with the university to manufacture the system in Japan.

Efforts have also been made to harness solar energy on a much larger scale, and two thermal power plants have been completed. One, a 25-kW unit at White Cliffs, near Broken Hill in New South Wales, has been designed and built by researchers from the Australian National University. It uses 14 five-meter mirror-glazed collectors focused on a central solar absorber to produce steam at 550 degrees C. This is used to drive a steam engine. The plant was completed last year, but due to a number of problems still has not been commissioned by the NSW Energy Commission.

The other station is a 100-kW hybrid plant built at Meekatharra, 600 kilometers north of Perth in Western Australia, which uses German technology developed by MAN. It produces 50-kW of electricity using 960 square meters of parabolic trough-concentrating collectors and produces an additional 50 kW from waste heat recovered from the town's existing diesel system. The $3.6 million cost of the project has been jointly shared by the West German and Western Australian governments. It is part of a much larger remote area power supply investigation in the state that involves studies of wind, ethanol, photovoltaics and vegetable oils as alternative energy sources.

Telecom has already used solar cells extensively to extend its network of microwave communications systems in the outback. It recently opened the first stage of what is believed to be the world's longest solar-powered microwave network, which will span 1,595 kilometers, linking the remote Kimberley region of Western Australia with Port Hedland.

Very promising work is also under way at the University of New South Wales on an extremely efficient solar cell that uses a new method of cell fabrication. Efficiencies of over 18 percent have been achieved. The work is being supported by NASA and the U.S. National Bureau of Standards.

Wind power is also considered a viable possibility in Australia, and there are proposals for a $1 million wind farm for northwest Tasmania and a system for Lord Howe Island. However, overall government interest in solar research has waned in recent years, mirrored in the level of funding, which fell marginally last year from $1.9 million to $1.8 million.

Oil, gas and electricity prices will have to rise substantially before solar energy becomes a viable alternative for many applications in Australia.

Scientific Instrumentation

Australia's most successful and well-known invention is the atomic absorption spectrophotometer, which has revolutionized chemical analysis around the world. It was the result of a flash of inspiration in a Melbourne garden one Sunday morning back in 1952, when Dr. Alan Walsh realized that the best way of detecting minute quantities of metallic elements was to measure their absorption rather than their emission of radiation. By the following morning he had set up a successful experiment at his laboratory at CSIRO's Division of Chemical Physics, which laid the basis for what is now called the most significant advance in chemical analysis of this century.

At the time, no one either in Australia or overseas was interested in developing the instrument, but it was finally taken up by a small Melbourne company, Techtron Appliances. This was taken over by Varian Associates

Australian States: Recent Developments

NEW SOUTH WALES: Australia's richest and most populous state is in the forefront of the resources boom. Major restructuring of the steel industry is under way and more than $350 million in export sales are expected this year from the new CSR–Pechiney aluminum smelter in the Hunter region and an expanded potline at Alcoa's Kurri Kurri plant. In addition, rural output is expected to double with a record wheat crop of over 6 million metric tons, which alone will inject $86 million into the state's economy.

The state government, led by Mr. Neville Wran, has developed a technology development policy that aims to increase economic development and technical competitiveness, encourage technological innovation, identify specific technologies for development and attract new technology manufacturers to the state.

SOUTH AUSTRALIA: In its search for a new industrial direction, South Australia is adopting a high-technology profile. Technology Park Adelaide has been established and the government hopes the state's background experience of weapons research and related enterprise will encourage new companies to locate there.

British Aerospace, EMI Electronics, Fairey and Raytheon have been established for some years but the real basis for South Australia's future prosperity lies in two large resource projects, the Cooper Basin oil and gas liquids project and the Roxby Downs copper, gold, uranium and silver mines. When the Roxby Downs mine starts production in 1990, South Australia expects a jump in its export income of 50 percent. A tarmac road from Adelaide to Alice Springs is planned to link up with the present road south from Darwin by 1988.

WESTERN AUSTRALIA: The largest state, Western Australia is rich in oil, gas, diamonds, gold, uranium, nickel and aluminum.

Brian Burke, the premier, plans to set up a development corporation to administer the state's stake in the Argyle Diamond mine which is potentially the world's biggest. Situated in the northeast Kimberley area, the mine is expected to generate infrastructure development by mining companies including a new all-weather airport and eventually a hydroelectric generating plant on the Ord River.

Talks on a high-technology joint venture with Japanese participation with Western Australia's providing the software are under way. Foreign investment is regarded as essential, particularly in large resource ventures.

VICTORIA: Occupying only three percent of the Australia land mass, Victoria has over a quarter of its population. It produces most of the continent's oil, much of its natural gas and has vast reserves of brown coal for electric power generation.

With industry predominating, agricultural production represents just over one-fifth of the national total.

The state government is developing a science and technology strategy within the context of a much broader industry and regional development policy. A science and technology unit has been established within the Ministry for Economic Development and a biotechnology advisory committee has also been set up.

Many of Australia's largest companies are headquartered in the state and Melbourne, the capital, continues to be an important financial center.

QUEENSLAND: The sunshine state has a population of 2.4 million, which is growing at nearly twice the national average. State taxes are the lowest in Australia and in 1981–82 nearly 28,000 companies and more than 40,000 new business names were registered. Planned public sector spending on infrastructure is about $8,000 million. The state has the world's largest bauxite deposits, a fully integrated aluminum smelting industry and huge reserves of oil shale and coal, and in the past seven years the gross value of annual minerals production has nearly doubled to more than $2,000 million.

The last of the mainland state governments to move into the technology field, Queensland decided last July to establish a Science and Technology council and to appoint a minister for technology.

TASMANIA: The island state has embarked on an expansionist policy with the government promising funds for the creation of a Tasmanian Development Authority as well as for numerous infrastructural projects and measures aimed at stimulating industry and commerce. Among initiatives announced are a record government works program of $470 million, an upgrading of major fishing ports, a start on design work for a major irrigation dam, a 50 percent increase in funds for tourism promotion, a record road building program and planning for the upgrading of Hobart airport to international standards.

Gross value of farm production in 1982–83 was $313 million against $760 million for minerals and metal refining.

in 1967, and now Varian Techtron in Melbourne is the world's second-largest maker of spectrophotometers and a major producer of UV-visible spectrophotometers. Sales exceed $25 million a year, and nine out of 10 instruments are exported to Europe and the U.S.

Atomic absorption spectrophotometers can now detect 66 different elements, and a single apparatus can make up to 1,700 analyses of different samples a day.

Another major Australian achievement, also in the 1950's, was the invention of the flame ionization detector, which made the then-new technique of gas chromatography 10,000 times more sensitive virtually overnight. It was invented by Mr. Ian McWilliam of ICI Australia's Melbourne laboratories, who realized that the components of a sample mixture could be detected by conductivity measurements if they were ionized as they left the column. The development has made an enormous contribution to environmental protection around the world, and every day thousands of laboratories use the device to measure impurities in the air, water, drugs, food and chemicals.

Victoria is now at the heart of scientific instrument manufacture in Australia and has a number of firms such as GBC Scientific Equipment, Labtest Australia and R&D Instruments directly involved in spectrophotometer development.

Another Melbourne company also involved in the field is Scientific Glass Engineering, which has developed a number of microsyringes for gas chromatography and has chalked up a world first by using glass-lined stainless steel tubing.

Another company, Selbys Scientific, is producing novel optical fiber nephelometers for taking high turbidity measurements, such as the concentration of metallic wear particles in engine oil; another, Silenus, has developed a highly sensitive clinical viscometer, which uses an innovative magnetic stabilizing system.

Australia also excels in astronomical instrumentation, as shown by its involvement in the design and development of the Starlab instrument package (the joint U.S., Canadian and Australian project to launch a space telescope aboard the shuttle in 1990), the design of the Parkes telescope and the new Australian Telescope.

Another example of this expertise was the construction in the 1960's of a unique solar radioheliograph (the only one of its kind in the world) at CSIRO's Culgoora Observatory in New South Wales. It consisted of 96 aerials set around a 10-kilometer circle, which performed as one huge dish 3 kilometers wide and photographed the sun

(continued on page A49)

FROM AUSTRALIA

a bionic ear...for deaf people who get no benefit from conventional hearing aids

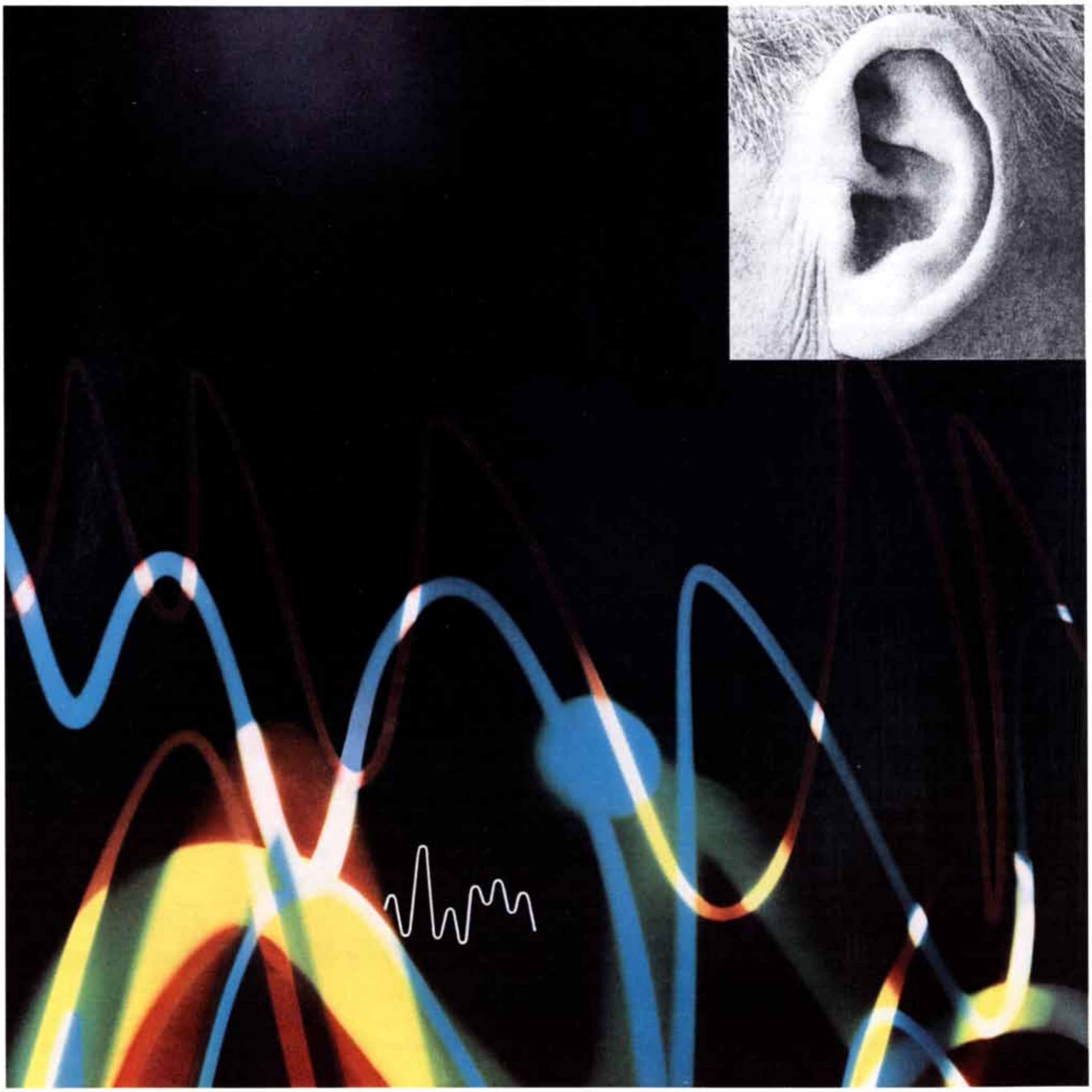
An Australian medical electronics company together with university researchers have developed a device called a cochlear implant. This consists of a set of electrodes which is inserted into the cochlea (inner ear). A small control unit connected to the electrodes is placed in the mastoid bone behind the ear. The patient wears an external speech processor similar to a body-worn hearing aid. This is connected to a small microphone/ transmitter unit worn behind the ear. Speech is picked up by the microphone. The speech processor extracts information from this and sends it to the implanted control unit by radio frequency transmission. Here the electrodes are activated to stimulate the hearing nerve to produce differing hearing sensations.

Not only is Australian high technology to the fore in medical and hospital use. It has also developed many innovative products and techniques for the scientific field, for astronomy, for industry, agriculture, mining and a host of wide-ranging general applications.

Find out just what Australia has available and how it can benefit you, by contacting the Australian Trade Commissioner.

Australian technology finds a better way

For information on Australian technology contact the Australian Trade Commissioner in your area. Australian Trade Commissioner offices are located in 42 countries



Technology Development Funding in Australia

by J. M. F. Keeney

One hundred years from now historians will be busy explaining our era. Of their many measurements of a society's vigor, the availability of venture capital will be an important one. Palo Alto, Boston and other such "neighborhoods" will be noted for the friendly interaction between financiers and technologists that spurred innovation. The correlation between venture capital and successfully commercialized technology will have been demonstrated by computer.

Europe has lately awoken to this equation. Sweden, Britain and France all boast of budding efforts—some of them government sponsored—that have supplied between $50 million and $100 million in each country. In comparison, Australia lags behind.

Australian bankers, made successful over many decades by investments in the several traditional sectors, are unaccustomed to dealing with fledgling technological enterprises. Many entrepreneurs complain of the banks' inability to understand technology and and of an obsession with "bricks and mortar" security.

The large corporations capable of amassing and channeling money into development have not shown the willingness of their European or American counterparts to diversify outside their industries, create specialized subsidiaries or actively engage in joint ventures. There is some evidence that the Australian entrepreneur does not enjoy the "folk hero" status one of his kind might in America; the stigma of failure is perhaps more daunting than in some other countries.

Such difficulties, until very recently, made the financial outlook for a young engineer with a good idea very dismal indeed. Even the chances of overseas cooperation were slight. Only two American venture-capital firms are known to have consistent interest in the Pacific region: Pacific Technology Venture of San Francisco and Boston-based Tucker and Anthony.

Now it appears change is under way. Although at present there is scarcely more venture capital in the country than a year ago, one can be guardedly optimistic about an improved state of affairs in 1984.

"Developing High Technology Enterprises for Australia," the most sweeping study of its kind, was prepared for the minister of science and technology and published in April, 1983. Informally known as the Espie report, after the study group's chairman, Sir Frank Espie, it devotes considerable attention to the plight of the average techno-entrepreneur in search of money, and to the paucity of venture-capital establishments.

Although the report does not quantify the extent of the problem, knowledgeable sources suggest that at the very least there are 100 qualified individuals with ideas, patents or prototypes in desperate need of finance. Venture capital available today is probably less than $5 million.

The Espie report concludes with a series of recommendations, which have been adopted in principle by the Hawke government. The report called for the creation of a Management and Investment Company Licensing Board to be made up of industry and business representatives with government involvement. The board will report to the minister of science and technology and issue licenses to venture-capital companies offering finance and management capabilities. Tax incentives will be generous to those licensed. The loss of revenues to government, estimated at $200 million over five years, should be more than offset by growth in the high-technology sector and the generation of taxable income.

A number of venture capitalists are now preparing to apply for licenses once the legislation establishing the scheme has been passed by Parliament. These include Business Loans and Equity Capital, which is 60 percent owned by Westpac and 40 percent by the U.K.-based Investors in Industry; Pratt and Co. Financial Services in

The $50 note portrays scientist Sir Ian Clunies-Ross, renowned for his work on parasites affecting livestock and his long association with the Australian Commonwealth Scientific and Industrial Research Organisation.

Melbourne; Citicorp Capital Investors, a branch of the U.S. company, and Western Pacific Management and Investment Co. of Melbourne. Terence Morgan, a Sydney-based merchant banker, is establishing a full-service bank that will, he states, "treat venture capital for high-technology enterprises as one of our priorities."

Among the most aggressive of venture capitalists is Pratt and Co., which intends to develop a portfolio of 10 enterprises including several in the high-risk, high-potential category. The major aim of the company is to retain control of the technology in Australia even if the product may have to be developed overseas. "We'll provide the missing ingredient whether it is finance or management skills or both," declares Mr. Robert Rees, one of the company's three directors.

The need for management guidance in addition to capital is a central concern of the Espie report. Dr. Craig Mudge, an electronics engineer who returned from the U.S. to direct a CSIRO-backed chip design facility, insists, "Australian financiers need to take a lesson from firms like Hambrecht and Quist in San Francisco, who employ engineers to help manage their portfolio companies."

The Australian Industry Development Corporation (AIDC), a statutory body recently revamped by the government, will soon be able to contribute to both management and financial solutions. The government has increased the corporation's lending powers substantially by lifting its authorized capital from $100 million to $150 million. Its gearing ratio has been increased so that it will be able to borrow up to 15 times the level of its capital and reserves. This will result in a total borrowing capacity of 1,700 million.

The AIDC has formerly concentrated on large capital-intensive projects such as mining and oil exploration. It will soon formalize a program to assist smaller, high-technology firms. Senior executive Barry Hilson describes the effort as "expanding our perspectives to include smaller loans than we have previously made and organizing an enlarged staff to deal with the needs of the small company."

One of the as yet unanswered challenges is the creation of an over-the-counter share market, which would allow people to realize their investments early. At present it may be difficult for investors to extricate themselves from a venture company in the short term. There are also fears that the limits on investment placed on the venture companies will mean that worthwhile higher-risk enterprises will remain unfunded and that without careful management the new scheme will become a tax haven.

While the Espie report has galvanized opinion and offered coherent solutions, a good deal remains to be done. For the pioneers in Australia's venture-capital business, 1984 will provide a significant testing ground. Decades from now it may be regarded as a benchmark year in Australia's efforts to retain control of its own inventions.

(continued from page A46)

once a second, providing a continuous record of the rapid outbursts of solar radiation. This could be played back as a fast-motion movie film. Unfortunately, CSIRO had to close the observatory due to budget funding cuts.

Mining and Mineral Processing

Australian-developed mining equipment and mineral processing techniques are now in use throughout the world, from the remotest parts of Africa to the heart of the U.S.. Many of the developments have revolutionized resource exploration and extraction methods, allowing much greater quantities of ore to be extracted, cutting costs, improving safety and increasing the economics of many projects.

One of the first innovations to have widespread impact around the world was the development of the flotation process for the separation of minerals, first patented in 1901 by a Melbourne brewer and chemist, Charles Potter. It was soon in use at Broken Hill, and by 1912 a selective flotation process for the separation of lead from zinc had been developed. It was further refined and is now the most widely used method in the world for mineral extraction and is used for the concentration of practically every mineral mined.

Since 1959 much of the work has been supported by the Australian Mineral Industries Research Association (AMIRA), which has a membership of 70 resource companies. AMIRA initiates and coordinates jointly sponsored R&D, which it contracts out to groups such as CSIRO.

One major advance of the 1960's was the development of a new system of on-stream analysis for measuring the different grades of ore in the raw mined material as it passed through the processing plant. The Australian Atomic Energy Commission pioneered the work, developing radioisotope X-ray and gamma-ray probes, first to determine the lead content in process streams and later to pinpoint a range of metals, including zinc, tin, copper, nickel and iron. The technique was later taken up by AMIRA in conjunction with the Australian Mineral Development Laboratory and the electronics group Philips, and a range of radioisotope immersion probes were developed to measure the levels of various metals in slurries.

Automatic computerized process control of mineral extraction now in use throughout the world has also been pioneered in Australia by MIM Holdings at its Mount Isa mine. Crushing, grinding and flotation of ores is computer controlled, which greatly improves the efficiency of processing, increases the level of metals recovered from the ore and allows the grade of mineral concentrate produced to be predicted.

One of CSIRO's successful inventions has been Sirotem—a computerized direct digital borehole logging system and data processing center, which can detect ore bodies covered by thick layers of conductive material.

The portable system was developed in the early 1970's by CSIRO's Division of Mineral Physics and is now in use in many parts of the world, including Scandinavia, Canada and China.

Another successful CSIRO development, QEM SEM (quantitative evaluation of minerals by scanning electron microscopy), is a fast, comprehensive mineral analysis system that can scan an ore sample and identify the number of minerals present, their types, quantities and position.

Others include Sirosmelt, a direct lead smelting technique, which cuts energy costs and increases metal recovery; Siroash, a gauge for on-line determination of ash in coal; Sirolog, a series of nuclear probes for accurately determining the grade of ores and the way minerals are arranged in rock, and Stratosnoop, for measuring the thickness of mudstone or shale in underground coal mines.

Australia is at the forefront of world soil and rock mechanics research and development, particularly in the design of open-cut mines. Techniques for the mining of very large openings, digital monitoring systems to detect slippage in high-wall or spoil heaps, new methods of underground mining and explosive technology and innovative ways of dealing with spoil heaps, tail-

(continued on page A51)

Sunrise Industries and Technological Autonomy

by Hon. Barry O. Jones, Minister for Science and Technology

Australia is an urbanized, technologically sophisticated nation with a strong agricultural base, rich mineral resources, a historic record of developing its intellectual capacity and substantial achievements in agricultural and medical research. The challenge to be faced in the 1980's and 1990's is this: can a nation with only 15 million people, a mere 4 percent of the English-speaking world, whose high-technology industries are overwhelmingly under foreign control, make a transition toward the newly developing brain-based high-technology ("sunrise") industries as wealth generators, to compensate for the long-term decline in employment in traditional manufacturing industries? Can new technologies be used to regenerate smokestack industries?

In the late 19th century Australia began to adopt the colonial model of technology transfer, exchanging gold, wheat and wool for technological artifacts and know-how. There were important exceptions: Australia was a world leader in agricultural machinery, refrigerated ships and building and mining techniques. We had some of the first electrified cities, electric cables and telephones. Before 1950 Australia produced pharmaceuticals, automobiles, aircraft, electronic equipment and one of the earliest stored-memory computers (CSIRAC). However, under the long Menzies era, foreign corporations were encouraged to establish themselves in Australia, often taking over locally owned companies.

The commanding heights of the Australian economy are in foreign hands, including motor manufacturing, oil, computers, electronics, chemicals, plastics and pharmaceuticals. This leads to a major disincentive to investment in research and development by heavy industry in Australia, with a few notable exceptions.

Australian industry and society face massive threats, but also great opportunities. We can be locked into a global economy in a dependent role, with diminishing power to maintain our technological autonomy, or we can develop our own high-technology industries and attempt to penetrate niches in the world markets. The revolutionary productive techniques of the 1970's, if they are adopted wholeheartedly, could overturn much of the conventional economic wisdom about economies of scale, which have excluded many Australian products from the world markets. On the other hand, rapid adoption of the new technologies in existing industries could cause a massive loss of jobs, so that consideration must be given to choosing technologies selectively and distinguishing between "old" and "new" industries.

As a major international trader in raw materials Australia has been adversely affected by the shift in Western economies from a resources base to a skills base. We can no longer assume with the confidence we had in the 1960's that minerals, wheat and wool alone will carry "the lucky country" toward higher living standards. We have to diversify and extend our skills base and our international market commodities.

Since 1965 Australia has had a post-industrial economy. Virtually all new jobs created since then have been in "services." As early as 1900 a majority of Australia's labor force worked in services and the figure is now more than 75 percent. A four-sector labor force analysis showed that more Australians now work in "information" than in farming, mining and manufacturing combined (27 percent).

To complete the transition to an *active* information society, Australia will strengthen its knowledge base and technological skills. There is already a strong foundation to build on. The Commonwealth Scientific and Industrial Research Organisation (CSIRO) was founded in 1926, employs 7,500 people and has an international reputation for research in agriculture, animal and plant genetics, mining, earth sciences and radio astronomy.

Australia has 19 universities and 22 technologically based colleges of advanced education, and many departments have international reputations for research. So have the Walter and Eliza Hall Institute, the Howard Florey Institute, the Baker Institute and the Garvan Institute in medical research, where Australia developed an early expertise, probably because of the tyranny of "distance."

Australia, with .3 percent of the world's population, produces 2.0 percent of the world's scientific papers each year. With one-eighth the population of Japan, Australia has produced the same number of Nobel prizewinners in the sciences (Florey, Burnet, Eccles and Cornforth). However, past policies failed to encourage commercial exploitation of scientific discovery. Many managers have preferred to acquire licenses to foreign products and assemble them here behind high tariffs rather than develop our own skills in invention and product innovation and aim at an export market. We had the ideas, but until now we were lacking in product development. The greatest problem in technology transfer was essentially *domestic*.

The Labor government elected in March, 1983 came to office with a detailed recovery and reconstruction policy aimed at promoting Australia's technological autonomy. The science and technology policy identified 16 "sunrise" industries—all of them where Australian research is at the leading edge or where we are appropriately placed to exploit significant markets in our region. The list of 16 leads off with biotechnology, in which Australia has a very strong skill base in the universities and CSIRO. There are a number of major developments in biotechnology that seem likely to penetrate world markets: much of our research funding comes from abroad and the final stages of product development may be carried out offshore. Three microelectronics-based "sunrise" industries were nominated: computer software, custom-made VLSI chips and personal computers, although local developments are occurring so rapidly that it now seems unwise to limit the list to three. CSIRO's industrial ceramic PSZ (partially stabilized zirconia) performs better than metal alloys and steel under heat and stress. Solar technologies are well advanced in Australia, and we have significant exports in scientific instrumentation and medical technologies.

These and other "sunrise" industries will not be major employers in themselves, but we expect them to be major wealth generators, like mining, that will generate significant second- and third-order effects in employment.

In the 1983-84 budget the government provided greatly increased funds for research in the key areas of biotechnology, genetic technologies and advanced industrial materials. The government has also taken steps to create a venture-capital market to encourage the creation of new industries.

We are changing direction, and there are good prospects that our development of human skills will generate higher levels of economic growth than could be produced by physical resources alone.

ing dams and general reclamation of mined land have been developed.

Research is under way in a number of other areas such as slurry pipelining of coal, new gold recovery techniques, the use of remote sensing for mineral exploration, coal hydrogenation and pyrolysis, new hydrometallurgical and electrometallurgical processing techniques, coal outbursts and drainage, improved control of sintering, the stability of underground and open-cut metal and coal mines and the use of aquatic plants to purify mining wastes.

Conclusion

Australia, as it enters 1984, is at a crossroads, faced with far-reaching decisions on its economic and technological future. Should it continue as a resource-based economy dependent on income generated by its mineral and rural exports or should it take a new course and establish a skill-based economy founded on new high-growth, high-technology enterprises?

Opinion is mixed and debate widespread on whether the federal government should back the establishment of new "sunrise" industries or continue with its high protection policies for the declining traditional manufacturing sector.

Mr. Barry Jones, minister for science and technology, is leading a strong push into the new technology field arguing that unless Australia moves rapidly to adopt new technology and set up new industries its economic base will contract, its skill levels deteriorate and its standard of living will further decline.

He maintains that Australia has already entered a postindustrial society and that it must reject all hopes of what he describes as "nostalgia-led or obsolescence-led recovery" based on traditional manufacturing.

He says that Australia has already undergone a paradigm shift in its economic base and is experiencing much more than an industrial decline that can be fixed with temporary support. A new type of society based on a different economic base—knowledge and skills rather than raw materials and muscle power—must be developed.

He believes Australia must rapidly establish new high-technology industries as wealth generators to compensate for the decline in employment in the sunset of smoke stack industries.

But this view is not shared by all and many of the more conservative elements of industry, business, the financial community and government argue that Australia should continue to concentrate on more traditional areas, such as mining and agriculture, and to prop up its declining manufacturing sector with more protective tariffs. Their only concession is that perhaps new technology should be introduced to some of these industries.

They argue strongly against giving extra support to the new emerging industries, believing that this should be left to market forces.

But, under the strong influence of Mr. Jones, government opinion is shifting more toward the support of new technology, as seen by the recent decision to introduce a new tax-incentives scheme to boost a venture-capital market in Australia and moves to direct more money into high technology.

But government, business and community attitudes have a long way to go before they become aligned with those of Mr. Jones.

Management is still reluctant to invest in research and development or in new technology, financiers are wary of taking risks and backing new enterprises and unions are opposed in many cases to the introduction of new technology. However, attitudes are changing and unions are becoming more willing to negotiate with employers on the introduction of new technology, recognizing that it represents a choice between adopting new practices and company closure. Management and financiers are also slowly sensing the possibilities.

Australia also has other major problems to overcome in developing a technology-based economy, in particular its small domestic market, its distance from export markets and the predominance of foreign-owned companies in Australia.

The latter has led to increasing dependence by Australia on overseas technology as well as to the low level of private-sector industrial research and development, with multinationals preferring to carry out their research in their home countries.

This lack of technological sovereignty is a major concern of the government, as is the loss of much Australian-developed technology overseas due to the reluctance of Australian industry to take the risk to develop it locally. Another concern is the need to adapt Australia's educational structure to a skills-based economy, with more leisure, less people employed and a large sector of the work force concentrated in service and information industries.

Mr. Jones believes there will have to be radical changes in attitudes to the work ethic and total redefinition of work and the ways it is organized.

He admits that new technology will reduce employment but believes that Australia should start looking at alternatives such as early retirement, recurrent education, a diversity of lifestyles based on individual autonomy, a guaranteed income scheme and a redefinition of work to include domestic work and hobby activities.

He sees the first step as the "shock of recognition" that technological change is under way, that Australia is entering a new economic era and that new ideas and attitudes must be adopted. How successful Mr. Jones and the new government will be in moves to transform Australia from a resource to skill-based economy is still debatable—but it is sure that the next few years will be among the most interesting for science and technology in Australia since the early pioneering days.

As Prime Minister Hawke said at the recent national technology conference, "We have in this country the intellectual wherewithal to adopt a new attitude that allows us to make better use of new technology for our national purposes.

"We have good schools and universities, stable government and the obvious potential to build a great future, based on the development and appreciation of innovative industrial structures.

"Some will say to this that Australia is not yet ready, that the cost of adjustment is unacceptably high, that protection must be provided against the growing power of overseas competition. To them I simply say 'I hear you, but cannot agree.' "

We get our best ideas while America and Europe sleep

. . . in Australia our perspectives are different

The half of the universe that nightly wheels above our heads is largely hidden from northern eyes. We've found some remarkable things there. Soon we'll have a new radio telescope array to probe even deeper into the southern sky. Its computers will employ some specialized chips we designed ourselves. We gave Australian industry the same custom-design capability.

Resourcefulness. The Australian environment demands it.

Our landscapes dazzled LANDSAT with their brightness. We coped by devising a superior image enhancement system and sold it to the country that made the satellites.

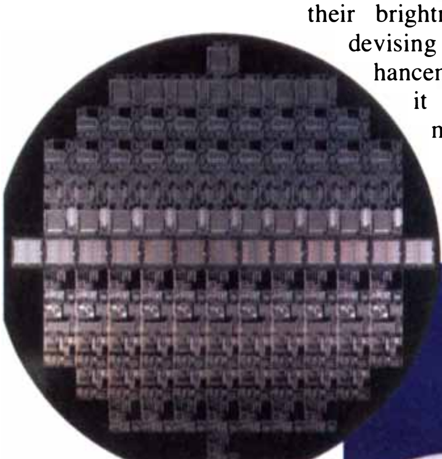
Our minerals are no longer easy to find. Today we're hunting for new deposits with a supremely sensitive device that detects their electromagnetic signature beneath the veil of salt in our ancient soils.

Our scientists invented the world's toughest ceramic, PSZ. You'll find it in tomorrow's super-efficient diesel engines and a range of other industrial products.

In agriculture, we've bred plants that can cope with our poor soils and erratic climate. The magic of somacloning, and direct gene insertion using transposable DNA elements, are just two trails our molecular biologists are exploring in the quest for tomorrow's crops.

We've devised ion-exchange water purifiers. A low-energy smelting system. A new generation of safe, computer-designed insecticides. A powerful mini-node for computer networking. And advances in a hundred other fields.

No giant steps. Just modest leaps.



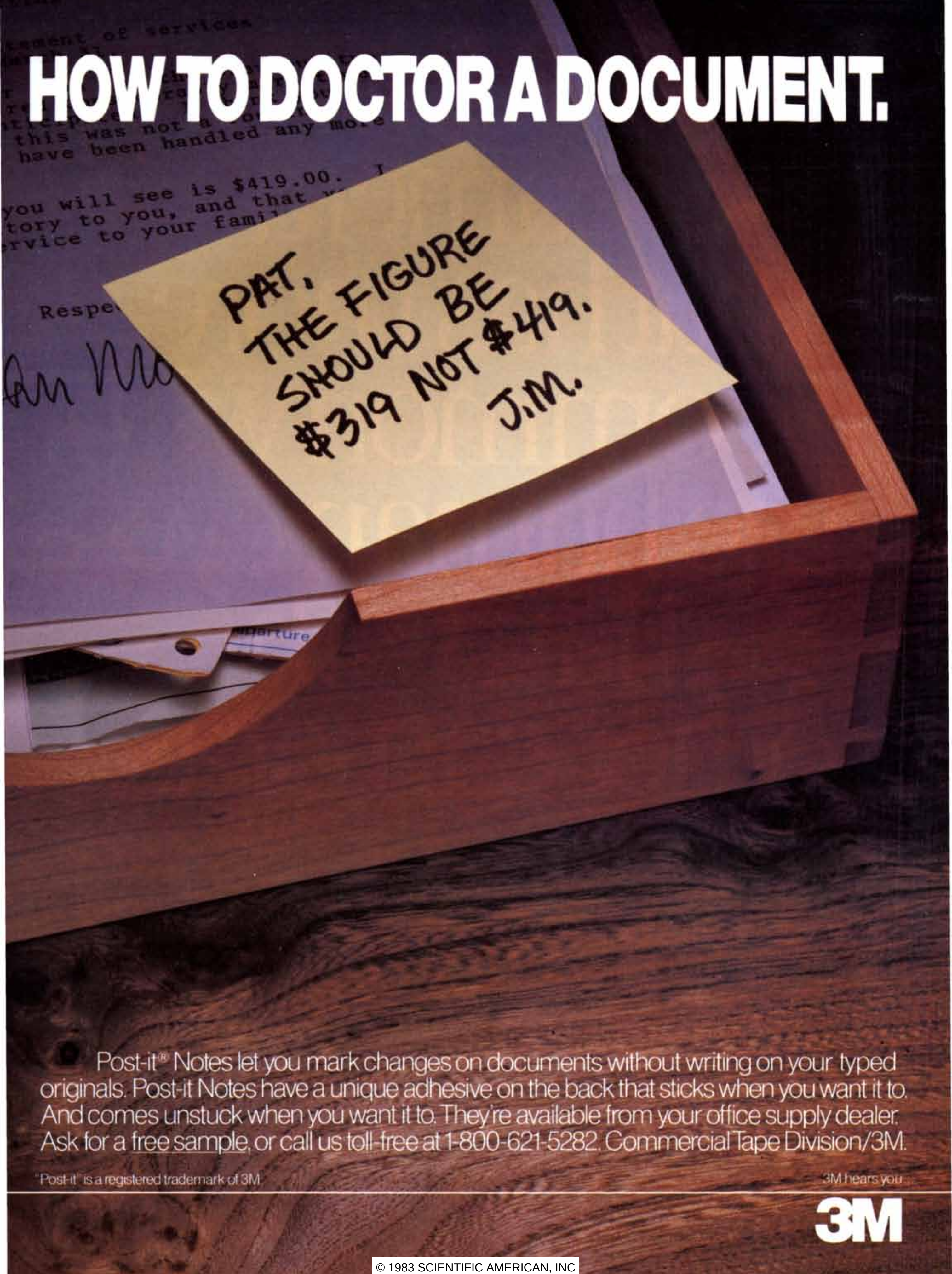
Australia's first multi-project silicon wafer.

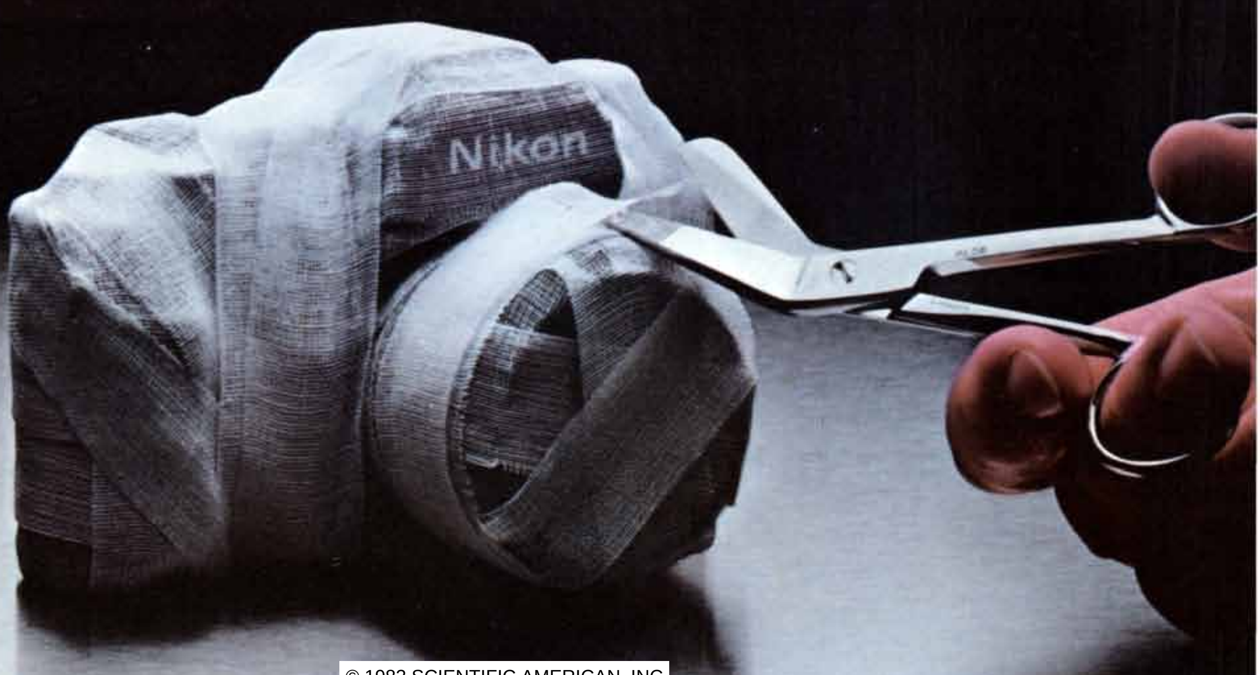
Components made from PSZ, the world's toughest ceramic.

CSIRO

Commonwealth Scientific and Industrial Research Organization

CSIRO Chief Executive:
Dr J. Paul Wild
Postal address:
PO Box 225, Dickson, A.C.T.
2602, AUSTRALIA.
Telephone: (062) 48 4132


HOW TO DOCTOR A DOCUMENT.
PAT,
THE FIGURE
SHOULD BE
$319 NOT $419.
J.M.
Post-it® Notes let you mark changes on documents without writing on your typed originals. Post-it Notes have a unique adhesive on the back that sticks when you want it to. And comes unstuck when you want it to. They're available from your office supply dealer. Ask for a free sample, or call us toll-free at 1-800-621-5282. Commercial Tape Division/3M.
"Post-it" is a registered trademark of 3M.
3M hears you...
3M


Nikon

High contrast. Shot in Aperture-Priority with the Nikon FA.

Back light. Shot in Programmed with the Nikon FA.

Available light. Shot in Programmed with the Nikon FA.

Sun-in-frame. Shot in Shutter-Priority with the Nikon FA.

Shot in Aperture-Priority with a leading multi-mode SLR camera.

Shot in Programmed with a leading programmed SLR camera.

Shot in Programmed with the biggest selling SLR camera.

Shot in Shutter-Priority with a competitor's "top-of-the-line" SLR camera.

The pictures you see here are actual unretouched photographs, shot simultaneously without any exposure compensation. ©Nikon Inc. 1984.

See.

The Nikon FA. The biggest advance in automatic photography since automatic exposure.

Until now, the metering system of any automatic camera could do just one thing. Measure light and give you a technically correct exposure.

But as any photographer knows, a technically correct exposure doesn't always give you the best picture.

That's why Nikon developed the FA. The first camera with AMP (Automatic Multi-Pattern) metering.

AMP is the only metering system that can automatically give you optimum exposure, not just technically correct exposure, even under extreme lighting conditions.

So what you see in your pictures is a lot more like what you saw with your eyes.

How AMP works.

AMP metering divides your picture into five segments and then individually measures and compares each segment, evaluating such factors as contrast ratios, variations in brightness levels and percentages of light and dark areas.

It then processes this information in its own Nikon microcomputer, comparing the components of your picture with those of nearly 100,000 photographs programmed into its memory, and instantly chooses the optimum exposure.

The FA gives you more choices than any other camera.

Shoot in the Dual-Program mode and the camera does it all for you. With one program for normal and wide-angle lenses and a high-speed program for Nikon AI-S and Series E lenses, 135mm and longer.

Or switch to Shutter-Priority.

With the FA's top shutter speed of 1/4000 of a second, there's not much you can't catch.

If you're most concerned about controlling the sharpness of foreground and background, Aperture-Priority is at your command.

And of course, you can also take full creative control in Manual.

Add other Nikon options, too.

When you shoot with the FA, you can take advantage of the most advanced photographic system in the world.

Use a Nikon motor drive and shoot up to 3.2 frames-per-second.

Or attach a variety of Nikon Speedlights to activate the FA's automatic TTL (through-the-lens) flash metering system, and shoot flash pictures at sync-speeds up to 1/250 of a second.

The FA is also compatible with all current and many older Nikon lenses, and a full range of Nikon accessories.

To find out more about the kind of pictures the FA can take, write to Nikon Inc., Dept. 55, 623 Stewart Ave., Garden City, N.Y. 11530.

Or better yet, just use your eyes.

Nikon
We take the world's greatest pictures.™

Rice

This member of the grass family, along with wheat and maize, is one of the three on which the human species largely subsists. Advances in the genetics of the plant have greatly increased its yields per acre

by M. S. Swaminathan

Among the major grain crops the only one that is grown almost exclusively as human food is rice. Indeed, rice constitutes half of the diet of 1.6 billion people, and another 400 million rely on it for between a fourth and half of their diet. As one might expect, the cultivation of the crop takes up a great deal of land: some 145 million hectares (358 million acres), amounting to 11 percent of the world's arable land. In the year ending July 31, 1982, the harvest amounted to 410.9 million metric tons. It would have been far less than that, and millions of people would have been severely underfed, if it had not been for a series of remarkable genetic advances that have brought under cultivation high-yield varieties, resistant to disease and insect pests.

Rice is an annual grass belonging to the same family as barley, oats, rye and wheat and sharing many of their characteristics. The rice genus is *Oryza*. Its origins go so far back into antiquity that they will probably never be traced with certainty. My colleague T. T. Chang at the International Rice Research Institute in the Philippines has speculated that the original habitat may have been the ancient supercontinent of Gondwana. When Gondwana broke up and became Africa, Antarctica, Australia, Malagasy, South America and South and Southeast Asia, *Oryza* species drifted into distinct geographic habitats.

The two cultivated rices, each with a great number of varieties, are *O. sativa* of Asia and *O. glaberrima* of West Africa. *Oryza* also includes 20 wild species, which are scattered in Asia, Africa, Australia and Central and South America. Archaeological evidence suggests that the domestication of *O. sativa* began in Asia more than 7,000 years ago and that *O. glaberrima* was domesticated in Africa somewhat later.

Rice grows in a diversity of environments almost unparalleled in the plant kingdom. It originated in the hot, humid Tropics, where monsoon rains and floodwaters create an aquatic environment for at least part of the year. Natural dispersal and human selection have extended the cultivation of rice from the banks of the Amur River (53 degrees north latitude) on the border between the U.S.S.R. and China to central Argentina (40 degrees south latitude). Rice is cultivated in cool climates high in the mountains of Nepal and India and in the hot deserts of Pakistan, Iran and Egypt. It is grown as a dry-land crop in parts of Asia, Africa and Latin America. At the other extreme of cultivation are floating rices, which thrive in floodwaters three meters deep in parts of Bangladesh, Burma, eastern India, Thailand and Vietnam. Rice adapts well to diverse growing conditions and performs better than other grain crops in areas with unfavorable saline, alkali and acid sulfate soils. On the basis of this characteristic the Food and Agriculture Organization has predicted (in *Land Resources for Populations of the Future*) that in coming decades rice culture will expand in several countries.

Methods of growing rice range from highly mechanized systems in the U.S. to labor-intensive cultivation in many parts of Southeast Asia. Average rice yields vary from less than one ton per hectare in some African countries to more than six tons in Australia, Japan, North and South Korea and the U.S. In tropical Asia the yield for centuries was from one ton to 1.5 tons per hectare. From 1960 to 1980, however, average yields in that area increased 40 percent and production increased more than 60 percent. Countries such as Indonesia and the Philippines increased rice production and productivity more during those 20 years than in the preceding 7,000. Indonesia is now nearly self-sufficient in rice. The Philippines, which were a major rice importer until a few years ago, now frequently export more than 100,000 tons per year.

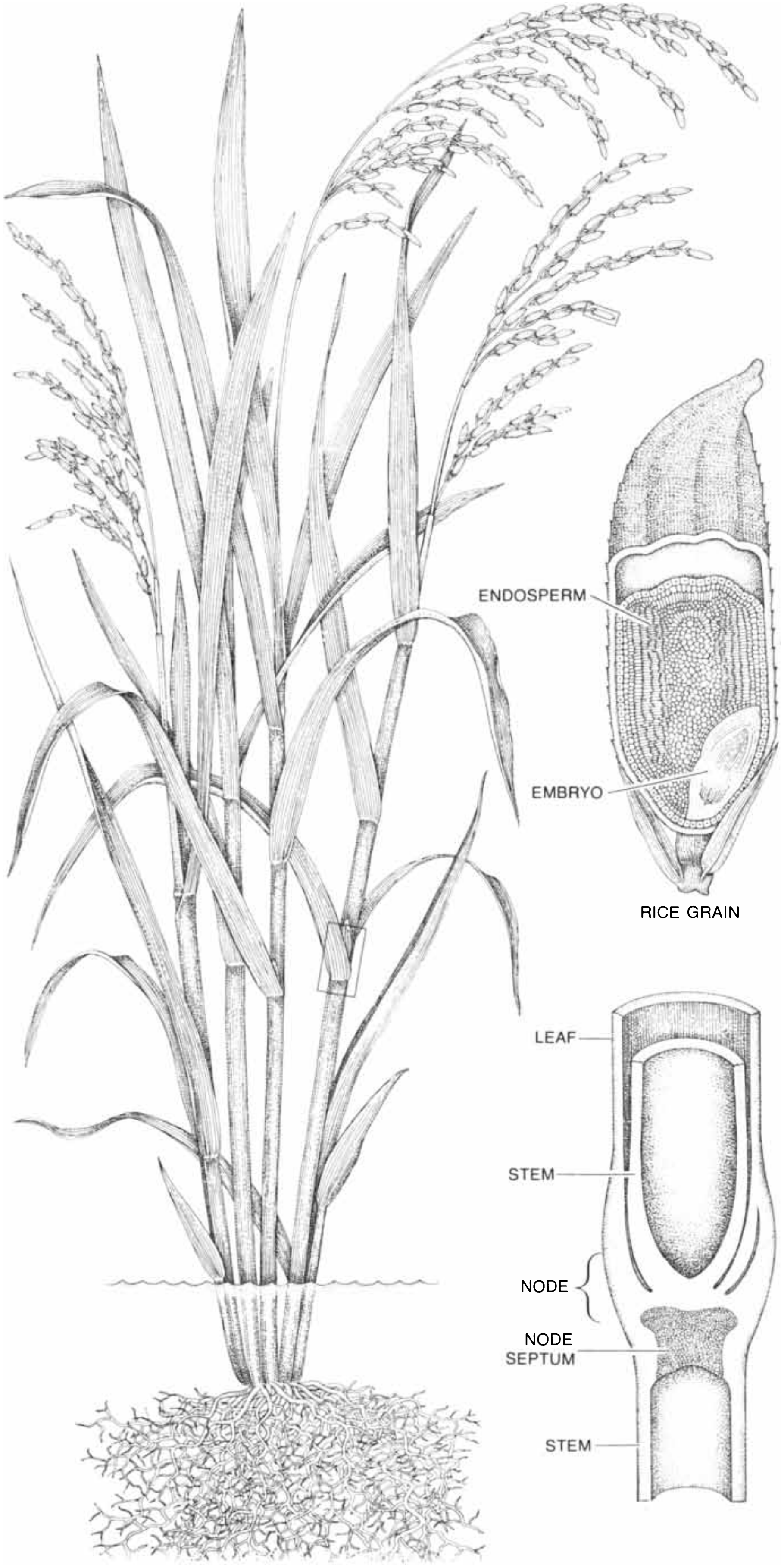
How did such a transformation take place? The primary catalyst of change was the development of high-yielding varieties of rice, supported by a trend toward scientific management of the soil and of the health of the plants. The governments of several countries in Southeast Asia introduced complementary policies to speed irrigation, land reform and such inputs as fertilizer, along with liberal credit arrangements and price supports. The result was a rise in production in most of the countries exceeding 5 percent per year.

Asian Rice

Over the millenniums the cultivated species of Asia (*O. sativa*) differentiated into three subspecies based on geographic conditions. They are indica, japonica (also called sinica) and javanica. A further classification, which emphasizes the habitat in terms of soil and water, is rain-fed lowland and upland rice and irrigated and deep-water rice. Indica rices were originally confined to the humid regions of the Asian Tropics and subtropics. Japonicas were cultivated in subtropical Temperate Zone regions. Javanicas flourished in the equatorial region of Indonesia.

In addition to their adaptation to climate the three races differ in characteristics of the grain, including the content of amylose (a derivative of starch), the elongation of the grain, the temperature at which the grain becomes gelatinous and the aroma in cooking. The cooking and eating qualities of rice are largely determined by the properties of the starch, which constitutes 90 percent of milled rice. Grain quality also must be

TERRACING FOR RICE greatly extends the area available for cultivation in mountainous regions of Asia. The terraces in the photograph on the opposite page are in the Philippines. Rice can be seen in several stages of cultivation in paddies holding from five to 10 centimeters (two to four inches) of water. The rice plant is able to thrive in water because of a system of air passages that carries oxygen to the roots. These diked terraces are filled by rainwater.



RICE PLANT is portrayed, together with a section of the tiller, or stem and leaf, and a grain. The rice genus is *Oryza;* this plant is *O. sativa,* the cultivated Asian species, of which there are some 120,000 varieties. The other cultivated line of rice is *O. glaberrima,* an African species.

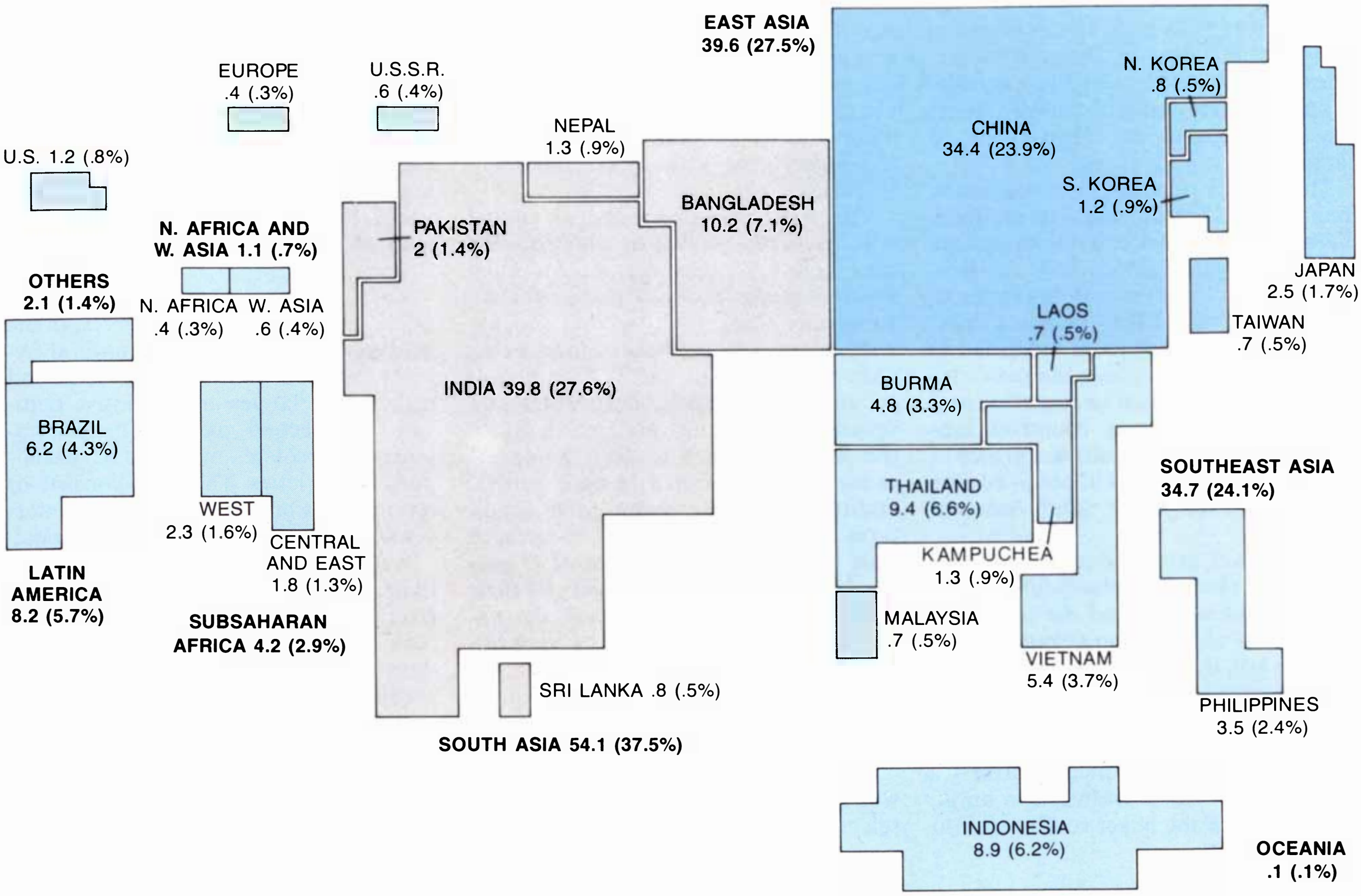
defined in relation to place. For example, to the people of Thailand rice is of good quality if it has long, slender, translucent grains that yield a fluffy and tender product when they are cooked. In Japan the preference is for rices with short, broad grains and sticky cooking properties that make eating with chopsticks easy. In parts of Pakistan and northwestern India premium prices are fetched by *basmati* varieties with long, slender, scented grains.

In tracing the spread of the Asian cultivated species *O. sativa* most investigators agree that the area with the greatest diversity of varieties is a belt extending from the Assam-Meghalaya region of India to the mountain ranges of Southeast Asia and southwestern China. Varieties that matured early and therefore escaped the most serious effects of periodic drought probably emerged between about 15,000 and 10,000 years ago along the southern and northern slopes of the Himalayas. Annual ancestral forms of *O. sativa* began to appear then at the periphery of the wild annual progenitors, mainly in the southern reaches of the Himalayas and to a lesser extent in southern and southwestern China. Alternating periods of drought and pronounced temperature variation accelerated the development of the annual forms of *O. sativa* in northeastern and eastern India, northern Southeast Asia and southern China. All the early cultigens were indicas.

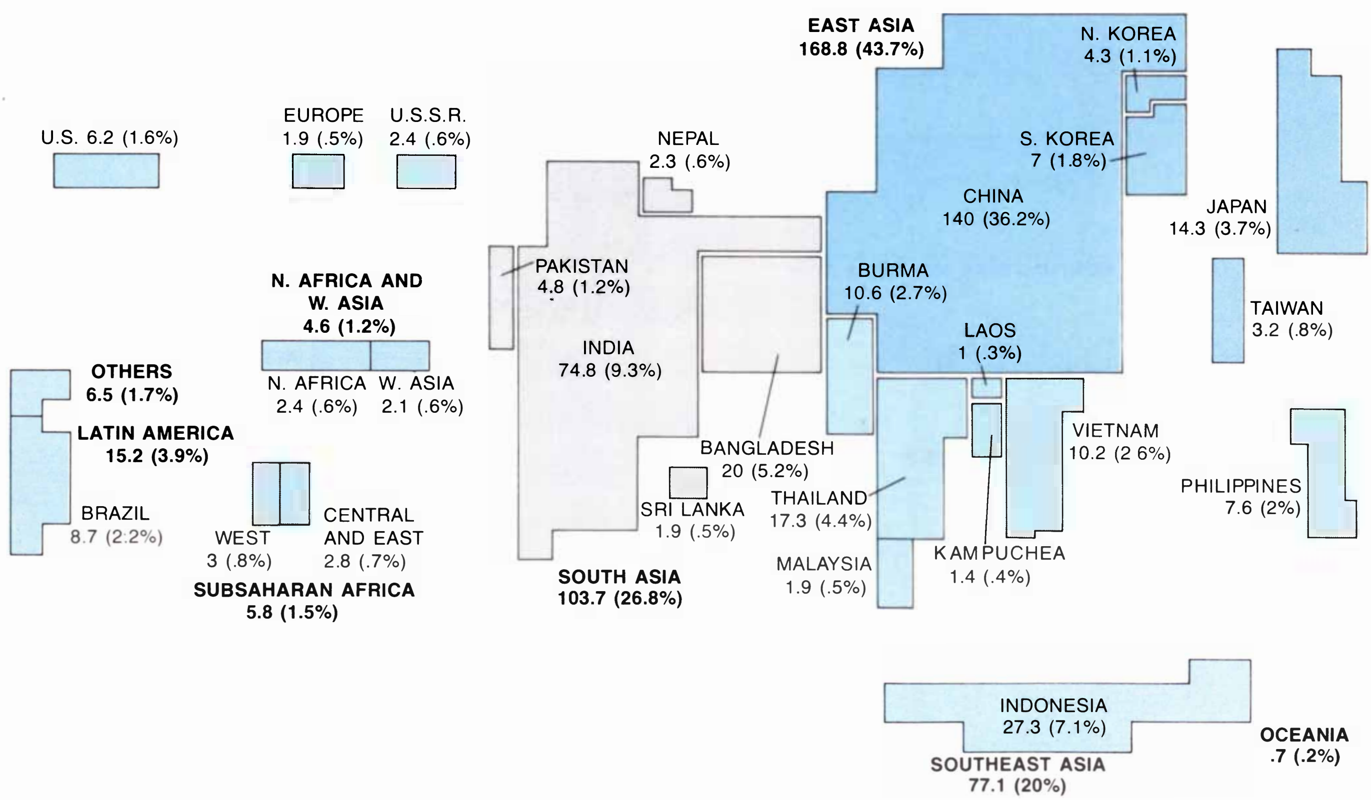
Cultivated varieties matured earlier than their progenitors, and so they survived better in drier, fluctuating climates and produced more seed. These factors helped them to move farther north than perennial forms. Increasing aridity on the northern border of the Himalayas and their associated mountain ranges in mainland Southeast Asia forced early settlers in China to move east and south toward more humid areas. As people migrated, plants also dispersed, accelerating the process of ecogenetic diversification.

Inside the present boundaries of China, along a branch of the Brahmaputra River (Hsiangchuna River), a Temperate Zone race of rice evolved from the primary tropical indica race. This race has been widely known as japonica since Japanese workers coined the term in 1928. Because the Japanese people obtained their rices from China, however, Chang argues that japonica rices should really be called sinica (keng) rice. Tall, large-grained rices migrated to the Indonesian islands from the east coast of South Asia and differentiated to become javanica rice.

The oldest remains of cultivated rice, both indicas, were found in eastern China and northern India and date back 7,000 years. The oldest sinica rice found in China is dated at 3300 B.C. Grains of what probably was wild rice have been



LAND DEVOTED TO RICE throughout the world is charted in a schematic map, wherein the space for a country is roughly proportional to the amount of land (shown in millions of hectares) planted to rice. Also shown is each country's percentage of the world total.



PRODUCTION OF RICE is charted schematically in terms of millions of metric tons. Each country's contribution to the world total is also shown as a percentage. The production measured in the tally is paddy, that is, rice grains that have been threshed but not milled.

excavated at Non Nok Tha in northern Thailand and date to about 4000 B.C. The antiquity of rice cultivation is widely known in Asia; rice farmers are sometimes referred to as "farmers of 50 centuries."

The later dispersal of Asian cultivars, mainly indicas, from India to the Middle East, North Africa and then Europe probably began in about 1000 B.C. Rice also traveled from India to Madagascar and East Africa. Many javanica cultivars found their way from Indonesia to Madagascar. West Africa obtained rice varieties from Europe or directly from South Asia. European countries supplied most of the first cultivars grown in South America. The U.S. obtained rice seed from Madagascar, South Asia and East Asia.

I have said little about the cultivated African rices (*O. glaberrima*) because they have never attained the economic importance of the Asian group (*O. sativa*). The late R. Porteres of the Institute for Agronomic Research in Paris postulated that the African cultigen *O. glaberrima* originated in the delta of the Niger River. The primary center of diversity for *O. glaberrima* is from the swampy upper basin of the Niger southwest into two areas near the Guinean coast. Rice was probably cultivated in the Niger basin in about 1500 B.C. and in Guinea about 400 years later. In West Africa *O. glaberrima* is a dominant crop grown in flooded areas of the Niger and Sokoto basins.

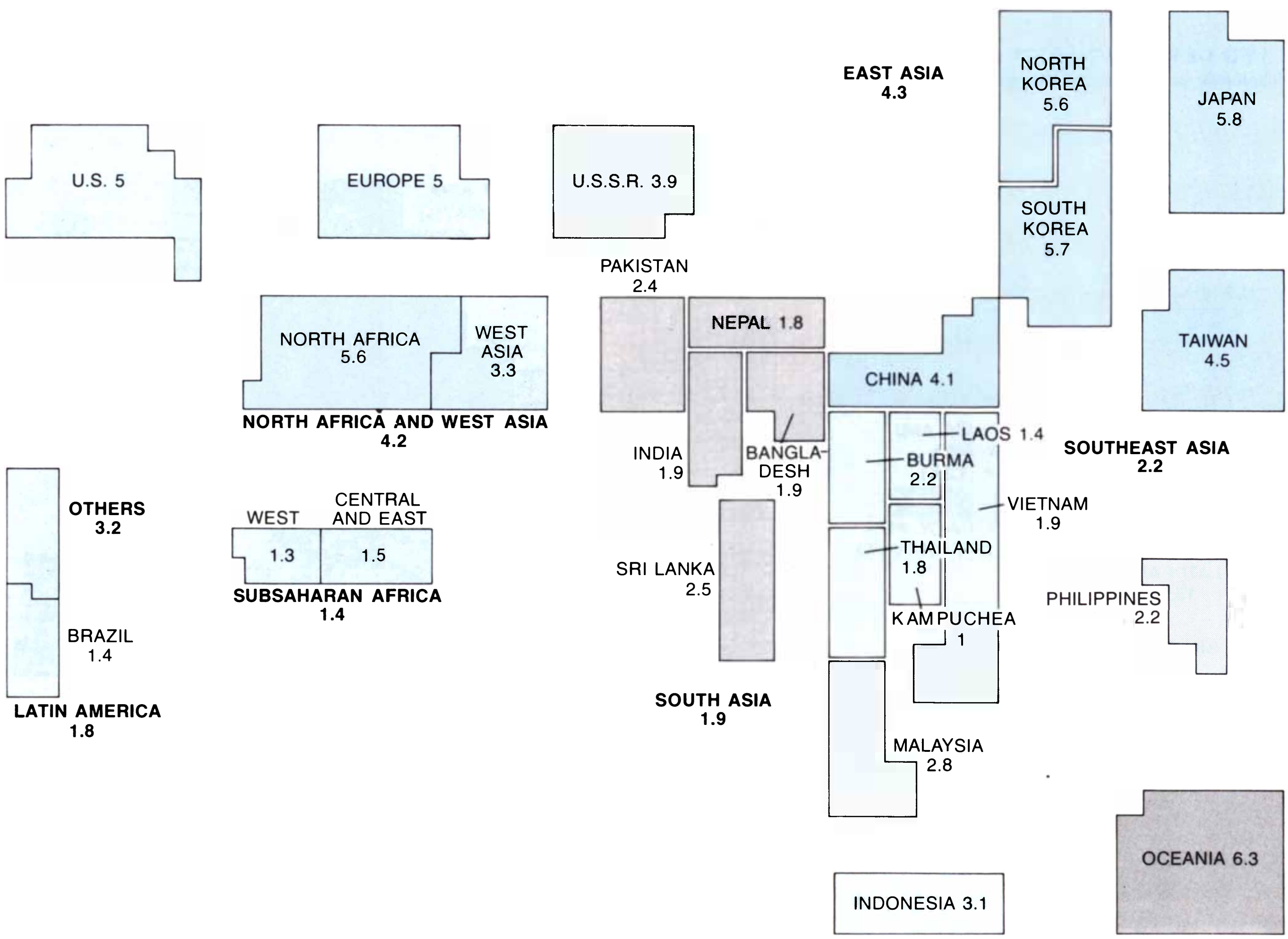
The African cultigen and its annual wild race (*O. barthii*) are less diverse than their Asian counterparts. Although Porteres recognized in *O. glaberrima* two subspecies designated *vulgaris* and *humilis,* other workers believe the *barthii-glaberrima* complex can be differentiated into two ecotypes: deep-water and upland. African cultivars are inferior to the Asian cultivars in deep water because they are limited in their genetic ability to elongate under such conditions. The differentiation of *O. sativa* in Asia probably predated that of *O. glaberrima* in West Africa, judging by their history of domestication and the extent of the diversification of varieties within species.

Preserving the Heritage

The ecogenetic differentiation of rice was achieved primarily through natural selection and human selection in different environments following hybridization, segregation or mutation. The many races of rice thus developed constitute a valuable pool of germ plasm. They must be collected, conserved and classified for current and future breeding programs. To this end the International Rice Research Institute has organized the International Rice Germ Plasm Center.

The full spectrum of germ plasm in *Oryza* derives from three sources. In the first group are wild species, natural hybrids between cultivated varieties and their wild relatives and primitive cultivars. The second group includes commercial, obsolete, minor and special-purpose varieties. The third consists of the varieties arising from human intervention, either unwitting or intentional.

It is estimated that the number of cultivars of rice may approach 120,000. Every major rice-growing country in Asia assembled germ plasm from substantial numbers of indigenous rices starting in the 1930's. The size of the national collections ranges from several hundred varieties in Laos to about 40,000 in China. India has assembled some 25,000 varieties and the U.S. maintains more than 7,000.



YIELD OF RICE is charted in metric tons per hectare. This chart and the two preceding ones are based on work by Adelita C. Palacpac of the International Rice Research Institute with data from the U.S. Department of Agriculture and the Food and Agriculture Organization.

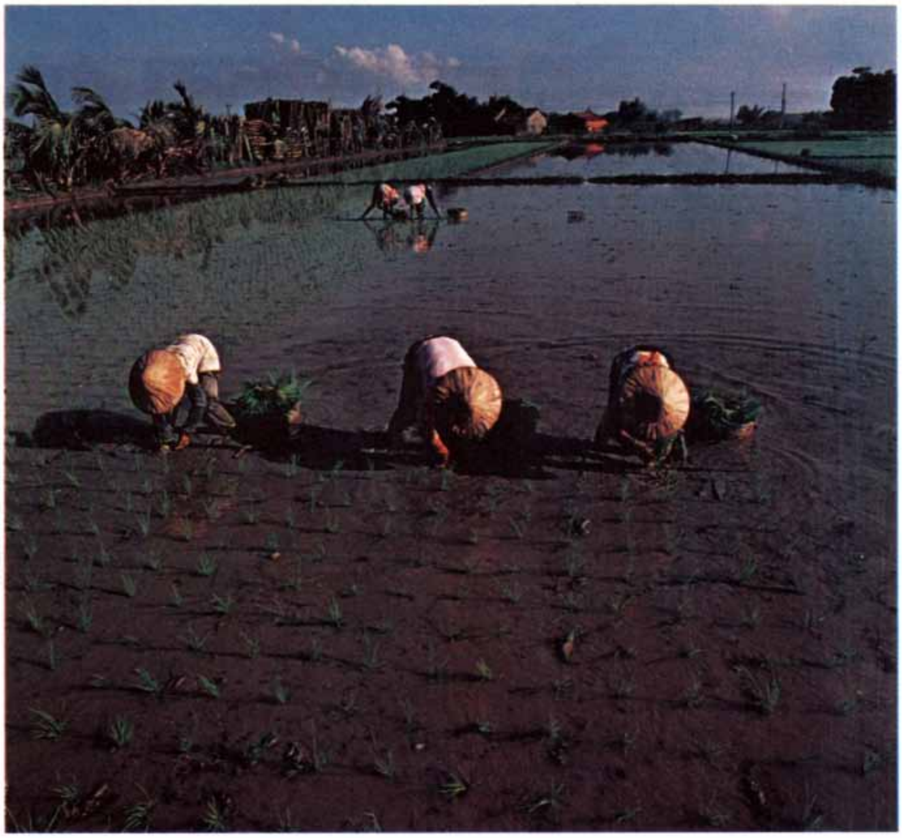
Most national collections in Asia have several common features. The main commercial varieties are well represented. Collections of minor and primitive cultivars are small. The collections include many duplicates, particularly of foreign varieties that have achieved an international reputation. Small numbers of breeding lines are maintained. Few wild rices have been collected. The lack of refrigerated facilities for storing seed has forced national centers to grow and renew their stocks almost annually, leading to mixtures, losses and mislabeling and a work load that has strained the local capabilities.

The International Rice Research Institute established its germ-plasm center in 1961. The collection grew to 12,000 accessions by 1970, assisted by excellent cooperation from various national centers. A cold-storage facility built at the start of the program has helped the bank to serve as a major seed-exchange center for many rice-growing countries.

Since 1971 the bank has been the central repository for the base collection of the world's rices. In 1983 the collection included 63,000 Asian cultivars, 2,575 African rices, 1,100 wild rices and 680 varieties maintained to test genetic traits. Thousands of breeding lines with one or more desirable traits are also preserved.

To provide extra security for the collection, which represents the products of thousands of years of natural and human selection, the institute deposits a duplicate set of seed at the U.S. National Seed Storage Laboratory in Fort Collins, Colo. A five-year plan has recently been developed, with the assistance of national rice scientists and the International Board for Plant Genetic Resources, to collect and conserve the remaining 30,000 to 40,000 varieties. When this task is completed by 1987, through the cooperation of scientists, farmers, students and military personnel who have access to remote areas, rice may be the first important food crop to be preserved for the 21st century and beyond in as complete a collection of naturally occurring genetic variability as can be achieved.

A recent example indicates the importance of this worldwide cooperation. Two years ago the International Rice Research Institute received a request from Kampuchea (through Oxfam, a humanitarian organization based in the United Kingdom) for a set of traditional rice varieties once cultivated in that country. Many local varieties had been lost to war. The institute's germ-plasm bank had 800 Kampuchean rice varieties as a result of a collection effort launched in 1973. Seeds of about 140 varieties required by the Kampuchean Ministry of Agriculture were supplied in 1981 and 1982; they are being multiplied locally. If it had not been for



RICE IS PLANTED by hand on a paddy in Taiwan. The scene is typical of the labor-intensive methods still employed in much of Asia. In a few places seed is dropped from airplanes.

RICE IS HARVESTED by hand near Chaing Rai in northern Thailand. This scene too could be duplicated in many of the world's rice-growing areas. In California, however, where much attention has been given to the mechanization of rice farming, the trend is to harvest with grain combines. The combine is usually fitted with tracks so that it can operate in the muddy paddy.

What if you chose as a technical

"At Union Carbide, an HP computer network automates production management and has paid for itself in less than two years."

Union Carbide's Linde Division in Tonawanda, N.Y., manufactures custom production equipment. Factory systems manager Doyce Coffman says, "When production demand fluctuates, it complicates our scheduling, product costing, and long range planning.

"To solve these problems we automated production and materials management with a network of HP 1000 and 3000 computers and 43 data capture terminals.

"Because these systems continuously monitor operations and respond in real time, we can identify and eliminate production bottlenecks on line. The payback? Improved efficiency factory-wide and a dollar savings that covered the cost of our entire HP network in less than two years."

What should you expect from a working partnership with Hewlett-Packard?

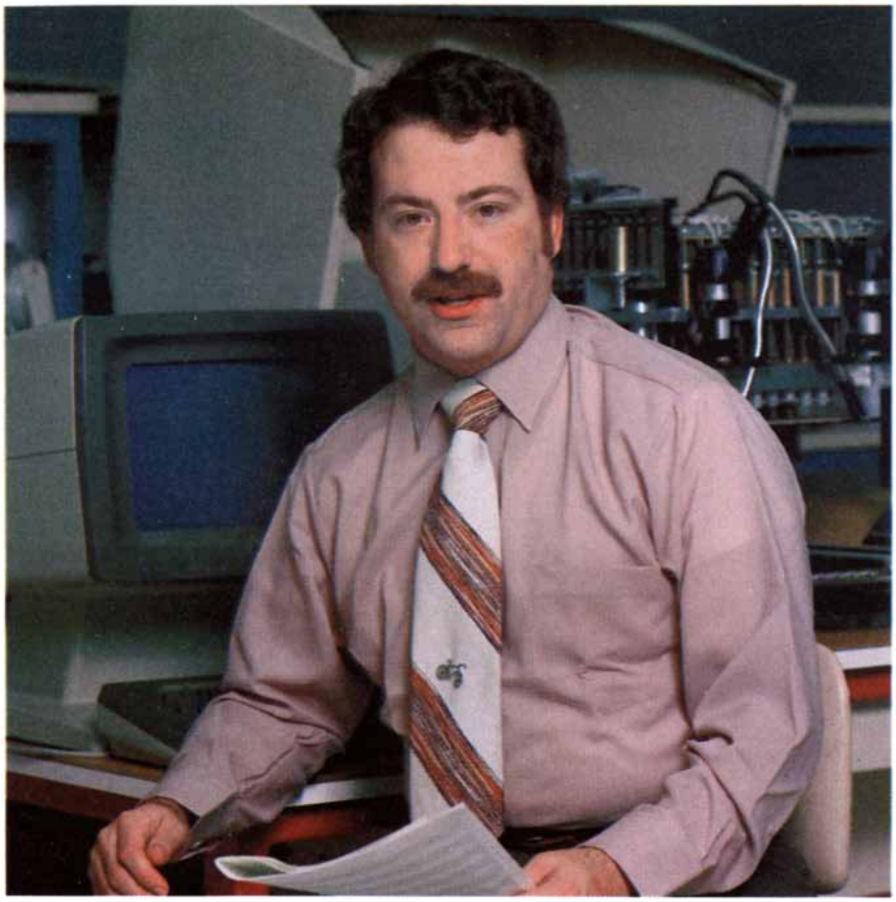
Results with assurance! DataPro Research Corporation's annual survey of computer users has ranked Hewlett-Packard *number one* in Customer Support Satisfaction for three years in a row. Why? Because the same quality and reliability engineered into HP products is built into our customer support programs.

From the outset, a Local Support Team of specialists who understand your needs will help you develop a cost-effective strategy to meet your goals—from selecting the right equipment to training your people.

You can tie your HP system into one of our Technical Centers where HP specialists will monitor your system's performance. Our routine, remote diagnostic tests can help iden-

The HP circle of quality customer support

Hewlett-Packard computer partner?



"At Amdahl, an HP computer network helps control testing. And it's paid for itself in two years."

Amdahl manufactures large-scale general purpose computers. In Sunnyvale, Calif., Bob Henderson, Manager of Test Systems Software, says: "We use a network of HP 1000 computers to control various tests on our Model 580 computers. With the HP systems, we can achieve the advanced testing and stringent quality control our high-speed, high-density products demand.

"Replacing batch systems with an HP computer network saved us nearly $100,000 in equipment costs. And by using HP 1000's we're obtaining test data ten times faster than before.

"Our test network helps streamline manufacturing processes and has paid for itself in just 2 years. In addition, the network saves us up to $34,000 a year by reducing paperwork and test cycle time."

tify and prevent problems before they occur.

The reliability designed into our systems and maintained by our worldwide support network enables us to offer Guaranteed Uptime Service—guaranteeing that covered systems will be up and running at least 99 percent of the time. **Results with assurance:** that's what a working partnership with Hewlett-Packard is all about.

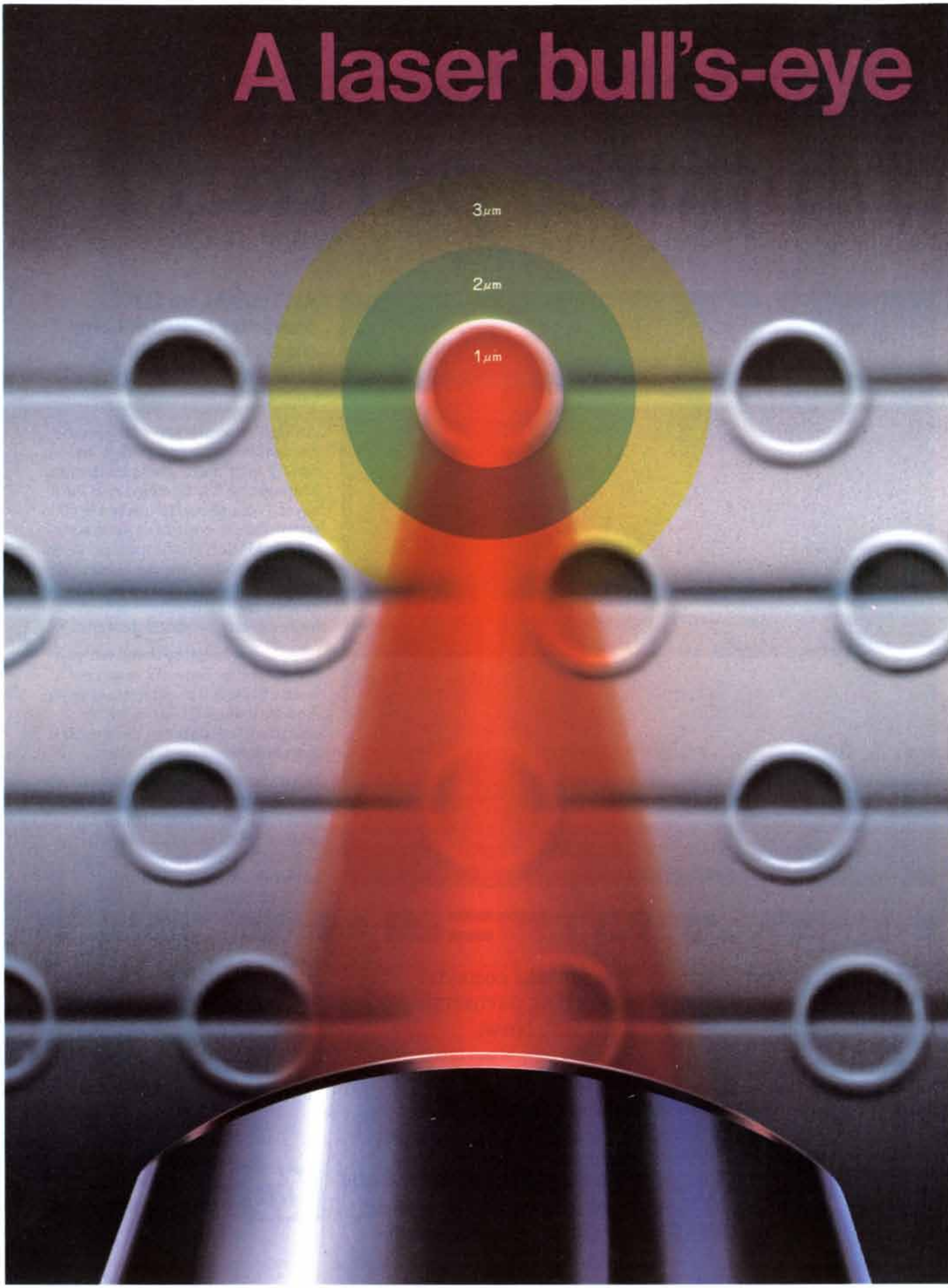
You should consider a working partnership with HP. Now.

For a free copy of our brochure explaining HP's practical, proven approach to meeting your long-term computer and information needs, write to A.P. Oliverio, Senior Vice President, Marketing, Dept. 207, Hewlett-Packard Company, P.O. Box 10301, Palo Alto, CA 94303.

When performance must be measured by results

HEWLETT PACKARD

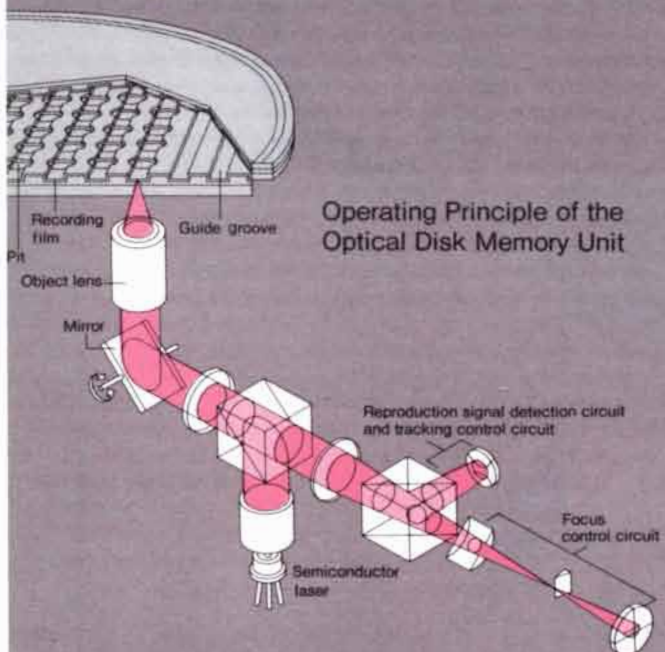
0002343


A laser bull's-eye
3μm
2μm
1μm

for optical memories

One error per million is too many

Optical memory systems allow nonerasable recording of up to 2.6 billion bytes of data on a single 12-inch disk. That's roughly equivalent to all the text appearing in 15.4 years of daily newspapers. But despite this incredible storage capacity, most systems available today have an error rate on the order of 10^{-5} or 10^{-6} during laser scanning of the disk's micron-sized digital pits, far from the 10^{-12} rate considered standard for conventional floppy disks and thus unacceptable in computer applications.



Operating Principle of the Optical Disk Memory Unit

Hitachi puts optical technology on target

To avoid error, the laser beam must be focused within 0.08 microns of dead center for each pit—a task as difficult as William Tell's if a grape had been placed on his son's head instead of an apple.

Hitachi engineers managed to achieve this amazing accuracy by using the beam itself to send two types of error signals to the scanner's control circuits. One is relayed to the object lens drive to refocus the lens. The other goes to the mirror drive to correct angular deviations of the laser by rotating its reflective mirror. These unique control techniques not only direct the beam to the bull's-eye, but also drop the error rate to an acceptable range of 10^{-11} to 10^{-12}.

Taking advantage of the digital difference

With this degree of precision, Hitachi's optical disk memory system can readily be applied in corporate record-keeping, libraries and archives. Unlike currently used magnetic tapes and microfiche, optical disks permit random access of information via computer, locating any bit of digital information in just a quarter of a second on the average. And once written, records are held permanently for an expected file life of ten years, adding longevity to the system's high-speed, high-capacity advantages.

Technical excellence is embodied in all Hitachi products

The development of a highly precise, optical disk memory system is just one case demonstrating Hitachi's technological strength. You'll find other examples in all of Hitachi's products, from Winchester hard disk drives and floppy disk units to a whole host of high-quality electronic appliances. Our comprehensive technological 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

the institute's germ-plasm collection, those locally adapted traditional varieties might have been lost forever.

As another example, if Indian scientists and the International Rice Research Institute had not preserved some of the unique genetic variability in rice that grows in the northeastern Himalayas and in the state of Kerala in southern India, it would have been difficult to stay one step ahead of the brown planthopper, an insect that preys on rice and is a menace to high-yielding varieties because it quickly develops resistance to pesticides and has the capacity to overcome genetic resistance based on single genes of the rice plant. Rices from this region possess genes for resistance not only to the brown planthopper but also to several other pests.

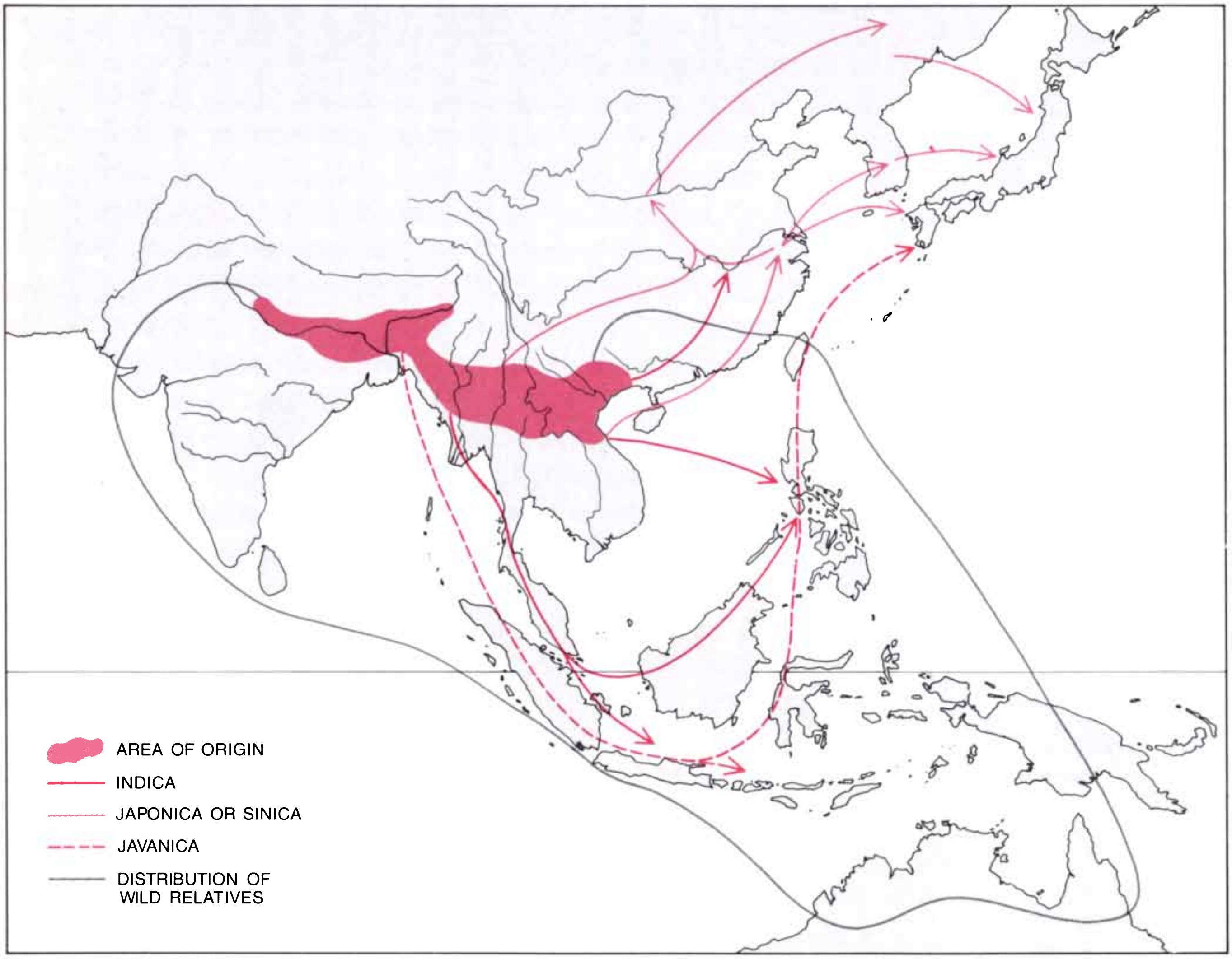
Why Rice Is Adaptable

Rice has an efficient system of air passage from shoot to root that makes it adaptable to a wide range of environmental conditions. The system enables rice to grow in waterlogged soils. Air enters the plant through the stomates of leaf blades and leaf sheaths and moves to nodes at the base of the plant. As the air moves from the shoot to the root, oxygen is supplied to the tissues, where it is utilized for respiration. From the roots the air diffuses into the surrounding soil. At least part of the shoot must be exposed to the air for the system to function efficiently.

Because of the air-passage system, the roots respire aerobically and utilize carbohydrates efficiently to produce needed energy even when they are growing in an anaerobic, waterlogged environment. Most other plants cannot grow in waterlogged soils because their air-passage system is less efficient. In rice the efficiency of oxygen transport from the shoot to the root is 10 times greater than it is in barley and four times greater than it is in maize.

In much of the world the rice crop is grown on soils with from five to 10 centimeters of water standing on them. The plant also yields well, however, on dry upland soils, where it is sometimes cultivated like wheat or maize, and in deeply flooded areas. For these reasons rice can be grown satisfactorily in mixed and multiple cropping patterns involving other crops such as wheat, maize and soybeans.

Floating rice can grow in water that is from 1.5 to five meters deep. It can withstand an abrupt rise in water level and grow rapidly above it, often growing as tall as six meters. It normally elongates at a rate of from two to 10 centimeters per day but can elongate as much as 25 centimeters per day in deep floodwaters. Workers at the International Rice Research Institute are collaborating with scientists in Thailand, Bangladesh and India to improve rice varieties adapted to areas flooded to a depth of one meter by incorporating the elongation genes of floating rice into improved high-yielding semidwarf varieties. These deep-water rices are short when the water is shal-



SPREAD OF ASIAN RICE (*O. sativa*) from its area of origin as a wild annual grass is indicated. Indica, japonica (or sinica) and javanica are the three major geographic races of *O. sativa*. The indicas were originally cultivated in the humid parts of the Asian Tropics and subtropics; japonicas developed in subtropical Temperate regions, and javanicas developed in equatorial Indonesia. Rice cultivation is believed to have begun in Asia some 7,000 years ago. The map is based on work of T. T. Chang of the International Rice Research Institute.

low, but they grow tall if the depth of the water rises because of flooding.

Nitrogen Fixation

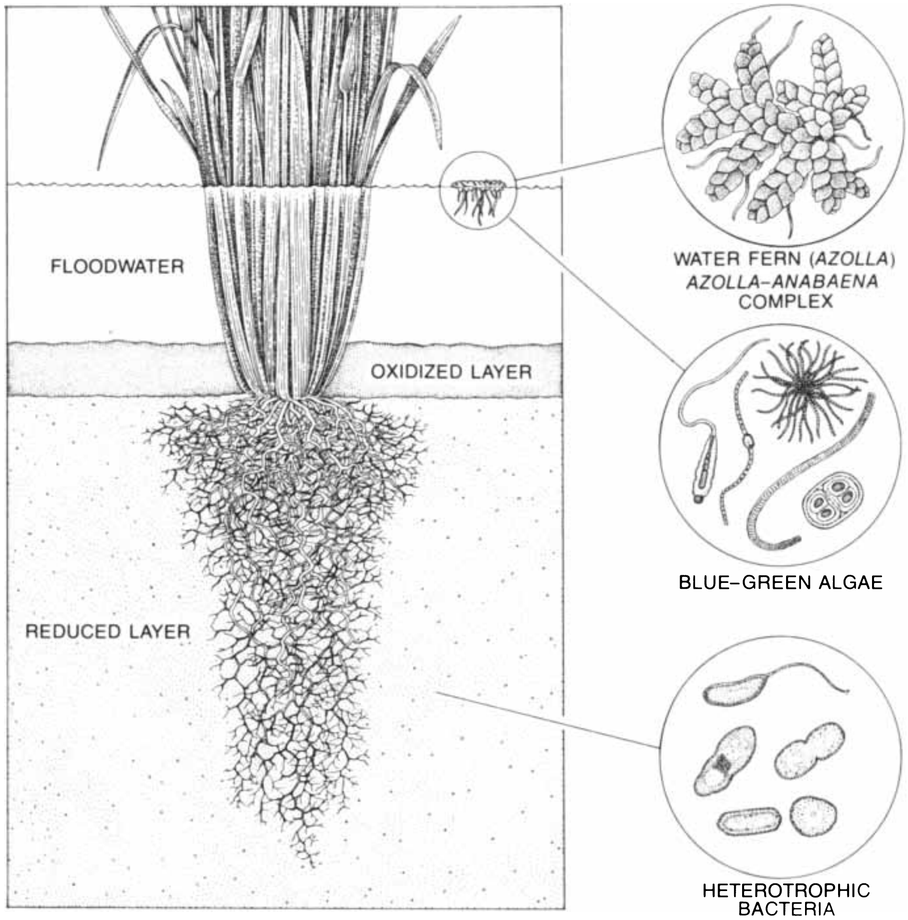
Some 80 percent of the atmosphere is nitrogen, which is not utilized directly by the rice plant. Nitrogen, however, becomes available through biological nitrogen fixation: the conversion of gaseous nitrogen (N_2) to ammonium ion (NH_4^+) by specialized microorganisms. Unique patterns of microbial metabolism are found in flooded rice soils, including transformations of nitrogen and sulfur, biological nitrogen fixation and the rapid decomposition of pesticides. Biological nitrogen fixation is the most intensively studied of these microbiological activities. Flooded rice soils encourage several nitrogen-fixing agents, including free-living blue-green algae, nitrogen-fixing bacteria and a symbiotic relation between nitrogen-fixing blue-green algae and the water fern azolla.

The symbiotic nitrogen-fixing relation between azolla and the blue-green alga *Anabaena azolla* is most active in flooded soils, where it can fix as much as three kilograms of atmospheric nitrogen per hectare per day. Nitrogen accumulated through this symbiosis is released into the soil and made available to rice plants when azolla decomposes. Because of its high nitrogen-fixing activity, azolla has served as a "green manure" for rice in China and Vietnam for centuries. (Vietnam has a temple dedicated to azolla.) The International Rice Research Institute is conducting collaborative research and training activities to disseminate knowledge of azolla and to encourage its use as a source of nitrogen for rice cropping throughout South and Southeast Asia.

Free-living nitrogen-fixing blue-green algae grow spontaneously in alkaline and neutral rice fields, where they can develop a large biomass. Trials are being conducted to improve their growth, either by means of cultural practices or by means of inoculation in rice fields.

Bacterial nitrogen fixation is a third source of biological nitrogen for the rice plant. Both aerobic and anaerobic nitrogen-fixing bacteria are present in rice-paddy soils. Surface and interior parts of roots and basal parts of shoots are colonized by nitrogen-fixing bacteria that are nourished by dead plant material and organic matter released by the living plant. Some 10^{10} cells of nitrogen-fixing bacteria are found per gram of rice root (dry weight). They probably provide a significant amount of nitrogen for the rice crop.

Rice plants need about 20 kilograms of nitrogen per hectare for each ton of rice produced. The different forms of biological nitrogen fixation enable rice to yield one or two tons of grain per hectare without supplementary mineral fertilizer. This is one of the reasons Asian farmers have harvested from one to two tons of rice per hectare for centuries without applying such fertilizers.



FIXATION OF NITROGEN by microorganisms in the soil and water means that a yield of about two tons of rice per hectare can be obtained without the application of mineral fertilizer. The organisms include free-living blue-green algae, *Anabaena azolla*, the same algae in a symbiotic relation with the water fern azolla, and a variety of bacteria. The habitats are also shown.

Improving Yields

Yields of rice were stagnant in most Asian countries until 20 years ago. Only in Japan had yields gradually improved from about 1.3 tons per hectare in A.D. 900 to 2.5 tons per hectare late in the 19th century, when the Meiji restoration led to the modernization of agriculture. Genetic improvements and better farm practices rapidly increased yields in the 20th century. Today yields in Japan, North Korea, South Korea, Australia and the U.S. (notably in California) average about six tons per hectare. Yields in most countries have increased gradually over the past 15 years, but many farmers still harvest only two tons per hectare.

The genetic key to these advances was the development of high-yielding varieties of indica rice. The traditional indica was tall and had a tendency to lodge, that is, to fall over because of wind, rain or the weight of the grain. Most of the modern varieties are semidwarfs.

Modern semidwarf indica rices were developed from semidwarf mutants such as Dee-geo-woo-gen, a Chinese variety. In the early 1950's Taichung Native 1, the first semidwarf indica rice bred outside mainland China, was developed and released in Taiwan. It resisted lodging, was responsive to applications of fertilizer and was fairly insensitive to differences in the length of the day with latitude.

When the Ford and Rockefeller foundations, with the cooperation of the Philippine government, established the International Rice Research Institute in 1960, one of the first projects set in motion by the institute was the achievement of dramatic improvements in yield by means of breeding programs utilizing the best available parents from many nations. IR8, a semidwarf, was a notably successful result. Released in the Philippines in 1966, the new variety set yield records and became known as "the miracle rice." Other high-yielding semidwarfs followed, including IR5, IR20, IR22 and IR24. They had progressively shorter growing periods and a better response to good management.

The dispersion of such high-yielding, short-duration varieties encouraged farmers to grow two or more crops of rice per year in dense plantings on irrigated fields with large applications of fertilizer. This technology brought more grain to market, but it also brought on increased problems with disease and insects. Workers at the institute therefore began to breed varieties with insect and disease resistance and to emphasize reductions in year-to-year variability as well as high yield in their strategy.

IR36, a variety now grown on more than 10 million hectares of the world's rice land, is a result of this strategy. It resists four major rice diseases and four serious rice insects, including brown planthopper biotypes 1 and 2. It grows well in a variety of cultural environments, tolerates several adverse soil conditions, has grain of good quality and matures in 110 days, which enables farmers to harvest as many as three crops in one year on irrigated paddies. IR36 is the progeny of 13 different varieties from six nations. Among its ancestors are IR8, Taichung Native 1 and a wild species from India, *O. nivara*.

Insects and Diseases

Breeding for pest resistance is complicated by the adaptation of insects and disease-causing organisms to resistant varieties of rice. For example, three distinct brown planthopper biotypes have been identified that directly damage rice plants and are also vectors of important virus diseases. When IR8 was first grown, only one brown planthopper was known to attack rice. Since then, as new resistant rice varieties have been developed, resistant biotypes of the insect have also selectively multiplied. Already brown planthopper biotype 3 may be starting to damage rice crops in certain areas of the Philippines and Indonesia. Fortunately variety IR56 resists brown planthopper biotype 3 and has been released in those areas.

It is essential that rice research remain one step ahead in the dynamic interaction of rice varieties and pest resistance. For this purpose workers in the institute's multidisciplinary Genetic Evaluation and Utilization Program systematically screen the worldwide collections of rice germ plasm for reaction to the major pests of rice.

Varieties with desirable resistance are crossed in a hybridization program to combine many positive traits in one variety. More than 4,000 crosses are made each year at the institute. The most promising material is tested at more than 300 locations under the International Rice Testing Program, encompassing some 800 plant scientists in 70 countries. Several successful varieties have been developed through this cooperative network.

New insect and disease problems are continuously monitored. As a result a new strain of grassy stunt virus, tentatively designated RGSV2, has been identified. This strain is found in several parts of the Philippines and may cause significant damage to rice. Vigorous screening is now in progress to identify sources of resistance to it.

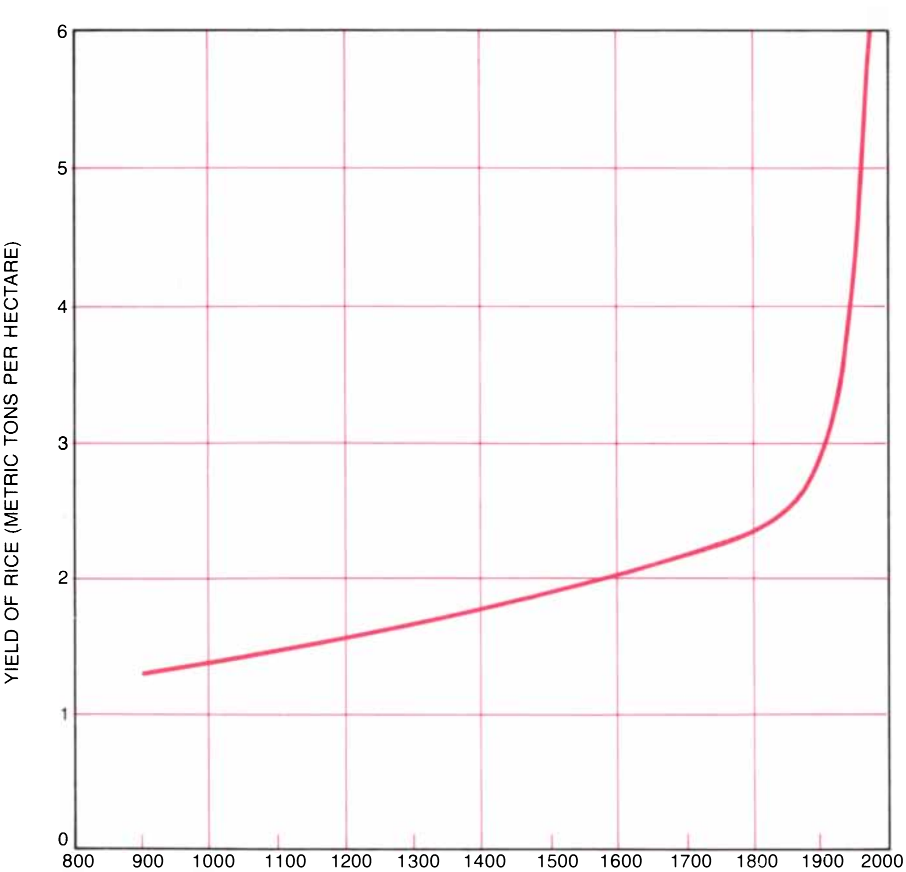
The never-ending contest between pests and human beings has led to the realization that the goal in farming should be pest management rather than pest elimination. To manage pests so that serious crop losses do not result it is necessary to adopt an integrated approach incorporating genetic, agronomic and biological methods of control, coupled with the prudent deployment of safe and effective chemical pesticides.

What Lies Ahead

The Food and Agriculture Organization has estimated that to keep pace with the rise in population an annual rate of increase of about 3 percent in rice production will be necessary over the remaining years of the 20th century. Production can surely be increased at that rate through the effective transfer of currently available technology, because modern technology and agronomic advances have not yet been fully disseminated throughout the rice-growing world. It would therefore be interesting to know how much rice can be produced exploiting the high level of technology currently available. Such information will provide a target for future yields. At experiment stations in Japan workers have achieved yields of from 9.3 to 10.2 tons per hectare with modern technology. Even that falls short of the highest yield ever recorded: 13.2 tons per hectare, achieved by a progressive Japanese farmer in 1960.

In developing realistic national plans for rice production it will be useful to know what the theoretical production potential is in each country. Since green plants convert solar energy into chemical energy, it will be worthwhile to estimate the potential rice yield under different levels of incident solar radiation. One important concept useful for the estimation of potential rice yield is energy-use efficiency, which can be likened to the energy-conversion efficiency for silicon solar cells.

Rice builds up most of the starch in the grain in the last stage of its growth. That stage is called the grain-filling period. To estimate the potential rice yield in tropical Asia the energy-use efficiency in the grain-filling period is assumed to be 3 percent for the wet-season crop (when most of the rice crop is grown) and 2.5 percent for the dry-season crop (the off-season crop, grown with irrigation). At many locations in tropical Asia solar radiation of 300 calories per



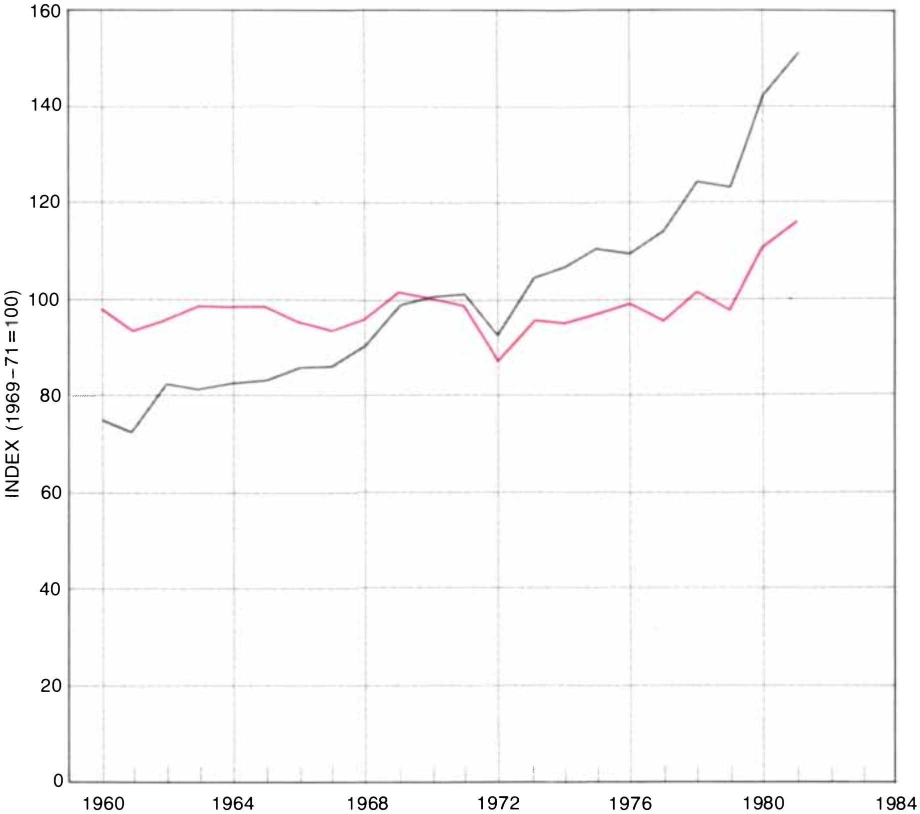
YIELDS OF RICE in Japan rose slowly from A.D. 900 to the late 19th century and then much more rapidly. The period of gradual rise is accounted for mainly by the spread of irrigation and by flood-control procedures. The recent rapid rise is due to the breeding of improved varieties, the trend toward the application of mineral fertilizer and better farm management.

square centimeter per day is common for the grain-filling period in the wet season. In the dry season, with fewer cloudy days, it should be possible to produce 12.5 tons per hectare with an energy-use efficiency of 2.5 percent under an incident solar radiation of 550 calories per square centimeter per day.

Another question of general interest is the relative yield potential of rice in Temperate Zone regions and the Tropics. The number of days for grain filling in rice is largely determined by temperature. It is about 25 days in the Tropics and 35 days in the Temperate Zone. If the daily incident solar radiation is the same for both regions, the total amount of incident solar radiation would be proportional to the effective number of days for grain filling. In other words, the potential rice yield would be higher by 40 percent in Temperate Zone regions than it would be in the Tropics, given that the daily incident solar radiation is the same. On the other hand, two or more rice crops can be grown in the Tropics if enough water is available.

Irrigated rice crops currently occupy about 30 percent of the rice-production area in developing tropical countries and account for 50 percent of the rice grown in those countries. Irrigated rice will assume increasing importance in the future because most developing countries are bringing more land under assured irrigation. Irrigation makes possible not only higher productivity but also higher cropping intensity, achieved by growing an irrigated rice crop, a nonirrigated crop of a grain legume or something else and then a second rice crop, all in one year.

In nonirrigated, rain-fed lowland and upland rice areas yields are low and unstable because the amount and distribution of rainfall are uneven and erratic. Such undependable weather conditions make nonirrigated rice cultivation risky. Hence farmers invest little in fertilizer, pesticides and other purchased inputs. The average yield of nonirrigated rice ranges from .5 ton to 1.5 tons per hectare for upland crops and is less than two tons per hectare for rain-fed lowland rice. The yield has increased little over the past two decades, particularly in upland rice. Intensive efforts are now being organized at the International Rice Research Institute and other institutions to increase and stabilize rice yield on nonirrigated lands.



RECENT TRENDS in total rice production (*gray*) and production per capita (*color*) are indicated for Southeast Asia. Notwithstanding the steady rise in total production, the rice available per capita has changed little since about 1969 because of the rate of increase in the population.

Genetic Engineering

Recent advances in biotechnology may offer new ways to improve yield. One is tissue culture leading to haploid plants: plants with only one complete set of genes instead of the usual two. With such plants one can stabilize vital characteristics because two competing genes for the same characteristic are not present. Moreover, the technique reduces the time needed to get a new variety from 10 generations with conventional breeding methods to as few as three generations. The Chinese have employed this procedure to create new varieties of rice, wheat and maize.

Most of the research on the haploid technique has been done in Japan, China and the U.S. with japonica rices. So far the procedure has not succeeded with indica rice because the indica varieties do not grow a useful number of calluses, an early stage of growth in which the tissue is undifferentiated; such tissue is needed for culture purposes. Much effort is being devoted to making the technique successful with indica rices.

Another promising technique is to modify plants by inducing mutations in tissue culture. Individual plant cells or pollen grains in culture can serve as mutable material that can be grown into whole plants. One advantage of this approach over the procedure that induces mutations in seeds and whole plants by irradiation or chemicals is that a particular trait can be selected at the cell level. Promising results for increasing the tolerance for salinity by rice plants have been obtained by this technique.

Among the many aims of genetic engineering the most ambitious goal is the incorporation of nitrogen-fixing genes into rice. The nitrogen-fixation system includes at least 17 genes, however, and it is too early to know whether the manipulation of such a large number of genes is feasible.

China has introduced hybrid rice into cultivation, opening the possibility of the remarkable improvements that have been achieved in this way with maize. The technique should not be confused with crossing, which consists in combining different varieties to achieve an enduring new line. In the hybrid technique only the first generation (F_1) of hybrids is sought; that generation is usually more vigorous than either of its parents. Hence the farmer must buy a new supply of F_1 seed each year.

The technique has been difficult in self-pollinating plants such as rice, but the recent development of the breeding system called cytoplasmic male sterility has made hybrid varieties possible in such plants. The system depends on the interaction of a gene that causes sterility in male plants and a factor in the cytoplasm of the female sex cell. The application of the technique to rice became possible after Chinese workers discovered male-sterile plants that can serve as female plants in crosses with suitable male parents. The advantage of the hybrid in terms of yield should be sufficient to offset the high cost of seed production. It is still too early, however, to forecast how successful the technique will be in China or to estimate how it will fare in other countries.

Unisexual Lizards

Populations of whiptail lizards in the southwestern U.S., Mexico and South America consist of females only. The animals are unusual among vertebrates in reproducing by parthenogenesis: virgin birth

by Charles J. Cole

In 1958 a zoologist at the Academy of Sciences in Leningrad, Ilya Darevsky, published a startling report on certain lizards of Soviet Armenia. He had found that several populations of the lizard genus *Lacerta* consisted exclusively of females, and he suggested the lizards could lay eggs that were viable although they were not fertilized. Many zoologists received Darevsky's finding with incredulity. At the time it was almost axiomatic that no vertebrate species could reproduce except through the union of the male's sperm with the female's egg.

To some biologists in the U.S., however, Darevsky's report was credible and welcome. For example, Sherman A. Minton of the Indiana University Medical Center and Richard G. Zweifel of the American Museum of Natural History had found populations of whiptail lizards (the genus *Cnemidophorus*) in the southwestern U.S. and northern Mexico that lacked males, although neighboring whiptail populations consisted of males and females in the normal one-to-one ratio. Minton, Zweifel and a few other biologists decided that further study of various whiptail lizards might well add another group to the roster of parthenogenetic species.

Parthenogenesis, or virgin birth, is a term that is often misapplied to what are only apparent forms of unisexual reproduction. For example, among some unisexual fishes the females produce diploid eggs: germ cells that contain two full sets of chromosomes. Thus the eggs are genetically complete, although they will not develop until they are activated—not fertilized—by contact with the sperm of a male. Since there is no fertilization, the process has been loosely classified as parthenogenesis. It is more correct to call it gynogenesis, a term that emphasizes the dominant role of the female. Among other unisexual fishes the females produce haploid eggs: germ cells that like the germ cells of species that reproduce sexually have only one set of chromosomes. These eggs do not develop until they are fertilized by a male belonging to a closely related species, a process that may be called hybridogenesis to emphasize the fact that the all-female offspring are hybrids. In addition at least one fish species, the rivulus (*Rivulus marmoratus*), reproduces as a self-fertilizing hermaphrodite. In all three instances, however, reproduction depends on the presence of sperm. Only reproduction in the certain absence of sperm can strictly be called parthenogenesis.

To prove that sperm are absent in the reproduction of a population in the wild is clearly impossible. It is not easy even under laboratory conditions. For example, females of some normally bisexual reptile species can store sperm in their oviducts for months or even years after mating and then produce viable fertilized eggs. Therefore to prove that some whiptail lizards were strictly parthenogenetic it would be necessary to raise first-generation females in the absence of males until they reached sexual maturity and then see whether they could produce a second generation in the continued absence of males.

So it was that early in the 1960's apparently unisexual whiptail lizards were captured in the wild and kept in laboratories for the purpose of studying their reproduction. Most of the early attempts were discouraging because no one knew how to maintain this kind of wildlife properly in captivity. At first some females did lay eggs in the laboratory, but most of the eggs spoiled. When this problem was overcome in Zweifel's laboratory, first-generation whiptail lizards were successfully hatched there and elsewhere but a second problem arose. Some 90 percent of the newly hatched lizards died within three months and most of the rest died within six months. Whiptails require almost a year to reach sexual maturity, and so it began to seem that getting second-generation whiptails would not be feasible.

In 1973, however, a factor vital to the lizards' survival in captivity became clear. The animals had already been provided with an appropriate living surface, shelter, warmth, food, water and generally hygienic conditions. What was lacking under routine laboratory lighting was exposure to the ultraviolet radiation in sunlight. The action of ultraviolet wavelengths on the lizards' skin brings about the synthesis of vitamin D (as it does in mammals, including man). That vitamin is essential to proper calcium metabolism. As the required radiation was supplied to the animals in the American Museum laboratory of herpetology Carol R. Townsend and I further discovered that different species of whiptails require different doses of it.

Once these details were worked out the raising of many unisexual whiptail species in captivity became routine. In our laboratory unisexual lineages have now reached not only the initial goal of a second generation but also generations unto the seventh, with offspring numbering in the hundreds and not a single male among them.

By the late 1970's the evidence that all-female whiptails were reproducing by means of strict parthenogenesis, that is, in the total absence of sperm, was becoming incontrovertible. It was not a conclusion to be taken lightly, because parthenogenetic species were previously unknown among the vertebrates. The best evidence of parthenogenetic reproduction came from comparing the patterns of development in unisexual females and bisexual whiptail lizards of both sexes, with particular reference to their reproductive tract. Laurence M. Hardy of Louisiana State University at Shreveport did most of this work, starting with a number of our second-generation offspring.

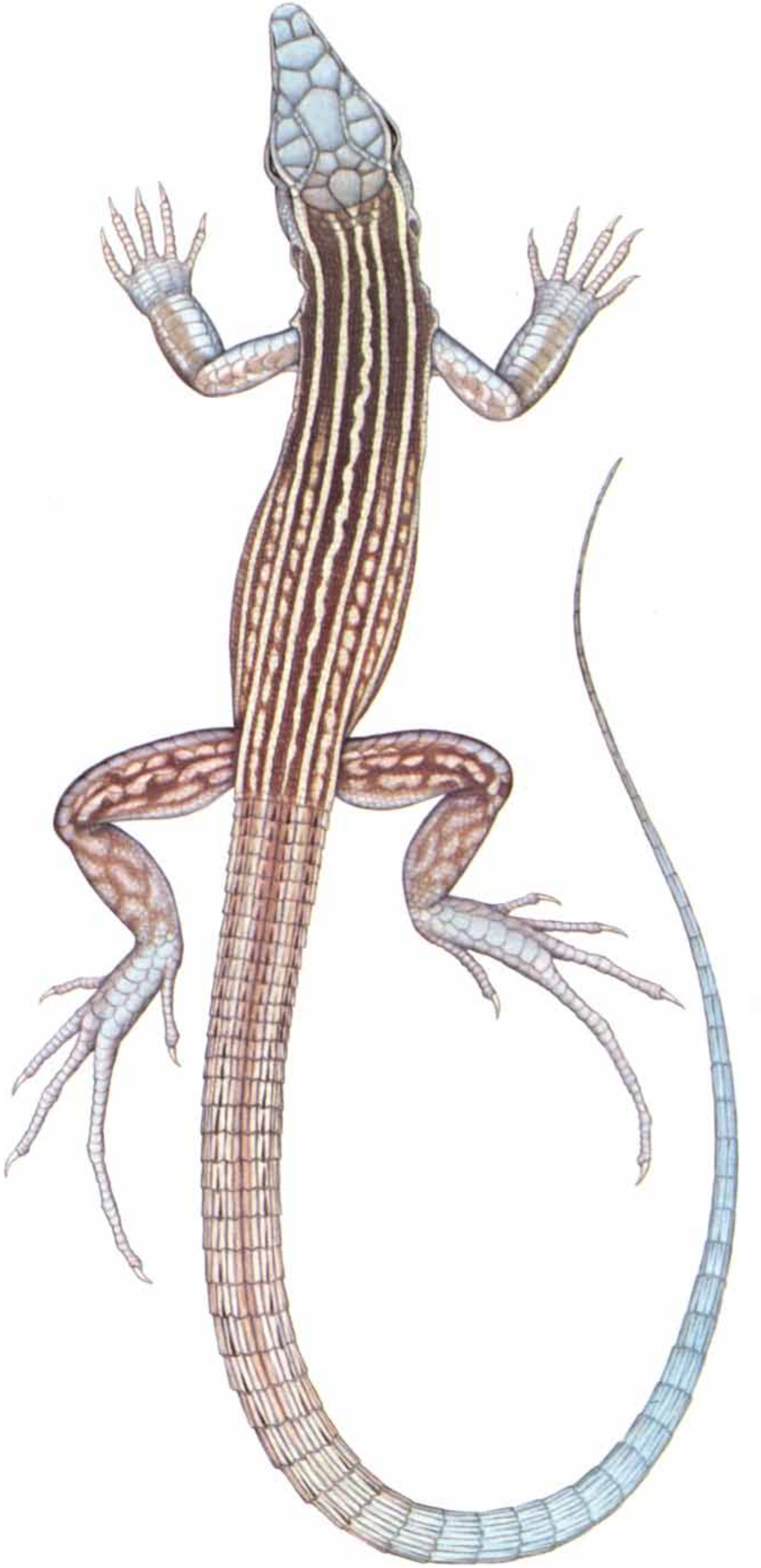
Hardy found that the reproductive tracts of the unisexual lizards were identical in every important detail with those of the females of bisexual species. He also found no trace of sperm in any of the unisexual whiptails, lending conclusive support to our view that they reproduce strictly by parthenogenesis.

In the years before laboratory colonies of whiptails were successfully established a number of biologists, including Zweifel, T. Paul Maslin of the University of Colorado at Boulder and William Neaves of the Harvard Medical School, conducted experiments with unisexual whiptails captured in the field to test the hypothesis that their populations were clones, that is, populations of genetically identical individuals related by common ancestry. The conclusions of their studies can be summarized briefly as follows. First, with respect to such characteristics as the number of epidermal scales around the middle of the lizard's body, the unisexual species show less variability than neighboring bisexual species. Second, the body cells of the lizards in the unisexual species contain marker chromosomes indicating the presence of two different genomes (complete sets of genes). This is to say the chromosomes are heterozygous. Moreover, when this is clearly the case, the heterozygous state appears to be fixed in the population because all individuals are alike. In contrast, the chromosomes of lizards in the bisexual species are homozygous: the animals have two similar genomes.

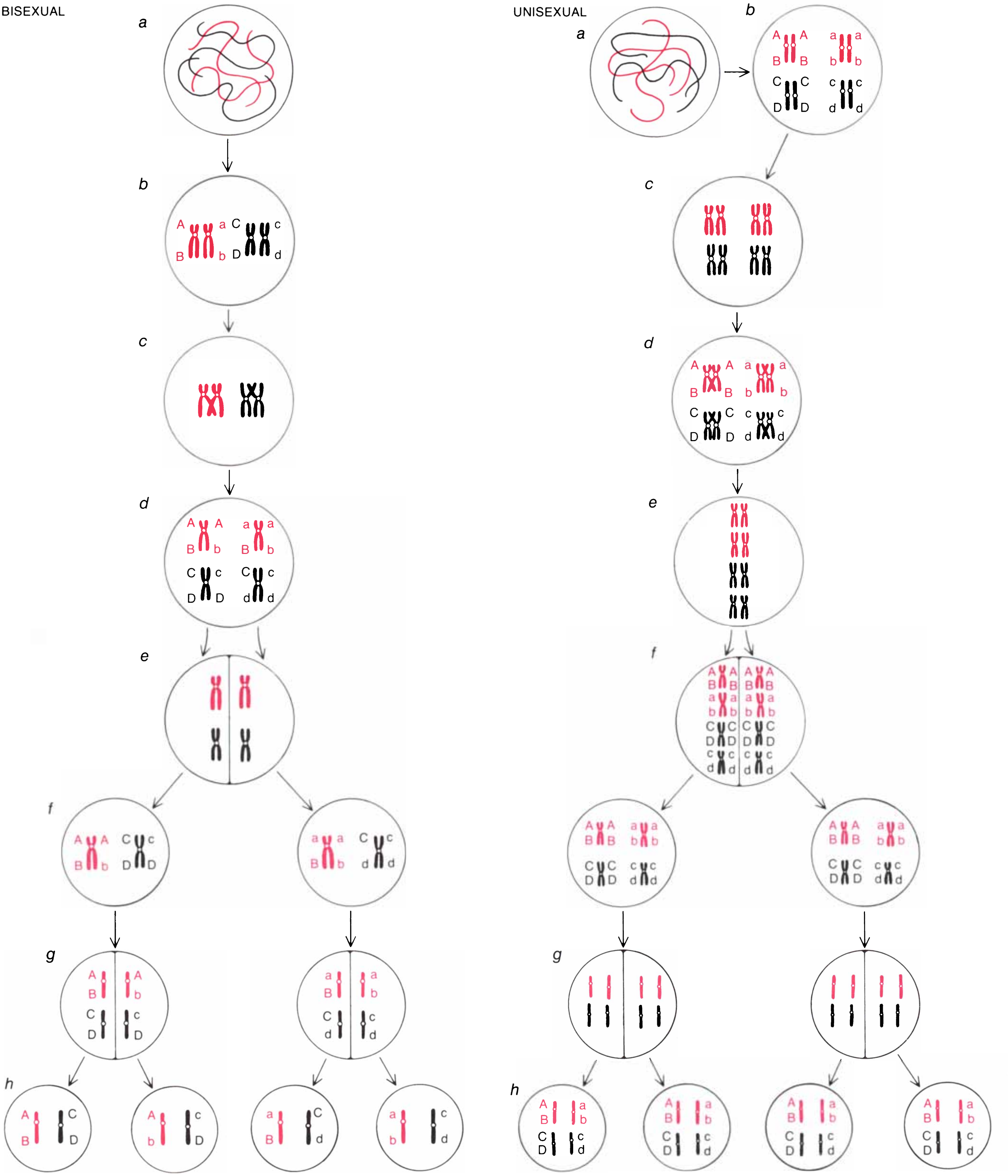
Third, the body cells of the lizards in many unisexual species have an unusually high chromosome count. Instead of showing the normal diploid count characteristic of bisexual species they are triploids. Fourth, when proteins of unisexual species are separated by the laboratory technique of electrophoresis, they reflect a high level of apparently fixed heterozygosity, whereas proteins of bisexual species do not. Finally, when a piece of skin is transplanted from one lizard to another of the same unisexual species, the graft usually takes, demonstrating that individuals of the same species have a substantial tissue compatibility. Exactly the opposite is true of bisexual whiptail species. All these findings are consistent with the hypothesis that unisexual reptiles are clones.

The laboratory colonies of unisexual whiptails have made it possible to collect added clone data that are simply not available from lizards captured in the field. For example, in our laboratory eggs and hatchlings are maintained in such a way as to permanently establish the individual lizard's ancestry and identity. Thus the genealogy of an individual and her relatives can be traced back to when the original female was collected in the field several generations earlier. This makes it possible to trace through a unisexual lineage such characteristics as external morphology, chromosome states and protein electrophoretic patterns.

The findings derived from our laboratory colonies are as follows. First, with respect to morphology, the color, the color pattern and the shape and size of the scales of individual unisexual lizards



WHIPTAIL LIZARD of the New World genus *Cnemidophorus* is drawn twice life size. Most species of whiptails are bisexual, and their reproduction requires fertilization of the female's eggs by the male's sperm. This female is one of the unisexual species: *C. neomexicanus.* Its eggs develop in the total absence of sperm. Its offspring will also be parthenogenetic females.



PROCESS OF MEIOSIS in bisexual whiptails is compared with the same process in unisexual species. On the left side of the illustration (*a*) two pairs of chromosomes (*black and color*) are shown schematically in an ovarian cell that is diploid: having two sets of chromosomes. Next the chromosomes duplicate and pair up (*b*). The chromosomes may then "cross over" (*c, d*), a process that can, as it does here (*d*), make the combination of genes on certain chromosomes (*indicated by the letters* A, a, B, b, C, c, D, d) differ from the initial one. The first cell division follows (*e*), with one chromosome of each set going to each new cell (*f*). A second cell division (*g*) gives rise to four new cells, one of which becomes a functional egg. Each cell is haploid: having one set of chromosomes (*h*). Now only union with a haploid sperm, that is, fertilization, will restore the egg cell to the viable diploid state. On the right side of the illustration is the process in a unisexual lizard's cell. The start (*a*) is similar, with two pairs of chromosomes in a diploid cell. The chromosomes then, however, duplicate and separate (*b*). Next they duplicate again and pair up (*c*), now in a tetraploid cell. Each new chromosome pairs with the original one that gave rise to it. Crossing-over (*d*) is followed by the first division of the tetraploid ovarian cell (*e*), with one chromosome of each duplicated set going to each new cell (*f*). The next division (*g*) gives rise to four new cells, one of which becomes a functional egg. Each egg is diploid (*h*) and genetically identical with the original ovarian cell. Development of these eggs begins without contact with sperm.

are nearly always identical with those of their mother. There are, however, some unexpected variations, and we are now seeking to find whether their basis is genetic or environmental. Second, with respect to chromosomes, whether the mother's species is diploid or triploid, she produces offspring with chromosomes identical with her own.

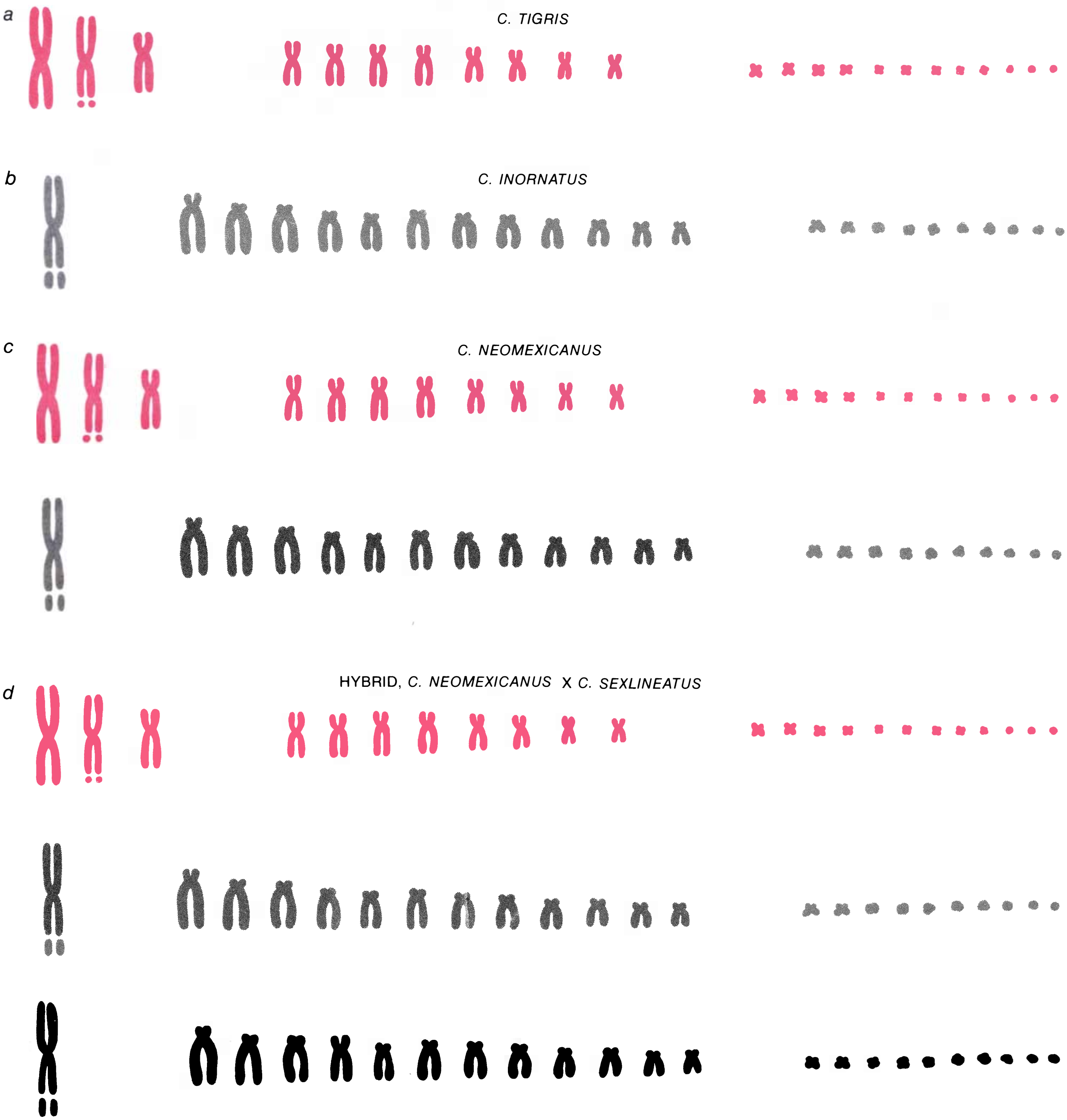
A third laboratory finding, in the field of biochemical genetics, has emerged from the studies of Herbert Dessauer of the Louisiana State University Medical Center in New Orleans. He has studied the proteins produced by more than 30 genes and has found that generation after generation the unisexual whiptail offspring show the same electrophoretic protein patterns as their mother. All the same genes are expressed, and all the gene combinations, including heterozygous states, are fixed. Hence the laboratory experiments strongly reinforce the conclusions of the field studies.

A few observations on how eggs are cloned by parthenogenetic whiptails indicate that the chromosomes duplicate themselves within the ovarian cells destined to become eggs. Thus these specialized reproductive cells become polyploid. In diploid species they become tetraploid, and in triploid species they become hexaploid. The chromosome counts, together with the evidence for clonal inheritance, suggest that in pairing to produce ova each duplicated chromosome pairs with the chromosome of which it is an exact copy. Therefore an exchange of DNA in "crossing over" among paired chromosomes has no genetic consequences because it involves identical strands of DNA. The crossing over is followed by the two cell divisions typical of egg maturation, so that the final product is an egg containing the same chromosomes and the same combination of genes as existed in the original ovarian cell before the chromosomes were duplicated. As a result the mature egg is genetically complete rather than haploid, as it is in females of bisexual species, and begins development without fertilization.

How did this peculiar parthenogenesis arise in nature? The whiptail lizard genus *Cnemidophorus* is confined to the New World and includes about 40 species. Most of the species are bisexual, but perhaps 12 are unisexual. Most of the unisexual species are found in an area of the southwestern U.S. and northern Mexico where plant communities have been shifting and mixing in periods of climatic change in recent geologic times. In some places one or two bisexual whiptail species share their range with a unisexual species.

Under these circumstances a parthenogenetic female is occasionally inseminated by a male of a bisexual species and hybrid offspring result. The offspring have polyploid chromosome counts: diploid and triploid unisexual females respectively produce triploid and tetraploid hybrid offspring. Dessauer has found that all the genes detectable by protein electrophoresis as being characteristic of each parent are functioning. The hybrids exhibit more of their mother's traits (she having contributed two or three sets of chromosomes) than of their father's (he having contributed only one set). If the egg is fertilized by a sperm carrying an *X* chromosome, the hybrid is a female; if it is fertilized by a sperm carrying a *Y* chromosome, the hybrid is a male.

The characteristics of such hybrids suggest that in nature unisexual triploid species of whiptails were cloned from hybrid female offspring of a diploid parthenogenetic female's insemination by



CHROMOSOMES of several whiptail lizard species are shown diagrammatically. At the top (*a*) are the 23 chromosomes of a *C. tigris* germ cell, the haploid number. Below them (*b*) are the 23 chromosomes of a *C. inornatus* haploid germ cell. Next are the 46 chromosomes in a typical body cell of the unisexual *C. neomexicanus*, the usual diploid count (*c*). Their similarity to the two sets of chromosomes above them is consistent with the hypothesis that the origin of *C. neomexicanus* was a crossbreeding of the bisexual *C. tigris* and *C. inornatus*. As is shown in the illustration on page 96, this combination of chromosomes is cloned generation after generation in *C. neomexicanus*. If one of these unisexual females were to crossbreed with a male of a bisexual species, a hybrid with three sets of chromosomes would be the result (*d*). The triploid array of 69 chromosomes is from a body cell of a laboratory crossbreeding of a *C. neomexicanus* female with a male of the bisexual *C. sexlineatus*. If a triploid were crossed with a normal male, the body cells of offspring would be tetraploid.

a male of a different species. But how did the diploid parthenogenetic female arise? This question was answered in 1966 by Charles H. Lowe, Jr., of the University of Arizona and John W. Wright, who is now at the Los Angeles County Museum of Natural History. One of the species they examined most closely is the unisexual New Mexico whiptail, *C. neomexicanus.* The species is found along streams in the Rio Grande drainage system and in nearby desert-grassland habitats, areas of transition between desert and grassland. One bisexual whiptail species, *C. inornatus,* occupies a grassland habitat it shares marginally with the unisexual *C. neomexicanus.* Another bisexual species, *C. tigris,* occupies a desert habitat it also shares marginally with *C. neomexicanus.*

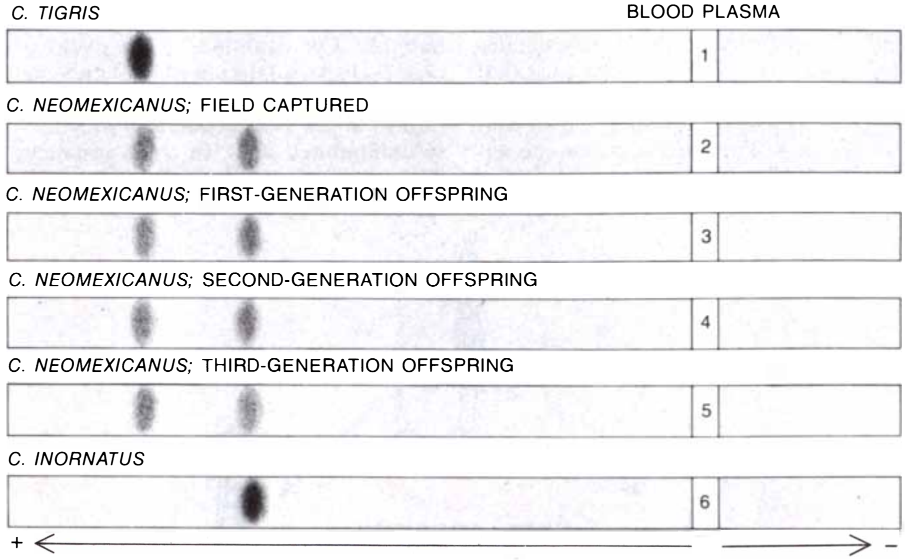
Lowe and Wright found that in such pertinent attributes as color, color pattern, scale shape, chromosomes and preferred habitats the character of *C. neomexicanus* appears to be that of a first-generation hybrid produced by the mating of *C. inornatus* and *C. tigris.* The *C. neomexicanus* populations have preserved their identity as first-generation hybrids generation after generation by parthenogenetic cloning.

Lowe and Wright suggested that at some time in the past the grassland species of whiptail interbred with the desert species, producing hybrids such as *C. neomexicanus.* As is the case with most interspecific hybrids, the initial crosses probably included both males and females, most of them sterile. At the same time the hybrids probably competed successfully with the nonhybrids in the mixed desert-grassland habitats. The first-generation hybrid males would have disappeared eventually, but any females capable of parthenogenetic cloning would have perpetuated their kind.

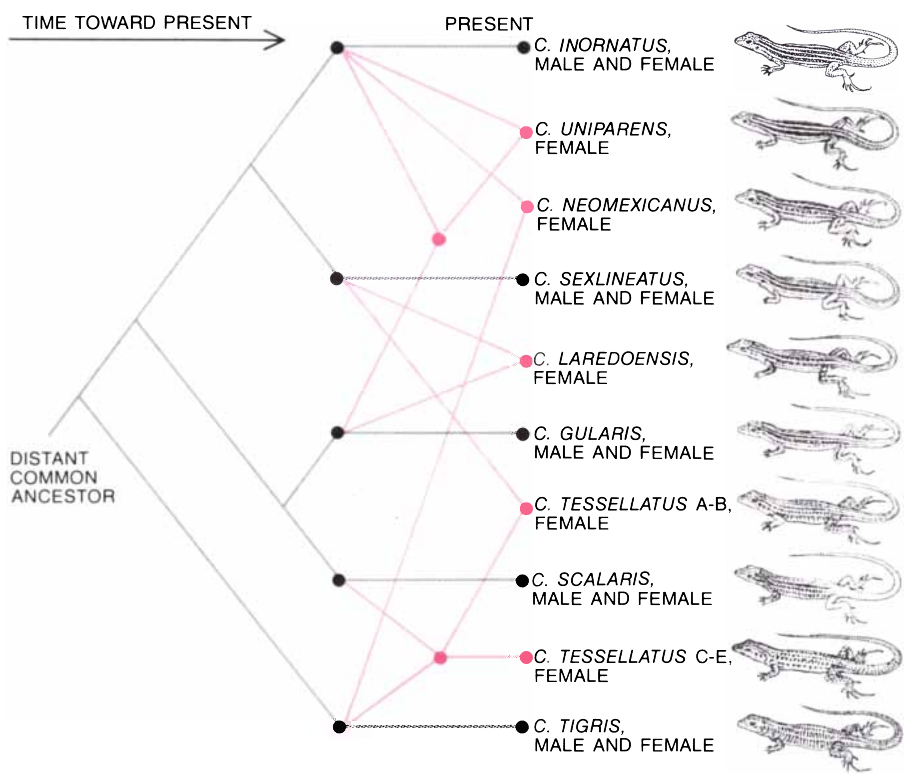
The sterility of interspecific hybrids is generally thought to be an example of natural selection in action. This is to say that individuals within a species that breed only with others of their own species generate the most fertile offspring, whereas hybridization, with its attendant sterility, would appear to have no adaptive advantage. In the case of the unisexual whiptails, however, the animals have continued to perpetuate themselves by cloning in spite of their hybrid origin. Do these parthenogenetic lizards enjoy any adaptive advantages?

One immediate advantage, of course, is the fact that parthenogenesis cancels the disadvantage usually imposed by hybrid sterility. Clonal inheritance also seems to maintain the advantage sometimes called hybrid vigor. This is to say that by cloning the animals do not break up a new combination of genes that has proved to be successful in a particular environment.

A third advantage arises from the liz-

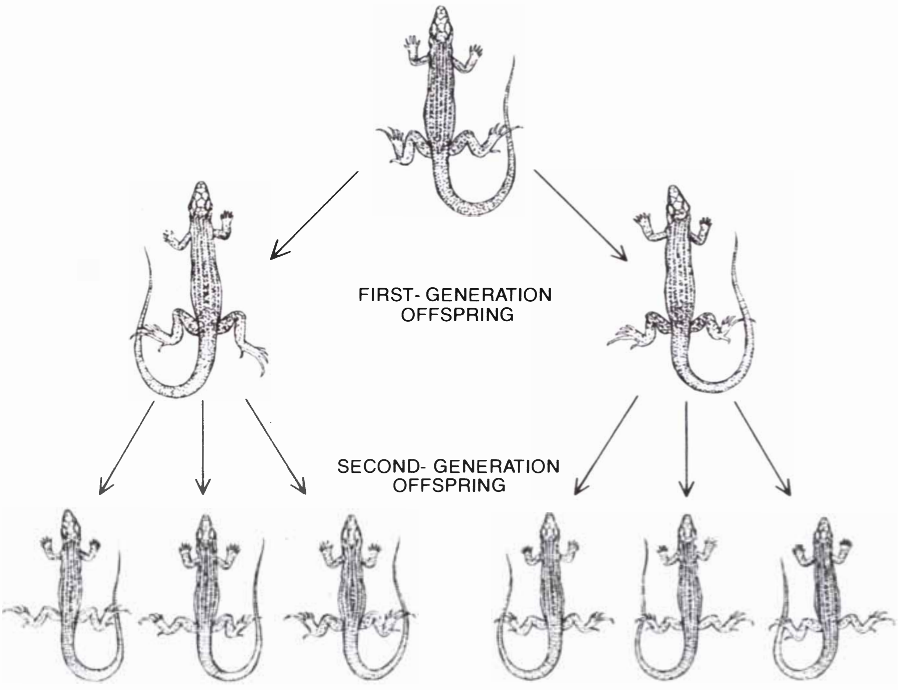


BLOOD PROTEINS OF THREE WHIPTAIL SPECIES were analyzed by electrophoresis on a gel; the result of the analysis is shown here schematically. Samples of blood plasma mixed with a radioactive isotope of iron were introduced through the six slots. After electrophoresis the radioactive iron exposed a sheet of photographic film laid over the gel, thus marking the positions of transferrin, an iron-binding protein. All these protein molecules have moved toward the positive pole of the gel. The two darker patches opposite slot *1* and slot *6* are respectively the fast-moving molecules of transferrin protein in the blood of *C. tigris* and slower-moving molecules of a structurally different transferrin in the blood of *C. inornatus.* The eight lighter patches mark the movement of molecules of the same proteins in the blood of the unisexual *C. neomexicanus.* The sample in slot *2* is from a specimen of *C. neomexicanus* taken in the field. The samples in the remaining slots are from its descendants of successive generations. Their identical displacement demonstrates clonal inheritance: the proteins, and therefore the genes that code for them, remain the same from generation to generation. Moreover, the unisexual *C. neomexicanus* expresses the two different transferrin genes present in its ancestral species. Many other proteins reveal that their genes are in the heterozygous state in unisexual species.

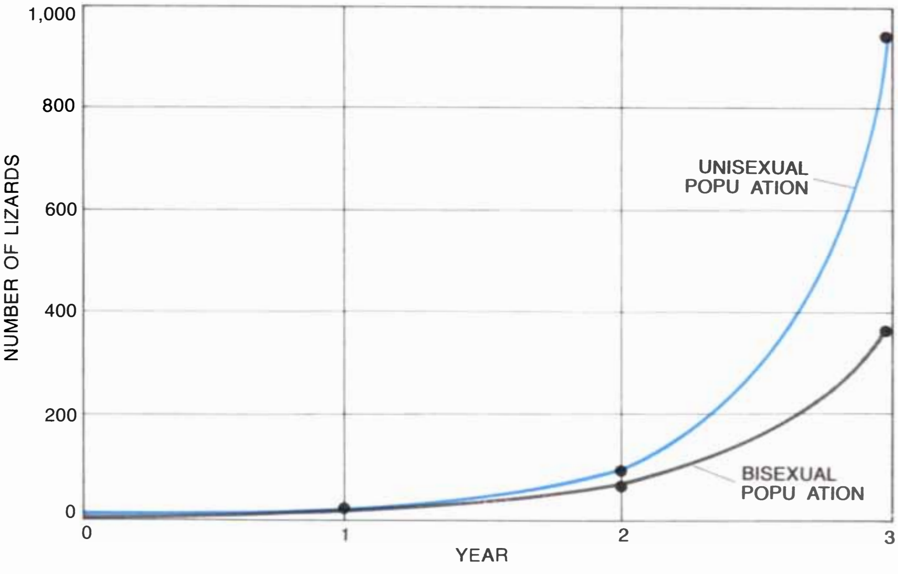


PHYLOGENETIC RELATIONS among nine species of whiptails, many of them unisexual, are outlined. The bisexual species (*black*) are related according to speciation theory: divergence from common ancestors. Here the widest divergence is between *C. inornatus* (*top*) and *C. tigris* (*bottom*). The unisexual species (*color*), derived through hybridization, show a more netlike phylogenetic pattern. Diploid unisexual species such as *C. neomexicanus* and *C. laredoensis* have arisen from a single hybridization, but triploids such as *C. uniparens* and *C. tessellatus* A-B have ancestries that include hybridizations and sometimes three bisexual species.

ards' reproductive potential: since the entire population consists of females that give rise to more females, unisexual species can reproduce in larger numbers than bisexual species, half of which cannot lay eggs. They are also at some advantage when it comes to colonizing new localities and perhaps even new habitats. For example, in the event of chance dispersal the survival of only one parthenogenetic lizard is sufficient to establish a new population in a previously uninhabited area. In some instances these last advantages might result in the establishment of parthenogenesis by means of mutation rather than hybridization; such may be the case with the South American whiptail species *C. lemniscatus,* which has been studied by P. E. Vanzolini of the University of São Paulo in Brazil.



UNISEXUAL WHIPTAILS OF THE SPECIES C. EXSANGUIS have now been raised in the laboratory through seven parthenogenetic generations. Illustrated here are the initial parent and some first- and second-generation offspring. All the offspring are females genetically identical with one another and with their parent. The six second-generation offspring are not identical with others in coloration because the adult markings do not appear until later.



HIGH REPRODUCTIVE POTENTIAL is one advantage of unisexual lizards. This graph compares the population growth of unisexual whiptail lizards (*color*) with the population growth of bisexual whiptails over a three-year period. The data are based on egg production by a laboratory-reared specimen of *C. exsanguis.* The point at year 0 on the graph represents the hatching of one egg. For the bisexual species an extra male would have to be present briefly for fertilization in year 1. At the end of the first year, assuming 100 percent hatching and survival, both populations would number nine. At the end of the third year, however, the unisexual lizards would be more than twice as numerous as the bisexual ones. The reason is that all the unisexual lizards deposit eggs whereas only 50 percent of the bisexual lizards do so.

At least in principle, however, not everything is advantageous for a parthenogenetic species. Clonal inheritance is a plus under fixed environmental conditions because it preserves gene combinations that are highly adaptive to those conditions. Environmental conditions, however, are rarely fixed. When they change quickly, as they often do, bisexual species are more likely than unisexual ones to produce some variant offspring that can adapt to the new conditions. The long-range probability of survival therefore may be less favorable for species that reproduce by cloning than for those that reproduce sexually with the features of Mendelian inheritance.

Still, not all the evolutionary cards are stacked against parthenogenetic species. Habitats often shift rather than being totally eliminated. Under such circumstances cloning lizards need only shift along with their habitat to remain within it. Furthermore, if any advantageous mutation were to occur in a unisexual lineage, cloning would ensure that the mutation would be transmitted to successive generations. Cloning may even confer some protection against disadvantageous mutations, particularly if they are recessive and occur in triploid individuals.

Twenty-five years ago a few biologists courageously raised the possibility that unisexual species of lizards existed. At that time no one would have predicted that such investigations would lead to the surprising discoveries of parthenogenetic cloning, polyploidy and the origin of new species of vertebrates by means of hybridization. Today interested workers find themselves in a position not only to ask new questions and design new experiments but also to utilize these specialized organisms in ways that could not have been imagined a few years ago. Among the possibilities that come to mind are gaining a better understanding of the role of sperm in fertilization, clarifying how it is that some animals are quite successful with multiple copies of their genes whereas others are not, studying switching mechanisms in embryonic development, producing cloned animals of known genetic composition for biological experimentation and even inducing cloning in normally bisexual species to increase productivity in animal husbandry. If unisexual reptiles contribute to progress in any or all of these areas, it will have begun with a few startling observations concerning a form of wildlife that practically no one considered significant.

Thunderbird Turbo Coupe. The understatement of the year.

"To say that this car is the best new Thunderbird in years is a dramatic understatement."

Road & Track,
January 1983

"The new Turbo Coupe is a most serious effort at returning to the concept of a 'personal car.' An enthusiast's performance GT."

AutoWeek,
January 1983

"Ford's new Thunderbird Turbo Coupe is pure driving excitement."

Popular Mechanics,
April 1983

"The Turbo Coupe's handling is remarkable..."

Motor Trend,
February 1983

"Ford had done its homework on this car, and we found it exciting, responsive, and fun to drive."

Motor Trend,
June 1983

"The new Thunderbird is Detroit's design triumph of the year."

Car and Driver,
July 1983

"There are absolutely perfect 'articulated' front seats with adjustments galore, excellent grip and comfortable support—and they look good."

Road & Track,
January 1983

Naturally, we are excited about what the experts had to say about Thunderbird Turbo Coupe. But we aren't surprised. From the Goodyear Eagle HR performance radials, to the gas-filled struts and shocks, to the quick-ratio power rack and pinion steering, the Thunderbird Turbo Coupe was designed to handle as well as many of the world's best sport coupes. And, its 2.3 liter turbocharged engine, when mated with either the 5-speed manual or optional 3-speed automatic transmission, delivers enough power to move the Turbo Coupe's aerodynamic shape quickly down the road. This combination of performance, handling and design makes the Turbo Coupe a true world class touring car.

"We trust the public will know a good thing when it sees one."

Car and Driver,
October 1982

Have you driven a Ford... lately?

Ford

The Control of Ribosome Synthesis

The particle of the living cell that translates RNA into protein is itself made up of three RNA's and 52 proteins. How is the assembly of ribosomes from these molecules adapted to the needs of the cell?

by Masayasu Nomura

The ribosome is the organelle of the living cell where proteins are made. On it amino acids are joined one at a time to a growing protein chain. The completed chain folds itself up into a molecule that can serve as part of the cell's structure or as an enzyme: the catalyst of a specific metabolic reaction. Since all such metabolic functions require enzymes, it follows that in order for a cell to grow and divide faster it must increase its capacity to make proteins. The ribosome, however, has a limited capacity for adding amino acids to the protein chain. For example, in the bacterium *Escherichia coli* each ribosome can add 15 amino acids per second at 37 degrees Celsius. If the temperature remains constant but the medium in which the *E. coli* live is enriched with nutrients, the bacteria begin to divide at a higher rate. Since the capacity of the ribosome is limited, if the cell is to increase the rate of protein synthesis, it must make more ribosomes. The ribosome is energetically expensive to manufacture, and making a large surplus of ribosomes would be wasteful. Therefore the bacterium must be able to adjust its ribosome output with considerable sensitivity.

Controlling the synthesis rate of ribosomes is by no means simple, because the organelle is a complex apparatus. Each ribosome includes 52 different proteins and three different molecules of RNA. The genetic code for the ribosomal proteins and the ribosomal RNA's (rRNA's) is carried by the DNA in the bacterium's single chromosome. In the making of the proteins the code is first transcribed into a molecule of messenger RNA (mRNA). The mRNA is then translated into protein on ribosomes made earlier. The rRNA's are produced directly from the chromosome by transcription. If the appropriate rate of ribosome synthesis is to be ensured, the rate of manufacture of each component must be controlled in relation to the rate for each of the other 54 components, much as automobile parts must be made in the right quantities and delivered to an assembly line at the right time for the automobile to be assembled with maximum efficiency. Furthermore, it is apparent that control could be exerted on at least two levels: the transcription of the genes for the 55 ribosomal components or the translation of the mRNA's for the 52 ribosomal proteins.

Perhaps the most intuitively obvious method of control is for the transcription of the ribosomal genes to be regulated directly. Indeed, in the early 1970's, when work on control of ribosome synthesis was begun, that was the main model of genetic control. It quickly became apparent, however, that direct control could not account for all of the observed pattern of variation in the rate of ribosome assembly. Further work, which is the subject of this article, showed that the synthesis of ribosomes is controlled by a system in which a fundamental role is played by two negative-feedback loops. As long as ribosomal-RNA molecules are available the ribosomal proteins bind to the rRNA's and the assembly of ribosomes continues. When no rRNA is available to be bound, certain ribosomal-protein molecules function as "translational repressors": by binding to the protein messenger RNA's they interrupt the translation of all the ribosomal proteins. That is the first feedback loop.

The second loop modulates the transcription of the ribosomal-RNA genes and certain other genes. Here the feedback control is exerted by free ribosomes: ribosomes that are not engaged in making protein. When free ribosomes are present, they turn off the rRNA genes. When the environment becomes richer in nutrients, the free ribosomes begin to make protein and the rRNA genes are turned on again. The combined action of the two interlocking feedback mechanisms enables a bacterium to adjust with great precision the rate at which ribosomes are made. Like many other fundamental scientific investigations, the elucidation of this system of genetic control has more than one outcome. The repression of translation by means of a feedback mechanism could provide a model for the control of the manufacture of other bacterial structures built up from several molecules. Furthermore, the results from *E. coli* could have implications for the control of ribosome synthesis in the more complex cells of multicellular organisms, including those of human beings. And gaining an understanding of such an intricate and elegant system of biological regulation yields an aesthetic satisfaction that is by no means the least significant fruit of such work.

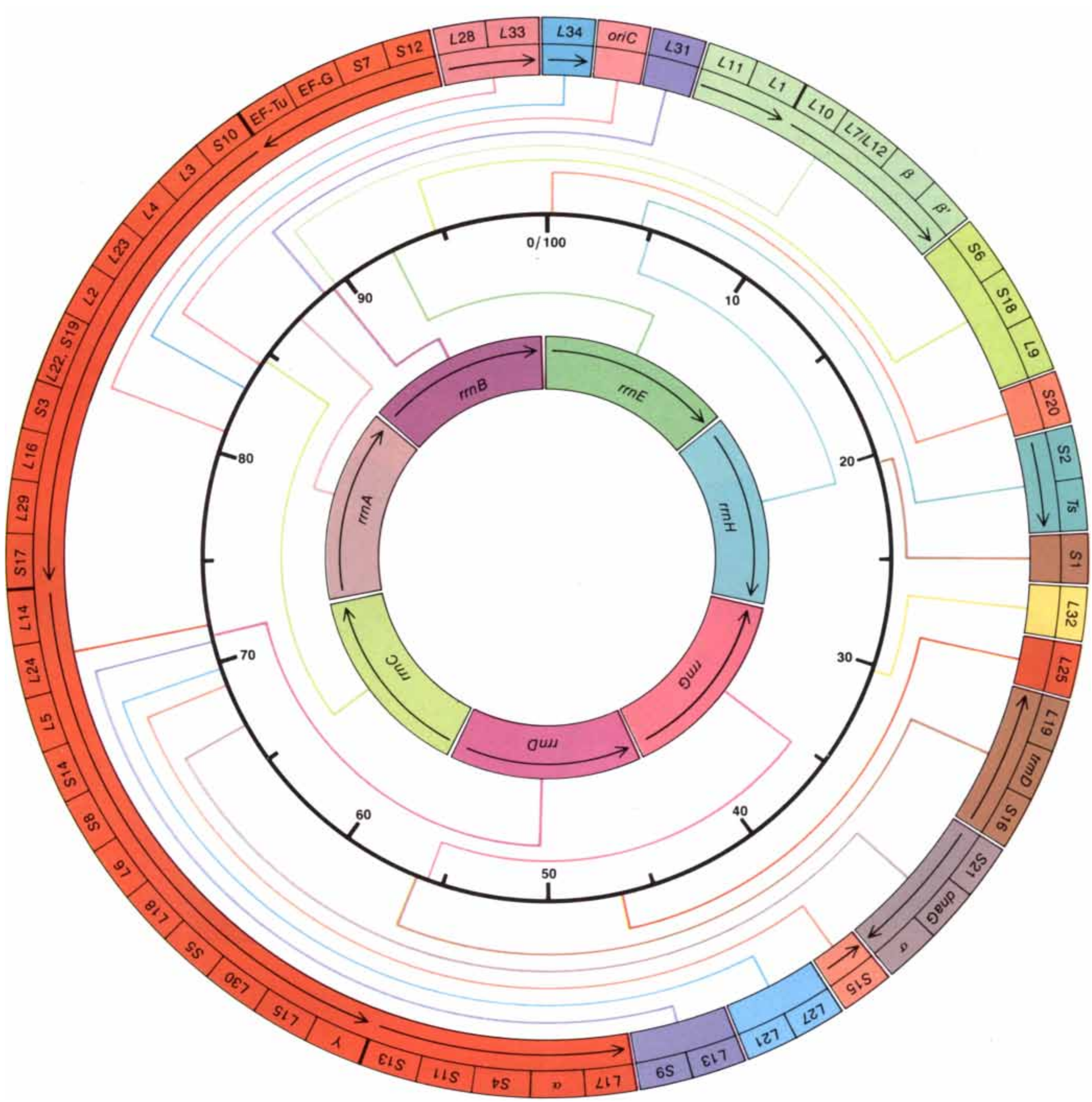
The structure, function and mode of assembly of the ribosome are so closely connected to the processing of genetic information that it is necessary to understand the basis of the genetic code in order to comprehend how ribosome synthesis is regulated. DNA, the primary genetic material, is a chain of four types of nucleotide bases. The bases are adenine, cytosine, guanine and thymine, abbreviated *A, C, G* and *T.* Because of their chemical structure *A* can generally form a bond only with *T,* and *C* can form a bond only with *G.* This restricted bonding, or complementarity, is the basis of the joining of two strands of DNA in a double-strand helix, such as the helix in the single chromosome that carries the genetic information in *E. coli.*

The *E. coli* chromosome is a piece of double-strand DNA in the form of a loop some four million nucleotides long. The sequence of nucleotides along the loop embodies the code for all the proteins required in the metabolism of the bacterium and also for several kinds of RNA. Either strand of the DNA helix can carry the coding information. The stretch of nucleotides that includes the code for a single product, whether it is a protein or a ribosomal-RNA molecule, is a gene. The genes for the 55 components of the ribosome scattered along the loop represent about 5 percent of the total information content of the *E. coli* DNA.

The sequence of nucleotides in a gene is transcribed into RNA by one of the enzymes called RNA polymerases. The polymerase binds to the DNA at a specific site and moves along the DNA strand, reading the code and catalyzing the addition of one nucleotide at a time to a growing RNA chain. Like DNA, RNA is a sequence of four nucleotides. The subunits in RNA are the same as those in DNA with one exception: in RNA uracil (*U*) takes the place of thymine. Like thymine, uracil is complementary to adenine.

When the polymerase moves down the DNA strand, the nucleotide added to the RNA is the one complementary to the nucleotide in the DNA. As a result the transcribed RNA is a negative replica of a stretch of the bacterial chromosome and as such preserves the genetic information encoded in that gene. When transcription is complete, the newly made RNA molecule is freed from the DNA.

If the transcribed RNA is messenger RNA, ribosomes quickly attach themselves to it and begin to translate



GENES FOR RIBOSOMAL COMPONENTS are shown in schematic form on a map of the chromosome of the bacterium *Escherichia coli*. The bacterial chromosome is a double-strand loop of DNA made up of four million of the subunits called nucleotide bases. The middle circle (*black*) is a map on which each division stands for 40,000 bases. A ribosomal gene is a stretch of nucleotides that bears the genetic code for one component of the ribosome: either a molecule of protein or a molecule of ribosomal RNA (rRNA). The genes for the 52 ribosomal proteins are shown in the outer circle. The proteins indicated by an *S* belong to the small subunit of the ribosome; the proteins indicated by an *L* belong to the large subunit. There are several nonribosomal genes among the genes for the ribosomal components. *OriC* is the point where replication of the bacterial chromosome begins. Genes shown in a single colored unit are transcribed together into a molecule of messenger RNA (mRNA), which is later translated into protein. The genes for the three rRNA molecules are shown in the inner circle. There are seven copies of each rRNA gene. The 21 genes are arranged in seven sets of three genes. For ribosome synthesis to be regulated the output of each ribosomal component must be adjusted in relation to the output of all 54 of the other components.

the encoded message. Short sequences of nucleotides in the messenger serve as signals for the initiation and termination of translation. When the ribosome attaches itself to the mRNA, the coding sequence in the messenger is read three nucleotides at a time. Each "codon," or group of three successive nucleotides, codes for a single amino acid in the protein chain.

A small RNA molecule called a transfer RNA (tRNA) serves as a hook to add the amino acid to the protein. Each tRNA is specific to one amino acid unit. Three bases on the tRNA complementary to the codon recognize the codon and bind to it. At the same time the amino acid on the other end of the tRNA is affixed to the end of the growing protein chain. When the ribosome advances to the next codon, the tRNA is ejected and replaced by a new one corresponding to the next amino acid.

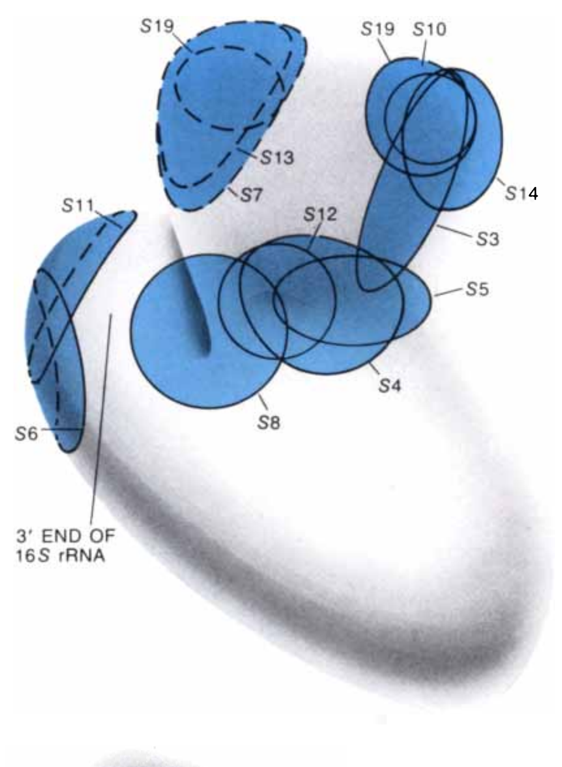
Like other bacterial cells, the *E. coli* is prokaryotic: it lacks a well-defined nucleus. In a prokaryotic cell transcription and translation are closely coupled. In many instances ribosomes begin to translate the messenger RNA before it has been freed from the DNA. After the mRNA has been freed its coding sequence continues to be translated. In general, however, the messenger lasts for only a few minutes before it is degraded by enzymes. In this period the coding sequence is typically translated between 10 and 20 times. Hence each mRNA can give rise to more than one protein molecule, with the exact number depending on how long the messenger lasts and on the rate of translation.

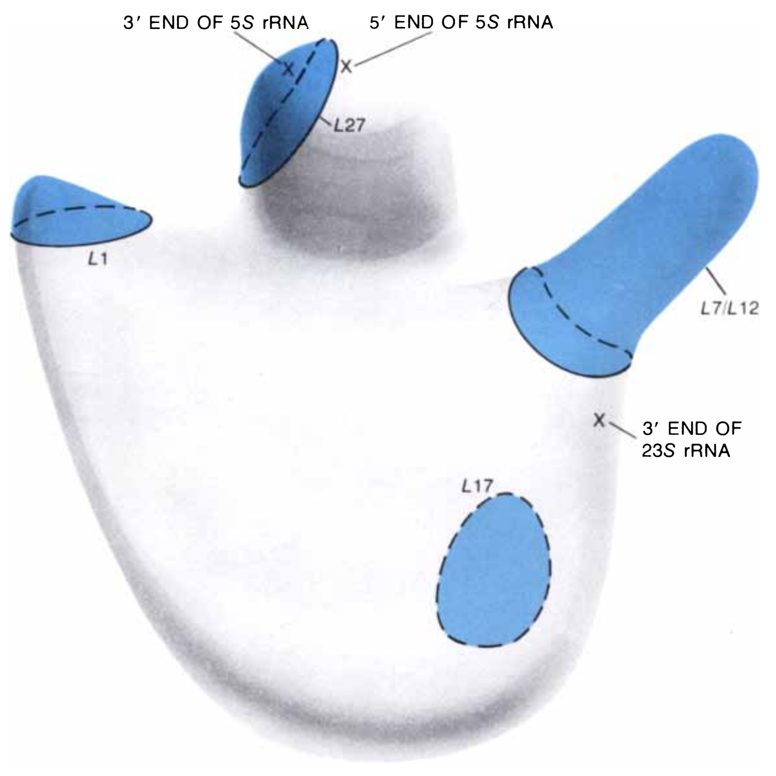
The cells of multicellular organisms, including those of human beings, are eukaryotic: they have a well-defined nucleus that is separated from the surrounding cytoplasm by the nuclear membrane. In eukaryotic cells transcription is accomplished in the nucleus and the messenger is transported to the cytoplasm, which is the site of translation. In such cells ribosomes are assembled in the nucleolus, an organelle within the nuclear membrane.

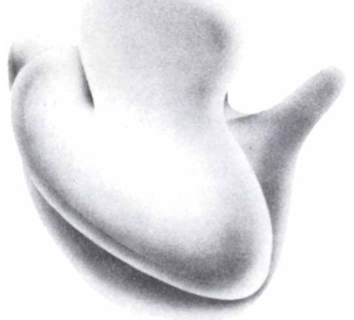
How do the fundamental genetic processes operate in the bacterial cell in the construction of a ribosome? The ribosome of *E. coli* has beên studied more intensively than that of any other organism, and the answer to the question is now emerging. Each *E. coli* ribosome consists of two subunits with different sizes and shapes: the 30*S* subunit and the 50*S* subunit. *S* stands for Svedberg unit, a measure of how fast a particle sinks when it is spun at high speed in a centrifuge. The value defined in this way is the sedimentation coefficient, and it is a measure of the size of the particle.

The 30*S* subunit, the smaller of the two subunits, includes one RNA molecule, the 16*S* ribosomal RNA, which is 1,542 nucleotides long. The other components of the subunit are proteins. There is one molecule of each of 21 different proteins. The proteins are designated *S*1 through *S*21, and here the *S* stands not for Svedberg unit but for small subunit. The single molecule of rRNA and the 21 proteins are attached to one another to form a particle with a complex shape and a molecular weight of 900,000 daltons.

The 50*S* subunit, the larger of the two subunits, includes two ribosomal RNA's: the 23*S* molecule and the 5*S* molecule. The two rRNA's respectively



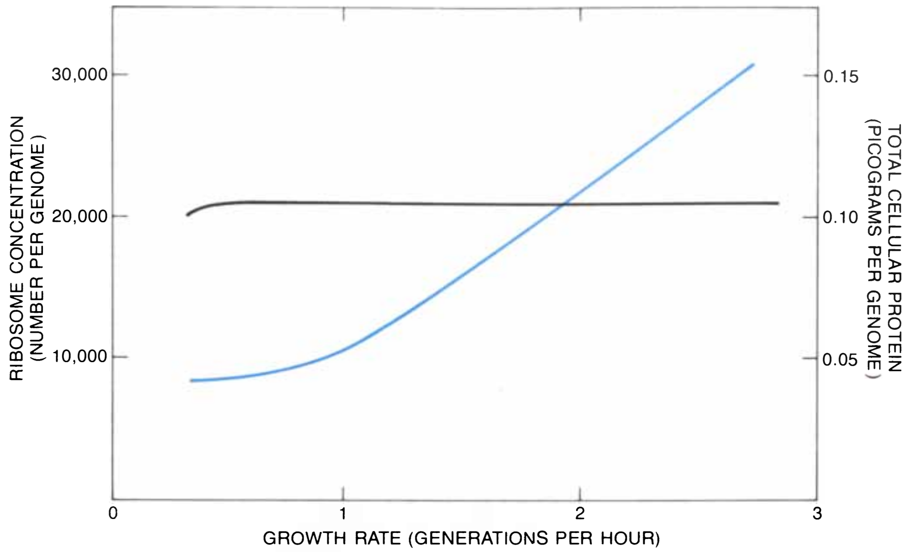




SUBUNITS OF THE RIBOSOME are shown with the location of some of their components mapped on them. Each subunit is made up of protein molecules and ribosomal RNA. The small subunit, designated 30*S*, is at the left. The large subunit, designated 50*S*, is at the right. The small subunit includes 21 protein molecules, all present in one copy each. The large subunit includes 34 protein molecules, all present in one copy each except for one molecule, which is present in four copies. The colored regions indicate where the proteins whose location is known intersect the surface of the subunit. The map was made on the basis of results obtained by immune electron microscopy; the colored areas are the sites on the surface of the ribosome where antibodies bind to the protein against which they were primed. The mapping was initiated by the author and his colleagues Lawrence Kahan and William A. Strycharz with James A. Lake of the University of California at Los Angeles; it was continued by Kahan and Lake. There is one molecule of rRNA in the small subunit and two in the large subunit. The *X*'s indicate where the ends of the rRNA's appear on the surface. The functional ribosome is at the left.

are 2,904 nucleotides long and 120 nucleotides long. The rest of the subunit consists of 34 protein molecules. The proteins are designated *L*1 through *L*34, with *L* standing for large subunit. Although there are 34 protein molecules in the 50*S* unit, there are only 31 types of protein. One protein is present in four copies per ribosome. The others are present in one copy each. The numbering system, however, extends to 34 because of misassignments in early work on the ribosome. The proteins and the two rRNA's are joined in a flattened spheroid that has several irregular protrusions on its upper surface. The large subunit has a molecular weight of 1.6 million daltons.

When the 50*S* subunit is attached to the 30*S* subunit, they form the functional ribosome, which has a sedimentation coefficient of 70*S* and a molecular weight of 2.5 million daltons. (The sedimentation coefficients are not additive because the shape of the particle influences how fast it sinks when it is spun in the centrifuge.) In the functional organelle there is a groove between the large and small subunits into which the messenger RNA and enzymes fit in the course of translation.



RIBOSOME CONCENTRATION in *E. coli* (*colored line*) rises with the rate of cell growth and division; the relation is a linear one. To grow and divide faster the bacterium must manufacture protein molecules more rapidly. Since each ribosome has a limited capacity to make proteins, *E. coli* must make more ribosomes in order to grow faster. Making surplus ribosomes and storing them would be wasteful; hence the bacterial cell must be able to adjust its ribosome output precisely. The protein concentration (*black line*), however, does not increase with the growth rate. Both ribosome concentration and protein concentration are expressed in relation to the mass of the bacterial genome (the complete set of genes). Rapidly growing cells have a larger volume, but the ratio of protein to DNA remains the same as in slower-growing cells.

Some of the earliest quantitative observations of the ribosome concerned the amount of RNA in the cell at different rates of growth and division. In the late 1950's Ole Maaloe, Niels O. Kjeldgaard and their colleagues at the State Serum Institute and later at the University Institute of Microbiology in Copenhagen examined the cellular contents of the bacterium *Salmonella typhimurium*. They cultured *Salmonella,* which is a close relative of *E. coli,* in various media and found there is a linear relation between the rate at which the bacteria divide and their total RNA content. At the time it was already known that most of the RNA in the cell is ribosomal RNA. Thus measuring the total RNA content yields an approximation of the concentration of ribosomes in the bacterium.

Subsequent work has shown that there is indeed a linear relation between the growth rate and the quantity of ribosomes in the cell: the faster the bacteria are dividing, the more ribosomes there are. (The linear relation does not hold, however, at very low growth rates.) The linear relation suggested to workers in the field that the control of ribosome synthesis is crucial to the regulation of cell growth and division. Indeed, some bacterial physiologists proposed that the ultimate limit on the rate of growth is the bacterium's capacity to make ribosomes. As we shall see, newer results strongly suggest this hypothesis is false. Nevertheless, the work by the Copenhagen school served to arouse considerable interest in the question of how ribosome synthesis is regulated.

In the 1960's and the early 1970's I worked on the structure and function of the ribosomes and on how the organelle is assembled from its many components. As I was studying how the ribosome is put together I became interested in the even more complex problem of how the manufacture of ribosomal components is regulated. When the problem of control is considered, several questions arise quite naturally. Among the most important of them is the question: Is the manufacture of the parts of the ribosome balanced and coordinated?

By balanced I mean that the components are produced in the same proportions that are present in the finished product. As a much simplified analogy, consider an industrial product with two parts. If 50 units of the finished product are desired, one way to ensure that this quantity is made is to manufacture 50 of each part and begin the assembly process. Assuming that no parts are broken along the way, when 50 units of the product have been made, assembly will come to a halt for lack of both parts. This is balanced control.

Coordinated regulation, on the other hand, means that the manufacture of the components is adjusted so that the ratio of the parts remains constant when the total volume of production increases or decreases. Hence in the example of balanced control given above if it were decided to make 100 of the finished product instead of 50, coordinated regulation would require that 100 of each component be made.

It can now be asked whether the manufacture of the 52 ribosomal proteins is balanced and coordinated so that the proteins are always made in the proportions in which they appear in the finished ribosome. It can also be asked whether the synthesis of the proteins is balanced and coordinated in relation to synthesis of the ribosomal RNA's, so that all the components of the ribosome are always made in equal amounts. As an alternative, the method of regulation could be unbalanced, with some components made in excess and later degraded by enzymes.

These were the basic questions that first arose when the regulation of ribosome synthesis was considered in detail. In its first stages the work concentrated on the ribosomal proteins. At about the time the problem of regulation was taken up E. Kaltschmidt and Heinz-Günter Wittmann of the Max Planck Institute for Molecular Genetics in Berlin devised a simple method for separating most of the ribosomal proteins. This method and others developed later were quite helpful in examining the synthesis of the individual ribosomal proteins at different growth rates.

In 1974 Patrick P. Dennis, a postdoctoral fellow in my laboratory at the University of Wisconsin at Madison, exploited the new methods of protein separation to observe the manufacture of the ribosomal proteins. He found that the synthesis rates of most of the proteins are balanced with respect to one another and that the rates are controlled in a coordinated manner in response to changes in the composition of the growth medium. With this knowledge we could approach the problem of regulation directly.

From very early in the work on the ribosome it had been apparent to me that in order to study the regulation of synthesis in molecular terms it was necessary to identify the genes for the ribosomal proteins and the ribosomal RNA's. Once the DNA of these genes could be identified and isolated it could be employed in experiments where proteins and rRNA are made in cell-free laboratory systems. Once such an in vitro system had been devised biochemical techniques would make it possible to measure the rate at which the ribosomal genes are expressed and to elucidate the control mechanism. The isolation of the genes for the 55 ribosomal components, however was a difficult task and took a long time to accomplish. Recombinant-DNA methods for manipulating genetic material had not yet been devised and the work had to be based on the methods of classical genetics.

Bacteriophage lambda, a virus that infects bacteria, was central to the investigation. When the phage invades a bacterial cell, the virus DNA is inserted into the chromosome of the host. After the insertion the phage can coexist with the bacterium quite peacefully for long periods. Certain triggering events, however, such as an injury to the host's DNA caused by ultraviolet radiation, can cause the phage to take over the bacterium's genetic apparatus and begin to reproduce itself.

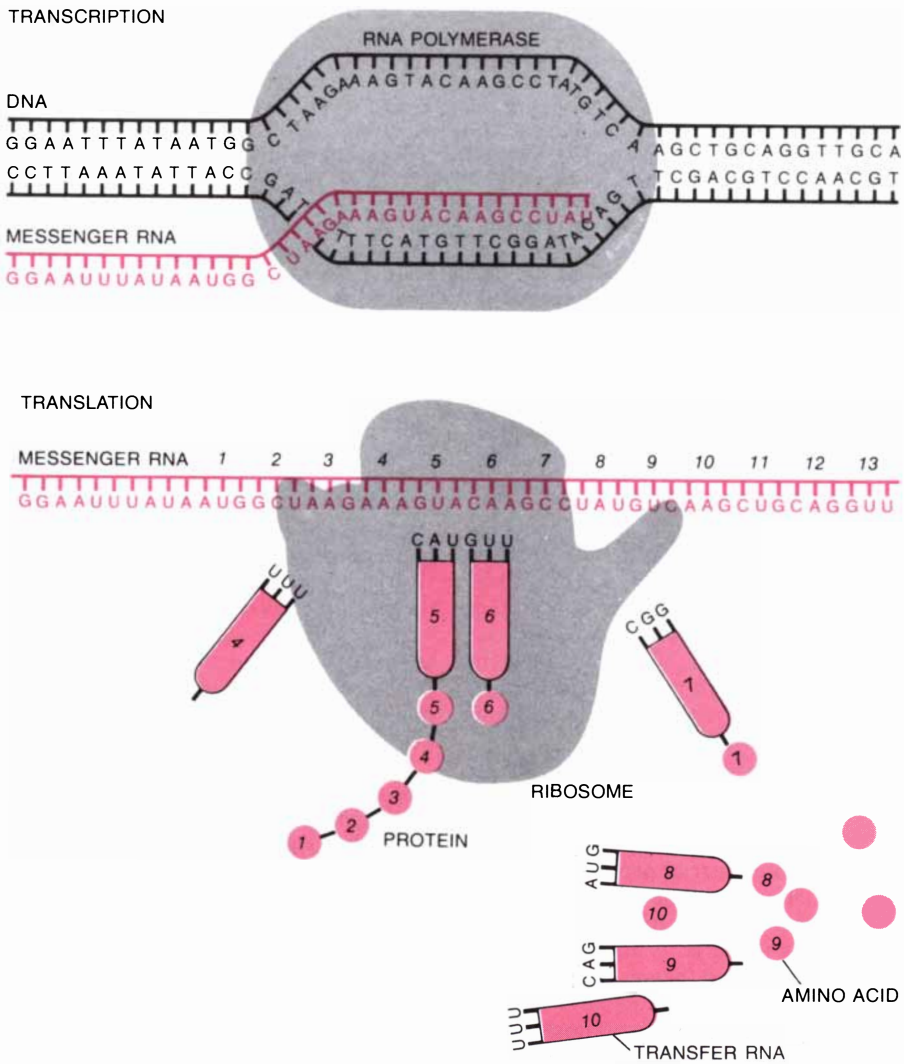
The first step in the reproduction of the virus is for the phage DNA to be excised from the bacterial chromosome by enzymes. In some instances the excision is done incorrectly and pieces of the bacterial chromosome adjacent to the phage DNA are removed along with the phage DNA. The excised DNA is then incorporated into a new virus particle. If both virus DNA and bacterial DNA are incorporated, the phage can serve to "transduce" the bacterial DNA into another bacterium. By the time the attempt to isolate the ribosomal genes was begun there was evidence that many of the ribosomal-protein genes are clustered in one segment of the *E. coli* chromosome. The region offered a ready target for a transducing phage.

In late 1974 S. Richard Jaskunas, Lasse A. Lindahl, a postdoctoral fellow in my laboratory, and I succeeded in directing the phage to integrate into the bacterial DNA near the cluster of ribosomal-protein genes. We then isolated several phage particles containing one or more ribosomal-protein genes. The number of ribosomal-protein genes in the virus particle varied considerably: one phage was shown to carry the genes for 27 ribosomal proteins, more than half of the full complement.

This was the first breakthrough in the study of the regulation of ribosome synthesis. The significance of this work lay in the fact that it made possible the construction of the test-tube systems for studying how the manufacture of the parts of the ribosome is regulated. Since 1974 ribosomal-protein genes from other regions of the *E. coli* chromosome have been isolated by other investigators as well as by workers in my laboratory.

While the ribosomal-protein genes were being studied in the mid-1970's the genes for the ribosomal-RNA molecules were also being examined, notably by Norman R. Davidson and his colleagues at the California Institute of Technology and in several other laboratories, including my own. In contrast to the genes for the ribosomal proteins, which exist in a single copy each, there are seven copies of each of the three rRNA genes on the *E. coli* chromosome. The genes in each set are transcribed together into a single RNA, which is processed to yield the 16*S*, 23*S* and 5*S* molecules. By 1977 all seven sets of rRNA genes had been isolated, and the effort to determine their structure and nucleotide sequence was under way.

With the ribosomal-RNA genes and



TRANSCRIPTION AND TRANSLATION are the fundamental genetic processes in the making of a ribosome. Transcription (*top*) is the copying of the genetic information in DNA into RNA. The four nucleotides in DNA are adenine, cytosine, guanine and thymine (*A, C, G* and *T*). In RNA uracil (*U*) takes the place of thymine. *A* can form a bond only with *T* or *U*, and *C* can form a bond only with *G*. Bases that can form a bond are said to be complementary. In transcription the enzyme RNA polymerase binds to the bacterial chromosome and temporarily separates the strands of DNA. The enzyme then uses one strand as a template for adding complementary nucleotides to a growing chain of messenger RNA or ribosomal RNA. Translation (*bottom*) is the joining of amino acids into a protein according to the information carried by the mRNA. Each codon, or group of three nucleotides in the mRNA, codes for one amino acid. The small molecules called transfer RNA's (tRNA's) have anticodons: three bases complementary to the codon. Each tRNA is specific to one amino acid. The ribosome moves down the mRNA reading the codons in turn. The anticodon on the tRNA binds to the codon and the amino acid on the other end of the tRNA is attached to the growing amino acid chain. When the ribosome moves on to the next codon, the tRNA is ejected and a new one takes its place.

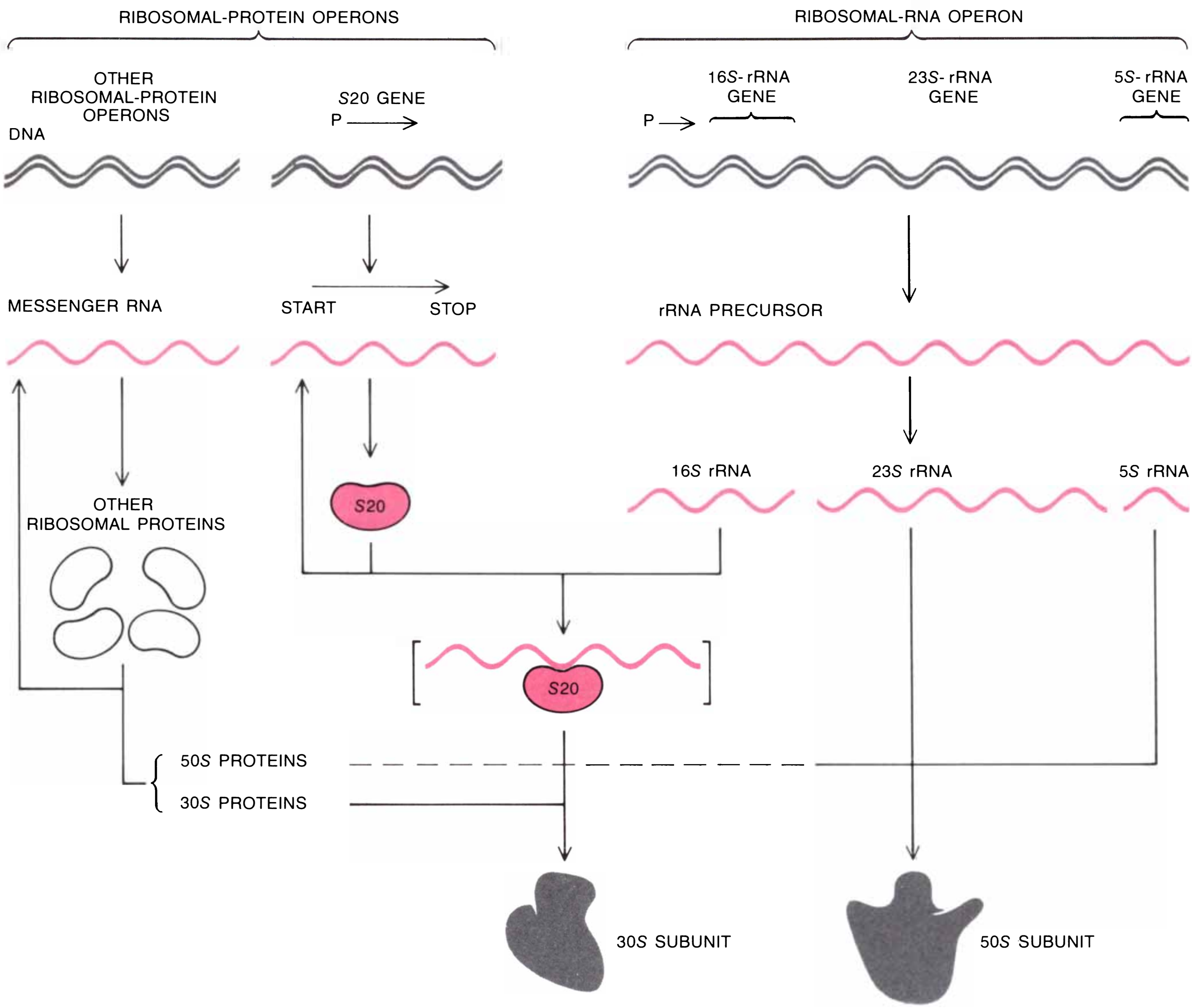
the majority of the ribosomal-protein genes purified, the question of how the ensemble of genes for the 55 ribosomal components is regulated to yield balanced and coordinated manufacture of the components of the ribosome could be taken up directly. At the time the most widely accepted concept of genetic regulation was the operon theory, which had been developed by Jacques Monod and François Jacob of the Pasteur Institute in Paris in the early 1960's to explain how *E. coli* cells control the metabolism of lactose.

E. coli can employ several molecules as a source of carbon, among them the sugars lactose and glucose. The utilization of lactose requires the enzyme beta-galactosidase. If the bacteria are grown in a medium that includes lactose as a carbon source, beta-galactosidase is found in the cell. The utilization of glucose, however, does not require beta-galactosidase, and if the bacteria are grown in a medium that includes glucose as the source of carbon, little beta-galactosidase is made.

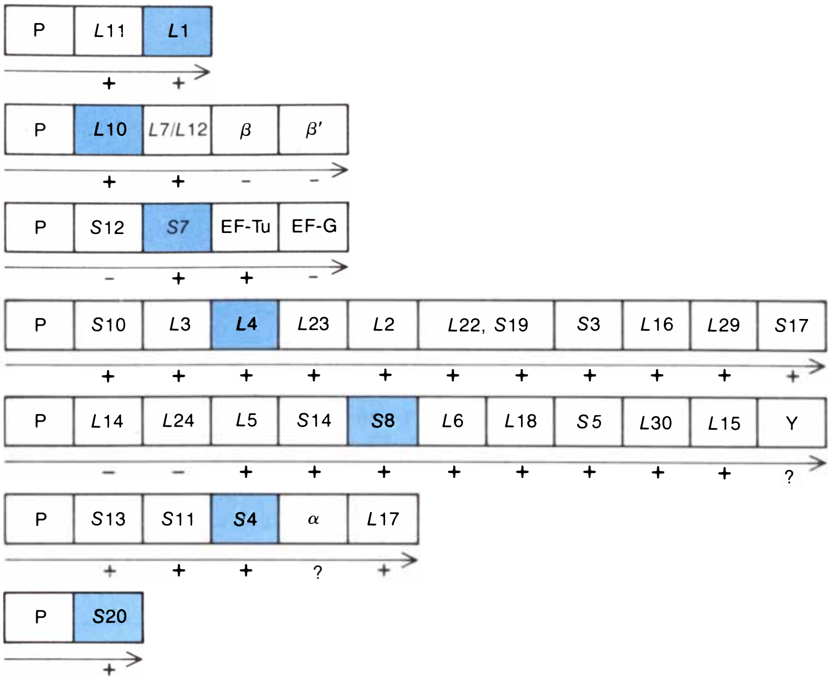
Monod and Jacob showed that three genes participate in lactose metabolism, one of which is the gene for beta-galactosidase. The genes form a genetic unit known as the lactose operon, or lac operon. They are next to one another on the bacterial chromosome and are transcribed together. The transcription of the lac operon is begun when a polymerase enzyme binds to a stretch of nucleotides called the lac promoter, which is near one end of the operon.

Close to the promoter is a group of nucleotides known as the operator. A protein called the lac repressor can form a bond to the operator site. When the repressor is bound to the operator, the polymerase cannot attach itself to the promoter and the lac operon cannot be transcribed. This is the situation when there is no lactose in the bacterial cell.

When molecules of lactose enter the bacterium, however, they are quickly modified. One of the modified forms, which is closely related to lactose in structure, acts as an inducer: by binding to the lac repressor it prevents the repressor from occupying the operator site. As a result the promoter is freed to interact with the polymerase and transcription begins. Thus the presence of lactose constitutes the signal to turn the



FEEDBACK LOOP enables the ribosomal protein *S*20 to regulate its own synthesis. The ribosomal-protein genes are transcribed in the groups called operons. Each operon is a unit made up of one or more genes that are transcribed together. An operon has its own promoter: the stretch of nucleotides on the DNA where transcription begins. The *S*20 gene constitutes such a unit; it is transcribed from its own promoter (*P*). Each set of three ribosomal-RNA genes also constitutes an operon. The rRNA's are transcribed together in a precursor that is divided to yield the three rRNA molecules. The *S*20 is a binding protein: in ribosome assembly it forms a complex with the 16*S* rRNA and is ultimately incorporated into the 30*S* subunit. As long as 16*S* rRNA's are available the *S*20 protein is made. When there are no 16*S* rRNA's to be bound, the excess *S*20 molecules interrupt further synthesis by stopping the translation of their own mRNA's.



TRANSLATIONAL REPRESSOR is a ribosomal protein that can stop the synthesis of itself and in some instances the synthesis of other ribosomal proteins. It does so by binding to the messenger RNA for its own operon and thereby halting translation. The translational repressors are indicated by colored boxes. Each horizontal row of boxes stands for a ribosomal-protein operon. Transcription begins at the promoter sequence (*P*); the arrows indicate the direction of transcription. In each operon the boxes marked with a plus sign correspond to proteins whose synthesis can be halted by the repressor. Not all the proteins in the operon are subject to repression in this way. For example, the synthesis of *S* 12 is not inhibited by *S* 7. It appears that *S* 12, *L*14 and *L*24 can interrupt their own synthesis. Thus the operon is divided into regulatory units. Genes for some proteins that are not components of the ribosome are included in the ribosomal operons. Alpha, beta and beta′ are subunits of RNA polymerase. *EF-Tu* and *EF-G* are "elongation factors" required for protein synthesis. *Y* has a role in protein secretion.

lac operon on and the absence of lactose constitutes the signal to turn the lac operon off.

The success of the operon model in explaining the regulation of lactose metabolism led most workers to focus their attention on the direct control of transcription. Partly as a result, in the late 1960's and early 1970's most of the known regulatory systems were based on the control of transcription by mechanisms similar to that of the original operon model. It was natural that when the control of ribosome synthesis began to be studied, an attempt would be made to apply the operon theory.

The first step in seeing whether the operon model could explain ribosome synthesis was to identify the transcription units and promoters for the ribosomal components. By then it had been established that the ribosomal-RNA genes are arranged in seven sets, each of which makes up an operon, or transcription unit. The work in my laboratory therefore concentrated on the ribosomal-protein genes. It was soon found that within the main protein-gene cluster the genes are arranged in several operons. Each operon includes one or more protein genes that are transcribed together starting at a single promoter, and each operon can be transcribed independently of the others.

Knowing the functional arrangement of the ribosomal-protein genes made it possible to propose a hypothesis about how the genes are regulated. The hypothesis, which closely followed the operon model, was based on the direct control of transcription. A central feature of the hypothesis was the assumption that all the protein operons respond to the same chemical signals (inducer and repressor molecules). The presence of the inducer would serve to turn on all the protein operons simultaneously and the absence of the inducer molecule would serve to turn them all off.

If it was true that all the ribosomal-protein operons respond to the same signals, it followed that the operator and promoter nucleotide sequences would have significant common features. The similarity of the promoters and their simultaneous response to a single inducer would explain the coordinated synthesis of the ribosomal proteins.

Although the operon theory was formulated to explain the action of a few adjacent genes, there is no reason in principle why such a mechanism might not account for the operation of many genes, even genes widely separated on the bacterial chromosome. Indeed, examples of complex operon systems were subsequently found, including one made up of genes that are activated when the *E. coli* chromosome is damaged.

For ribosome synthesis to be regulated by a mechanism of the operon type, however, would call for some strict specifications. To explain the balanced synthesis of the ribosomal proteins it would be necessary to assume that the transcription of each of the ribosomal-protein operons proceeds with the same efficiency, that all the messenger-RNA molecules made from these operons have the same average lifetime and that all the coding sequences in the mRNA are translated at the same rate. If any of these conditions were to be violated, the ribosomal proteins would be made in different amounts.

Furthermore, there is evidence that the rate of manufacture of the ribosomal RNA's and the ribosomal proteins is about the same, at least in healthy growing cells. To maintain the operon hypothesis it would have to be assumed that the promoters for the rRNA operons share features with the promoters for the ribosomal-protein operons. This would be necessary for the operons to respond in a coordinated way to the same signals. In addition, since the rRNA's are transcribed directly from the DNA, whereas the proteins come from the messenger molecule, which can be translated more than once, the efficiency of the promoters for the rRNA operons would have to be greater than the efficiency of the promoters for the protein operons, and the ratio between them would have to be such as to yield an equal quantity of the two types of components.

Such theoretical difficulties did not seem insurmountable, and in any event the operon model provided a valuable starting point. If the operon mechanism were the correct one, supporting evidence would be the similarities between the promoters for all the ribosomal operons. Fortunately for the work on the ribosome, at about the same time that the ribosomal genes and promoters were isolated two powerful new techniques for determining nucleotide sequences were developed, one by Frederick Sanger and his colleagues at the Medical Research Council Laboratory of Molecular Biology in Cambridge and the other by Allan M. Maxam and Walter Gilbert of Harvard University.

Leonard E. Post, a graduate student in my laboratory, employed the new techniques to find the nucleotide sequences of the promoters for the ribosomal-protein operons. The complete sequences for five promoters were determined. Herman A. deBoer and Scott F. Gilbert in my laboratory also determined the nucleotide sequence of the promoters for two sets of ribosomal-RNA genes, and other investigators determined the promoter sequence for other sets of rRNA genes.

The sequencing work failed to reveal any obvious common characteristics that distinguished the ribosomal promoters from other known conventional promoter sequences. Thus the hypothesis that balanced and coordinated synthesis of the ribosomal components is achieved by an operon mechanism became less appealing.

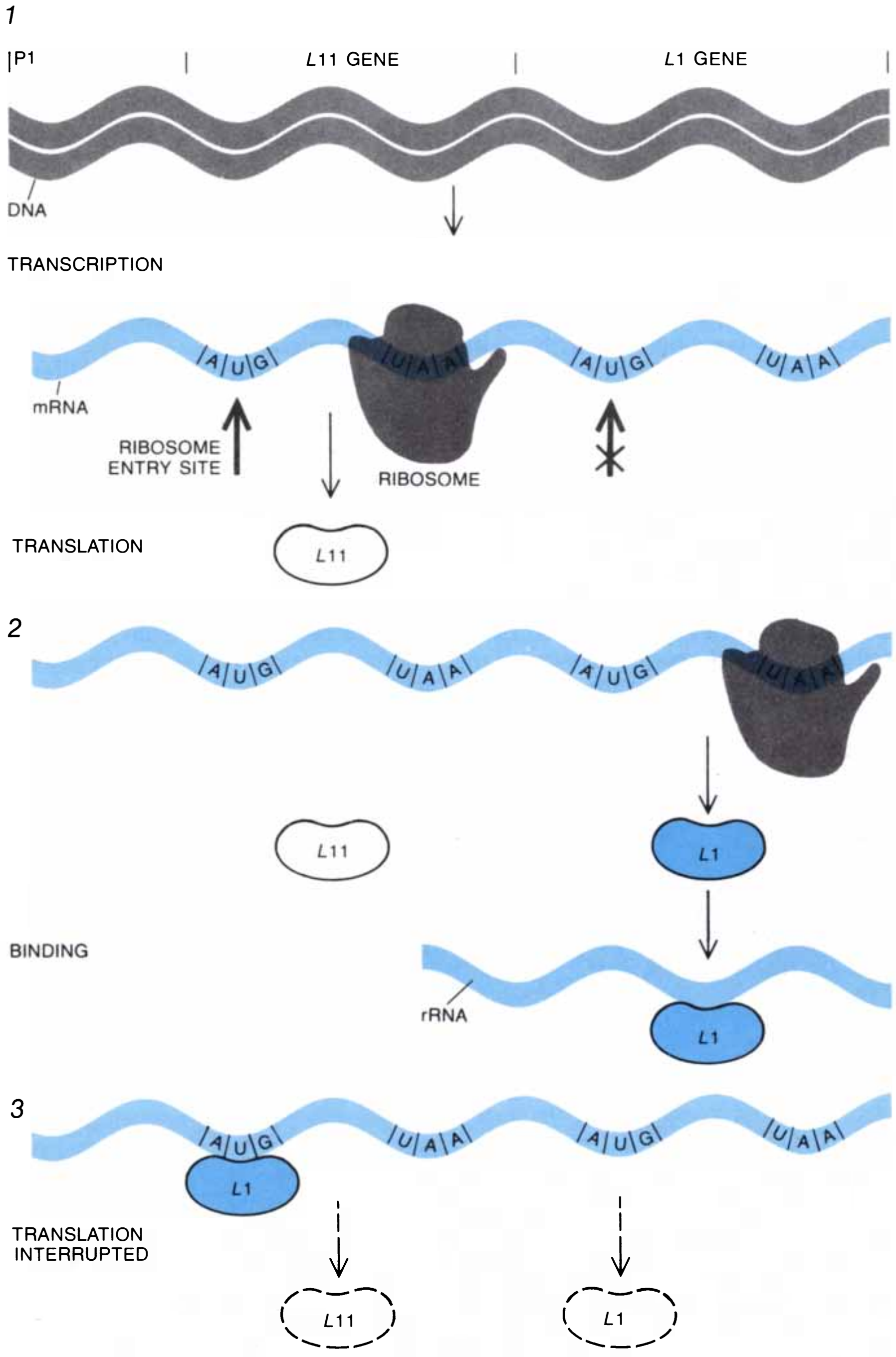
When the direct regulation of the complete ensemble of ribosomal genes was put in doubt, two alternative possibilities remained. The first was that ribosomal-RNA synthesis is regulated directly and protein synthesis is regulated as a secondary consequence of the control of rRNA manufacture. The second possibility is the converse of the first: that the ribosomal-protein genes are regulated directly and the rRNA genes are regulated as a secondary consequence of the regulation of the protein genes.

Since there are many ribosomal-protein operons, the second possibility had the same drawback as the basic operon model. It entailed the existence of many promoters with the same structure, a possibility that had already been shown to be improbable. The first mechanism, however, seemed to be promising. It appeared that a simple way to control the overall rate of ribosome synthesis was for ribosomal proteins to be made in marginally larger quantities than ribosomal RNA's. The quantity of rRNA would limit the rate of assembly of ribosomes. The very small surplus of ribosomal proteins could serve as repressor molecules to interrupt the synthesis of unneeded additional proteins. It was assumed that the regulation of ribosomal-protein synthesis was carried out at the level of transcription.

The first attempt to find out whether ribosomal-protein synthesis is controlled by a feedback system was carried out by Ann M. Fallon, Sue Jinks-Robertson and Geneva D. Strycharz, working with me. The work was designed to provide an additional and more conclusive test of the operon hypothesis. If identical promoters sensitive to the same inducer or repressor were the basis of balanced synthesis of the ribosomal proteins, then adding extra copies of a ribosomal-protein operon and its promoter to the bacterium would unbalance the manufacturing process. The proteins corresponding to the added genes would be made in larger quantities than the other ribosomal proteins.

If the synthesis is balanced and coordinated by a feedback mechanism, however, no significant effect of the increased "gene dosage" would be noted because the feedback loop could repress the action of the extra genes and their promoters just as easily as it could the action of the original genes. Hence balanced synthesis would be maintained.

Copies of some of the ribosomal-protein genes were introduced into *E. coli* by putting the genes into bacteriophage lambda or a plasmid, a circlet of DNA separate from the bacterial chromosome. As we have seen, the bacteriophage can infect bacterial cells and coexist with the bacterium. The plasmid can also be taken up by the bacterial cell. The extra copies of the ribosomal-



SEQUENTIAL TRANSLATION, in combination with the feedback repression of translation, ensures that the ribosomal proteins are always made in equal quantities. The *L*11-*L*1 operon includes a promoter and the genes for the *L*11 protein and the *L*1 protein in that order (*1*). At the beginning of each coding sequence in the messenger RNA are the nucleotides *AUG*, which serve as the signal to start translation. At the end of each coding sequence are the nucleotides *UAA*, which serve as the signal to stop translation. The ribosome can begin translation of the *L*11-*L*1 message, however, only at the beginning of the *L*11 sequence, at a point called the ribosome entry site. Only after *L*11 has been translated can translation of *L*1 begin. Each translation of *L*11 makes possible only one translation of *L*1 (*2*). When the two proteins have been made, *L*1 binds to the 23*S* ribosomal RNA as long as rRNA's are available. When the supply of 23*S* molecules has been used up, *L*1, which is a translational repressor, binds to the *L*11-*L*1 mRNA near the ribosome entry site and no *L*11 or *L*1 is made (*3*). Thus if *L*11 and *L*1 are being made, they are made in the same quantity. If translation is repressed, no *L*11 or *L*1 is made.

protein genes on the phage or the plasmid are expressed along with the information in the chromosome.

When the phage or the plasmids with the ribosomal-protein genes were put into the bacteria, the rate of transcription and translation was measured. The rate of transcription of the messenger RNA for the proteins increased in proportion to the number of copies of the gene present in the bacterial cell. The rate of translation of the proteins from the mRNA, however, was unchanged. The proteins whose genes had been added were not made in greater quantities than the proteins whose genes were present in the usual number of copies.

These results were consistent with the feedback concept of regulation. It was evident, however, that the feedback process was operating at the level of translation rather than at the level of transcription. A few examples of genetic control exerted by means of the regulation of translation were known at the time. Our findings, however, were unexpected and dramatic to workers accustomed to the concept of transcriptional control. After repeating the experiments many times we proposed the translational-feedback model as the explanation for balanced synthesis of the ribosomal proteins. Experiments similar to ours were done at about the same time by Dieter Geyl and August Böck of the University of Regensburg and also by Niels Fiil of the University Institute of Microbiology in Copenhagen, working with Patrick Dennis (who was now at the University of British Columbia). Their results were in agreement with the idea of translational feedback.

In the translational-feedback theory the free ribosomal proteins (those not bound to ribosomal RNA) serve as feedback inhibitors of the translation of their own messenger RNA's. Ribosomal-protein synthesis is coupled to the process of ribosome assembly. As long as the assembly of ribosomes, that is, the binding of the ribosomal proteins to the rRNA's, removes the free ribosomal proteins, the messenger for the proteins continues to be translated. When there is no more rRNA to be bound, the free protein stops the translation of its own messenger.

Direct proof of the translational-feedback-regulation hypothesis came from in vitro tests done by John L. Yates in my laboratory. Yates mixed the DNA bearing the code for the ribosomal proteins with ribosomes, RNA polymerase, other enzymes and other molecules needed for transcription and translation. Amino acids that had been labeled with radioactive atoms were added to the test-tube mixture as a means of detecting the proteins into which the amino acid units were assembled.

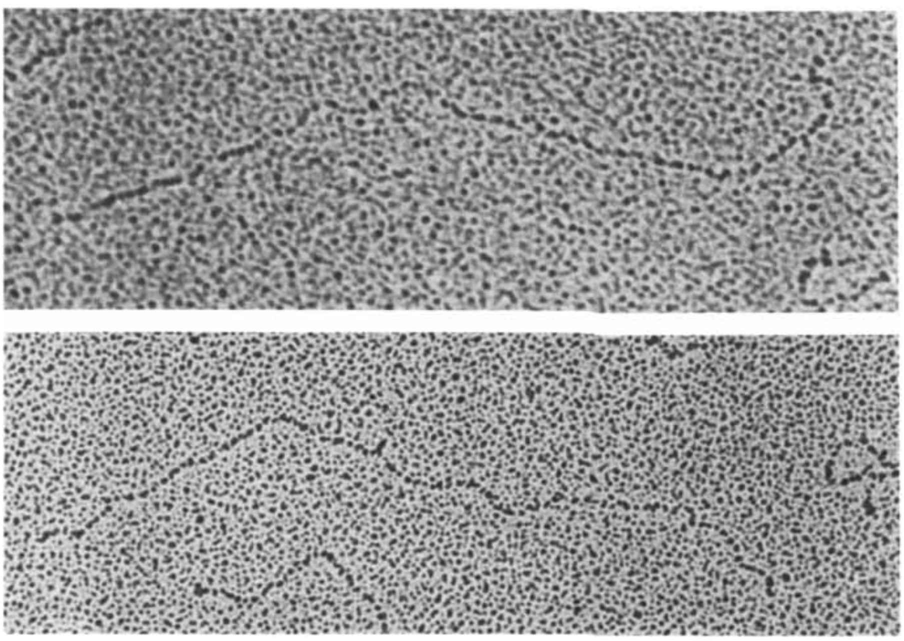
By 1979 most of the ribosomal proteins had been purified. When small quantities of the proteins that had been isolated were added to the in vitro mixture, it was striking to observe that some proteins were capable of inhibiting translation of the protein messenger RNA's. Not all the ribosomal proteins had this property, but the ones that did have it inhibited the synthesis of several other proteins in addition to themselves.

The genes for the proteins whose synthesis was interrupted were always in the same operon as the gene for the repressor protein was. (This repressor is called a translational repressor to distinguish it from the transcriptional repressors of the classical operon model.) For example, when protein *S*4 is added to the test tube, it stops not only its own translation but also that of proteins *S*13 and *S*11; all three of the proteins are in the same operon. Protein *L*1 inhibits its own synthesis and that of *L*11, with which it shares a transcription unit. The presence of protein *L*1, however, has no effect on the synthesis of proteins *S*4, *S*11 or *S*13, and the presence of *S*4 has no effect on the translation of *L*1 or *L*11.

Proteins *S*8, *S*4 and *L*1 were immediately identified as translational repressors. Soon *S*7 and *L*4 were so identified in my laboratory. At about the same time Ryuji Fukuda of the University of Kyoto and independently Herbert Weissbach of the Roche Institute of Molecular Biology and his colleagues demonstrated that *L*10 can halt the translation of its own message. Böck and Reinhard Wirth of Regensburg showed that *S*20 is also a translational repressor. Each repressor regulates the translation of all the proteins in its operon or a subgroup of those proteins. Each group of proteins regulated in this way is called a unit of regulation, and there is only one repressor in each unit of regulation. When translational repression was discovered, it appeared likely that the repressor halts translation by binding to the messenger RNA transcribed from its operon and thereby keeps the ribosome from attaching itself to the messenger and beginning translation.

Since the repressor protein can inhibit the translation of several ribosomal proteins encoded in a single messenger RNA, the repression could be achieved in two ways. The repressor could bind to a single target site in each unit of regulation on the mRNA, preventing the translation of all the coding sequences at once. Alternatively, the repressor could interact with several sites on the mRNA, with each site corresponding to a single coding sequence.

Work by Yates, Dennis Dean and Gail A. Baughman in my laboratory suggests that there is only one binding site in each unit of regulation and that the binding of the repressor to this site simultaneously interrupts the trans-



RNA MOLECULES are long, single-strand chains under chemical conditions where base pairing between adjacent regions of the chain is prevented. The upper micrograph shows 16*S* ribosomal RNA's under such conditions. When 16*S* rRNA's are mixed with the *S*8 protein, which is a translational repressor, however, the result is an RNA with a short double-strand region near the middle of the RNA chain; such molecules are shown in the lower micrograph. The *S*8 protein, which is not visible in the micrograph, is bound to the double-strand region. The capacity of the *S*8 protein to form a complex with either the 16*S* rRNA or the messenger RNA for its own operon enables the feedback loop to operate. The electron micrographs were made by Michael Beer of Johns Hopkins University, Theo Koller of the Swiss Federal Institute of Technology and workers in the author's laboratory at the University of Wisconsin at Madison.

lation of all the proteins in the unit. It turns out that when the messenger is translated, the ribosome can begin translation only at a single position within the regulatory unit: the ribosome entry site. Only when the first coding sequence after the entry site has been translated can the second sequence after the entry site be translated.

This process, which is termed sequential translation, or translational coupling, was first demonstrated for the operon made up of the *L*11 and *L*1 genes. In the *L*11-*L*1 messenger the ribosome entry site is near the beginning of the *L*11 coding sequence; protein *L*1 can be made only after protein *L*11 has been translated. The synthesis of both proteins is probably accomplished by the same ribosome, which continues on to the second coding sequence when it has reached the end of the first. It appears that when the ribosome moves along the messenger, each act of translation of protein *L*11 makes possible only one translation of protein *L*1.

The principle of sequential translation probably holds in other units of regulation in ribosomal-protein operons that are controlled by the translational feedback mechanism. Indeed, other work has shown that the principle holds even in some nonribosomal operons. In all these instances a unit of regulation consists of several genes and each coding sequence on the messenger RNA must be translated in turn before the next one can be read by the ribosome.

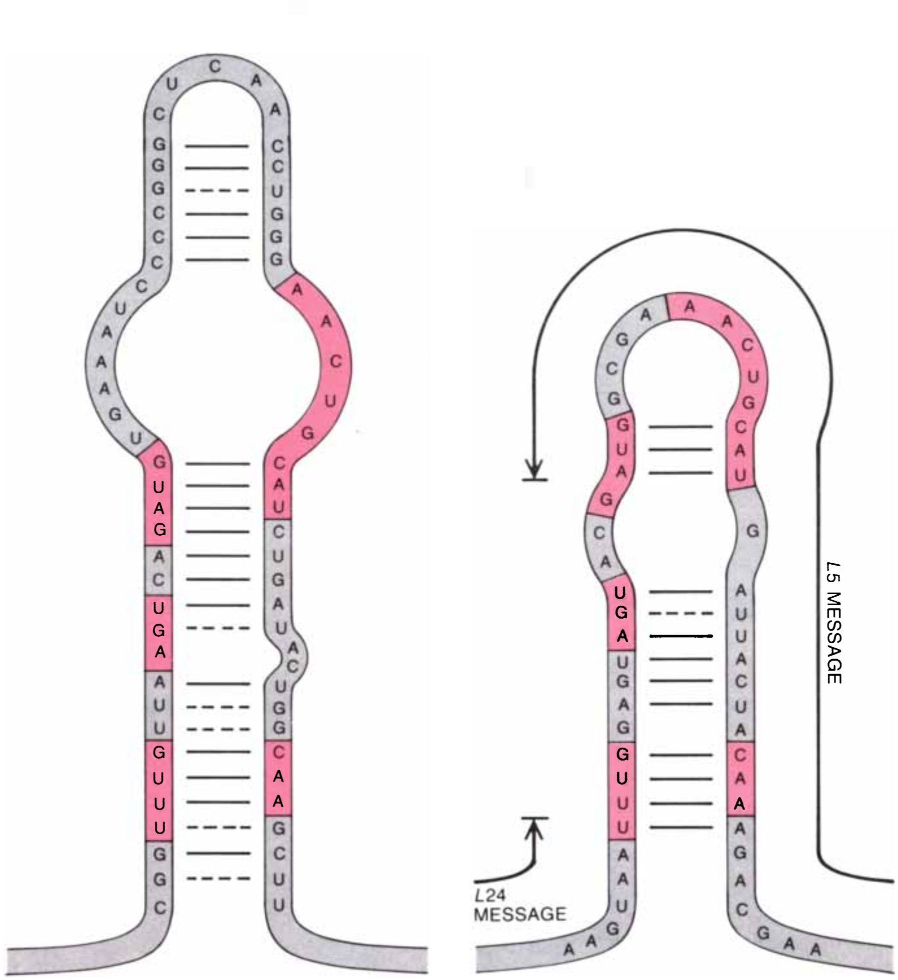
It can readily be seen that the combination of sequential translation and translational repression yields balanced and coordinated manufacture of the ribosomal proteins in relation to the ribosomal RNA's. Translational repression ensures that the ribosomal proteins are made in the same quantity as the rRNA's: when the supply of rRNA's for binding has been exhausted, the repressor ribosomal proteins immediately bind to their own messenger RNA's and synthesis of new ribosomal proteins stops. If the rate of synthesis of rRNA increases and more rRNA's become available for binding, the translation of the ribosomal-protein mRNA's begins again. When translation of the protein mRNA's is under way, sequential translation ensures that equal amounts of all the ribosomal proteins will be made.

When translational repression was discovered, the interaction between the repressor and the messenger immediately became a subject of great interest. The significance of the interaction stems from the fact that when ribosome assembly is in progress, ribosomal RNA's and messenger RNA's for ribosomal proteins are both present. An essential feature of the translational-feedback theory is the competition between the rRNA and the mRNA to bind the repressor ribosomal protein. For this competition to exist the repressor protein must be capable of binding to either RNA molecule. As a result of the study of ribosome assembly the binding of the protein to the rRNA was fairly well understood. Little was known, however, about the interaction of the repressor protein and the messenger.

Clues to the nature of the interaction began to accumulate; in the process new light was shed on old data. The translational repressors identified in the early work on ribosome synthesis all belonged to the category of "binding proteins." It was already known that binding proteins can form a specific complex with RNA in the test tube in the absence of other ribosomal proteins. Such interactions had been studied years before translational repression was discovered and some of the binding sites for the proteins on the RNA molecules had been characterized.

The information that had been gathered on the target sites for the binding proteins was exploited to help understand the competition between the ribosomal RNA and messenger RNA in binding to the repressor protein. The simplest way to achieve such competition would be for the same region of the repressor protein to interact with either of the two RNA's. If the same region of the protein served for binding, it seemed likely there would be structural similarities between the target site on the rRNA and the target site on the mRNA.

Similarities were indeed found in the instances where information was available on the makeup of the sites on the ribosomal RNA and the messenger RNA that interact with the same repressor ribosomal protein. The resemblances were of two kinds. Some similar nucleotide sequences were found at the binding site in both types of RNA. In addition there were similarities in the

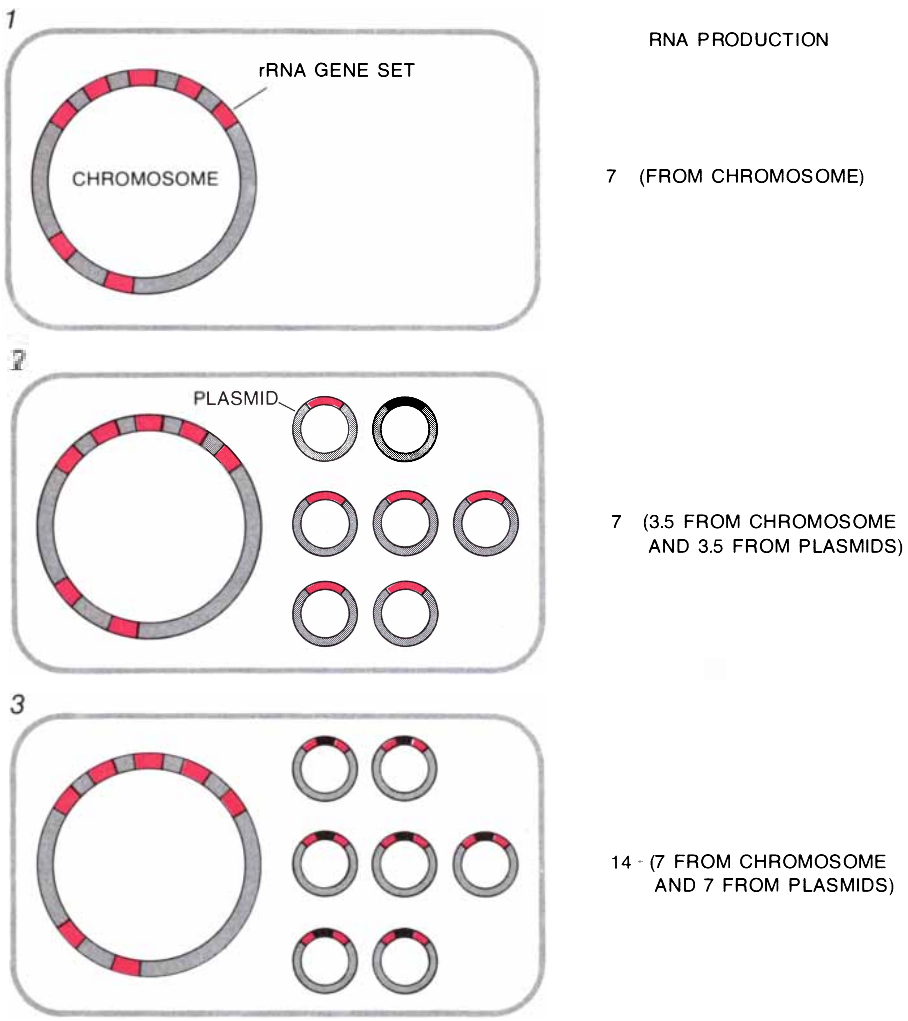


BINDING SITES for the translational repressor on ribosomal RNA and messenger RNA show significant similarities. In ribosome assembly the ribosomal protein *S* 8 forms a complex with the 16*S* rRNA. It can also form a complex with its own mRNA. When *S* 8 is bound to the mRNA, it prevents the translation of several proteins in its operon, beginning with *L*5. In both complexes the *S* 8 molecule interacts with a double-strand region of the RNA chain. The structure of the double-strand regions has been worked out. The one for the 16*S* rRNA is at the left; the one for the mRNA is at the right. The double-strand regions arise from pairing between complementary bases (*solid lines*) and also from weak, nonstandard pairing between *G*'s and *U*'s (*broken lines*). There are striking similarities between the regions in nucleotide sequences (*colored areas*). The resemblances enable the translational repressor to bind to either RNA.

secondary structure of the RNA chains.

The secondary structure of a section of an RNA molecule corresponds roughly to the folding of the molecule in space. The configuration that arises from the folding of the chain is influenced in part by the capacity of complementary stretches of the RNA strand to form bonds with each other. In areas where such bonding takes place the RNA molecule forms a double-strand structure; such structures can be part of a site where only a specific repressor protein can form a complex with the RNA.

The similarity between the target site on the messenger RNA and the target site on the ribosomal RNA makes the operation of the negative feedback loop possible. As long as rRNA's are being transcribed the repressor ribosomal protein recognizes the binding site on the rRNA and binds to it. When the rate of transcription of rRNA diminishes, the result is a small excess of the repressor. Some repressor molecules cannot find an rRNA to bind to; these proteins recognize the mRNA binding site and form a complex with it instead, thereby interrupting the translation of the proteins from the mRNA. Since the same process is going on simultaneously for all the repressor proteins, the translation of all the ribosomal proteins is halted.



SECOND FEEDBACK LOOP regulates the transcription of ribosomal RNA in *E. coli*. An experiment that served to establish the existence of the second loop is shown schematically in the panels of the illustration. Three groups of *E. coli* cells were grown on the same medium. All three made the same quantity of ribosomes. One group (*1*) had only the seven sets of rRNA genes (*color*) carried on the bacterial chromosome. The rRNA transcribed from the chromosome was incorporated into ribosomes. Free ribosomes, the ones not engaged in making protein, acted as repressors to prevent the transcription of unneeded additional rRNA's. Plasmids, or loops of free bacterial DNA, were inserted into a second group of cells (*2*); the plasmids carried extra sets of rRNA genes (*color*). The second group of bacteria made the same amount of rRNA as the first. Half of the rRNA came from the chromosome and half from the plasmids. RNA from both sources was incorporated into ribosomes that could function as repressors. A third group of cells also had plasmids inserted into them (*3*), but deletions had been made in the rRNA genes carried by these plasmids (*black*). As a result the rRNA derived from the plasmids could not be incorporated into ribosomes, and the RNA from the plasmids could not contribute to regulating the output of rRNA. Therefore the third group of bacterial cells made twice as much rRNA as the other two groups, including both functional and defective molecules.

It is apparent that when ribosomal RNA's and messenger RNA's are both available, the probability that the repressor ribosomal protein will bind to the rRNA must be considerably greater than the probability that it will bind to the mRNA. In other words, the binding of the repressor protein to the rRNA target site must be much more favorable in terms of energy than the binding of the repressor to the mRNA target site. If this were not so, the binding of the repressor ribosomal protein to the messenger would interrupt translation without regard to whether rRNA's were available to be assembled into new ribosomes. The system made up of the repressor and the two RNA's would therefore cease to function as a negative feedback loop.

This inference is supported by experimental results. In the test tube the repressor action of protein *L*1 on the synthesis of proteins *L*11 and *L*1 can be eliminated by the addition of a quantity of 23*S* ribosomal RNA to the reaction mixture. The 23*S* rRNA is known to form a complex with *L*1 when the ribosome is being assembled, and as we have seen *L*1 serves as the translational repressor of its own operon. Evidently the formation of the *L*1-23*S* complex is strongly favored in the competition with the complex between *L*1 and the mRNA for the *L*11-*L*1 operon.

Feedback mechanisms based on competition between two similar interactions could turn out to be a common means of regulating the expression of genes for the proteins that are included in cellular assemblies or are components of complexes of nucleic acids and proteins. At least one other instance of such a system is already known.

Since the repression of the translation of the ribosomal proteins depends on the supply of ribosomal RNA's, the question of how the quantity of rRNA is controlled becomes quite significant. It is a striking feature of ribosome synthesis that many different types of growth medium can result in the same rate of manufacture of ribosomes. For example, a medium with a poor carbon source and a good nitrogen source can yield the same quantity of ribosomes in *E. coli* as a medium with a good carbon source and a poor nitrogen source, provided the two mediums are on the whole equally favorable for growth. The same is true of many other substances needed for cell growth.

As we have seen, the operon model relies on the presence of a chemical signal (the inducer) to turn the genes on and off. The inducer is a chemical derivative of a substance in the environment of the bacterium. Since a medium with a particular chemical composition can yield the same rate of ribosome synthesis as a medium with a quite different

composition, it follows that if the manufacture of ribosomal RNA were regulated by a mechanism of the operon type, the two different mediums would have to give rise to the same chemical compound in the same quantity, which would serve as an internal signal in the regulatory process. This appears to be implausible, although not impossible.

It seems more likely that a second feedback mechanism determines the rate of transcription of the ribosomal RNA's. Suppose the *E. coli* cell generally has the capacity to make more ribosomes than it is making at any particular time. The bacterium's capacity is rarely tested, however, because free ribosomes (the ones that are not making proteins) usually limit the rate of transcription of rRNA and hence construction of new ribosomes.

To test this hypothesis Jinks-Robertson, Richard L. Gourse and I devised a gene-dosage experiment for ribosomal RNA. Plasmids carrying a complete set of rRNA genes were introduced into *E. coli.* The plasmids exist in multiple copies in the bacterial cell. In our experiment there were seven copies of the plasmid for each bacterial chromosome. Each transformed cell therefore had a total of 14 rRNA operons, its own seven and the seven added by the plasmids.

The doubling of the ribosomal-RNA operons caused no increase, or only a slight one, in the total rate of rRNA synthesis. The plasmid-borne rRNA operons were active: it was found that rRNA was being transcribed from them at approximately the same rate as it was being transcribed from the sets of genes on the bacterial chromosome. The 14 rRNA promoters were all functioning at about half the usual rate, yielding the same overall synthesis of rRNA as they would in bacteria without plasmids.

The outcome of the gene-dosage test is consistent with the feedback model for limiting ribosomal-RNA synthesis. According to this model, the rate of total rRNA synthesis will not significantly exceed the rate required to make the optimum number of ribosomes. If an excess of rRNA is made, it will accumulate in the form of free ribosomes, which will inhibit further synthesis of rRNA from both the plasmid-borne operons and the chromosomal operons.

This finding, however, is also consistent with another explanation. The alternative account entails the existence of a molecule needed for the transcription of ribosomal RNA. The quantity of the hypothetical factor would be determined by the composition of the bacterium's environment and therefore would not be affected by the number of sets of rRNA genes. This might be called a factor-limiting mechanism.

To evaluate the validity of the two hypotheses we deleted certain nucleotides in ribosomal-RNA genes that were inserted into plasmids. The plasmids were then put into *E. coli.* As a result of the deletions the rRNA transcribed from the operons on the plasmids was not functional: it could not be assembled into ribosomes. Therefore the bacteria into which the plasmids were put had extra rRNA genes and promoters but only the usual number of genes capable of making RNA that could become part of a ribosome.

If the factor-limiting account were correct, the total amount of ribosomal RNA made in the transformed cells would be equal to the amount made in cells with the usual complement of seven rRNA operons. The reason is that both groups of cells would be grown in the same medium; growing at the same rate, they would both get the same amount of the hypothetical factor that limits rRNA transcription.

If the feedback model is the right one, however, a quite different result would be observed. The defective ribosomal RNA derived from the operons on the plasmids could not be assembled into ribosomes and so would not serve to make proteins or to inhibit the synthesis of more rRNA. Only functional RNA incorporated into ribosomes that are then accumulated as free ribosomes can serve as a repressor. Therefore the cells with the defective rRNA genes would make more rRNA in total (including both defective and functional rRNA) than the unaltered bacteria do.

The gene-dosage test showed that there was a notable increase in ribosomal-RNA manufacture in the cells with the defective genes, which supports the feedback hypothesis. It is still not known, however, whether the control of rRNA synthesis is accomplished directly by the free ribosomes or by an intermediary molecule formed in response to the increase in free ribosomes.

Although some details must still be filled in, the question of how the synthesis of ribosomes is controlled in *E. coli* can now be answered in outline. Consider *E. coli* cells in a medium favorable for growth. Under such conditions the ribosome concentration is high but almost all the ribosomes are making proteins; there are few free ribosomes in the cell.

If the bacteria are transferred to a medium less favorable for growth, the rate of protein synthesis declines, and as a result there is an increase in the quantity of free ribosomes. The free ribosomes directly or indirectly inhibit ribosomal-RNA synthesis and also the synthesis of some other RNA's, including transfer RNA's, which are not relevant to the current argument. The absence of newly transcribed rRNA leads to the interruption of ribosomal-protein synthesis by the translational feedback cycle. The accumulation of ribosomes is halted and the ribosome concentration gradually becomes commensurate with a lower rate of growth.

If *E. coli* growing in a poor medium are put into one that favors growth, however, the rate of protein synthesis increases rapidly. The immediate increase is small because there is generally only a small quantity of free ribosomes available to be put to work. Even a small increase in protein synthesis, however, entails a large reduction in the concentration of free ribosomes, and that reduction results in a burst of ribosomal-RNA synthesis and therefore in a quickening of the pace of ribosome accumulation.

This pattern is consistent with the observation that when bacteria enter a richer medium, the rate of ribosome synthesis increases faster than the rate of protein synthesis. The rate at which proteins are made is initially increased only by the utilization of the few free ribosomes; larger increases must await the assembly of new ribosomes. Eventually a steady state is reached in which ribosome manufacture is in balance with the volume of protein synthesis determined by the new, favorable conditions for cell growth and division.

Hence the two intertwined feedback systems can explain how *E. coli* cells adjust the assembly of ribosomes to achieve a particular growth rate. The presence of a considerable quantity of free ribosomes in the cell when growth is very low suggests that in general cell growth is not limited by the capacity to make ribosomes. On the contrary, the bacterium appears to adjust the rate at which ribosomes are made to achieve the highest growth rate possible in the particular environment. The mechanisms found in *E. coli* for adjusting the ribosome synthesis rate must have conferred considerable adaptive advantage on the ancestors of the current *E. coli* lines when such mechanisms evolved in competitive natural surroundings.

Having understood in a general way how the manufacture of ribosomes is carried out in a bacterium, a logical next step is to attempt to understand how such regulation is achieved in eukaryotic cells. Because of the division of the cell into the nucleus and the cytoplasm, the control of ribosomal synthesis in eukaryotic cells must be somewhat more complicated than it is in the bacterium. In spite of such complexities the eukaryotic cell, like its bacterial relative, appears to adjust the rate of ribosome synthesis to match the environmental conditions and yield a specific rate of cell growth. Although the mechanism of the regulation remains a puzzle, having penetrated the mystery of how *E. coli* manages the feat makes it possible that the control of ribosome synthesis in eukaryotic cells will be understood in the not too distant future.

SCIENCE/SCOPE

An advanced infrared seeker now being developed would improve the operating range and accuracy of future air-to-ground missiles and guided bombs. Hughes Aircraft Company is producing a scanning focal plane array (FPA) seeker to demonstrate advanced infrared imagery. The sensor, the size of a collar button, consists of tiny infrared detectors on one side and a like number of signal-processing elements on the other. Because the sensor would be more sensitive than existing devices, it can stay locked on small targets more easily, distinguish between targets and background clutter more easily, and detect targets from farther away. The seeker also promises benefits in weight and cost. Hughes also will conduct a study to determine whether the seeker would be feasible for a variety of weapons planned by the U.S. Air Force and Army for between 1990 and 2000.

Productivity will be improved by a new computer system that will give management access to vital manufacturing information within minutes instead of hours or days. Hughes is developing a system that will collect data directly from numerical-control machines, machine operators, material movers, assemblers, inspection stations and operators, and product test stations. Information will be entered into the system via bar code wands and scanners, keyboards, and voice input devices. In the interest of speed and accuracy, work station terminals will be programmed to determine if the data seems reasonable. If not, it will ask for corrections before entering the data into the system.

A weapon-locating radar can pinpoint the sources of enemy artillery, rockets, or mortars -- often before the first projectile hits. The AN/TPQ-36 Firefinder radar sweeps a series of pencil-shaped radar beams, adjustable according to the terrain, along a 90-degree sector of the horizon several times a second. When an object breaks through this curtain, the system instantly transmits a verification beam. If this beam detects a target, the system's computer unleashes a rapid succession of tracking beams. While tracking this target, the radar continues scanning, locating other targets and developing tracks on them as well. Hughes builds the radar for the U.S. Army, Marine Corps, and selected allies.

An infrared sensor made of standard components turns night into day for tanks and other combat vehicles. The compact device, called Hughes Infrared Equipment (HIRE), was designed to be low in cost yet high performing. It can be adapted to periscopes to let gunners of such tanks as the M48 see through darkness, haze, or battlefield smoke. HIRE can be mounted in laser tank fire control systems, light armored vehicles, or used as a target acquisition/fire control sight for anti-aircraft, ship, and helicopter applications. The design uses U.S. Army common modules, the standard building blocks for thermal imaging systems.

Hughes Research Laboratories needs scientists for a spectrum of long-term sophisticated programs. Major areas of investigation include: masked ion beam lithography, liquid-crystal materials and applications, submicron microelectronics, ion propulsion, and new electronic materials. For immediate consideration, please send your resume to Professional Staffing, Hughes Research Laboratories, Dept. SE, 3011 Malibu Canyon Road, Malibu, CA 90265. Equal opportunity employer.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

For additional information please write to:
P.O. Box 11205, Marina del Rey, CA 90295

The Packing of Spheres

What is the densest way to arrange identical spheres in space? There has been much progress on the problem, particularly in 24 dimensions, and the results can be applied to digital signaling

by N. J. A. Sloane

A manufacturer of ball bearings is asked to deliver as many balls, all the same size, to a foreign port as he can ship on a day's notice. Although the ball bearings are ready, only one ship is available, and the shipper explains to the manufacturer that the draft of the ship would be too deep for the local channel if the balls were to fill more than three-fourths of the volume of the hold. The manufacturer is unperturbed. "Your ship is safe," he replies. "Let the hold be filled to the hatches." Should the shipper believe his client?

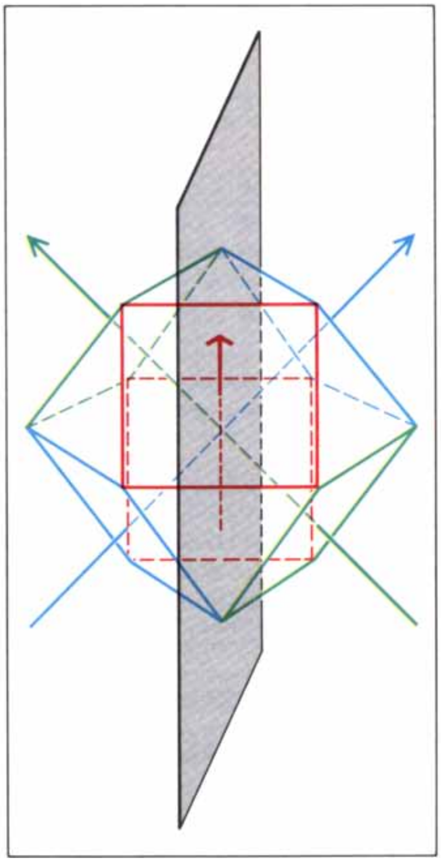
The solution to this problem depends on finding out how densely a large number of identical spheres can be packed together in space. If, instead of ball bearings, the ship's hold were to be packed with cubes all the same size, the answer would be easy. Since the cubes would fit together with no wasted space in between, the hold could be essentially filled with cubes (ignoring the small spaces that might be left around the walls and ceiling) and the manufacturer's assurances would clearly be wrong. Balls, however, cannot be packed without wasting space. If in spite of arranging the ball bearings as densely as possible the wasted space still exceeds a fourth of the volume of the hold, the shipper can safely fill the hold with balls and proceed out of the channel.

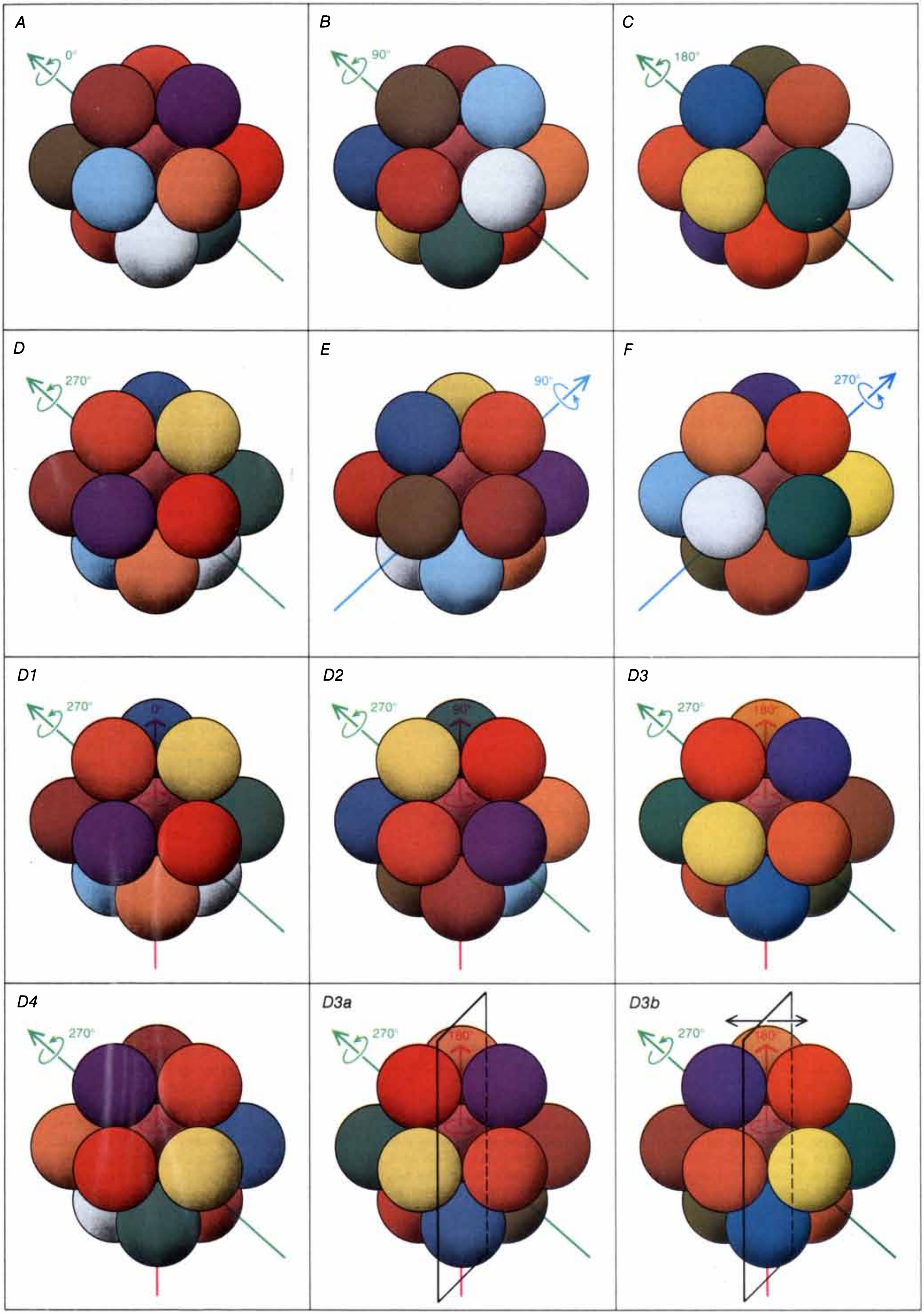
A few minutes spent experimenting with oranges or billiard balls is enough to mislead many people into thinking the problem is trivial. Arrange three spheres on a flat surface so that their centers form an equilateral triangle; continue adding spheres on the surface so that each new sphere touches at least two spheres already in place. Build a second layer of spheres by placing each new sphere in the "deep hole," or depression, left at the center of any triangular group of spheres in the first layer. The finished second layer is identical with the first layer, although it is shifted in the horizontal plane. If more layers are added in the same way, the packing of spheres that results is called the face-centered-cubic packing, which is familiar to chemists and crystallographers; it fills just over 74 percent of the volume of the space. As far as anyone knows, it is the densest packing that can be achieved.

I am sorry to report, however, that this density has never been mathematically proved to be maximal. The least upper bound on the density obtained so far was found in 1958 by C. A. Rogers of the University of Birmingham, who proved that no packing of spheres can have a density greater than about .7796. The result is not particularly helpful to anyone looking for a more efficient way to pack ball bearings. Rogers' proof offers no construction of any sphere packing that comes close to his bound, and in the paper announcing the proof he remarked that "many mathematicians believe, and all physicists know," the correct answer is about 74 percent. In the quarter century since Rogers made the remark his assessment stands unchanged; the sphere-packing problem, so simple to state and so difficult to solve, remains one of the basic unresolved problems in mathematics.

Densely arranged configurations of hard spheres have been studied for many years in part because of their broad implications for an understanding of the behavior of solids and liquids. For example, the molecular properties of

FACE-CENTERED-CUBIC packing of spheres, often seen in fruit stands or in piles of cannonballs at war memorials, is thought to be the densest packing of spheres in three-dimensional space. In spite of centuries of effort, however, a proof of its maximal density has never been given. Each sphere in the packing "kisses," or touches, 12 other spheres; a proof that this number is maximal was not given until 1874. If the center of one sphere is fixed, the set of all possible rotations and reflections that permute the 12 surrounding spheres is called the symmetry group of the packing. The symmetry group of the face-centered-cubic packing has 48 elements; they can best be understood if the center of each sphere is thought of as the vertex of the polyhedron at the right, which is called a cuboctahedron. Any one of the six square faces of the cuboctahedron can become the front face by an appropriate rotation of the figure about the green or the blue axis (*A–F*). Each set of four spheres that form a square face (say the fourth face) can then assume one of four configurations if the entire figure is rotated about the red axis (*D1–D4*). Finally, each configuration (say the third one) can be reflected about a vertical plane to give a new configuration (*D3a, D3b*). The total number of elements in the symmetry group is therefore 6 × 4 × 2, or 48. An analogous symmetry group, which describes the rotations and reflections of a dense packing of spheres in 24-dimensional space discovered by John Leech at the University of Glasgow, has been important in the mathematical theory of finite groups.



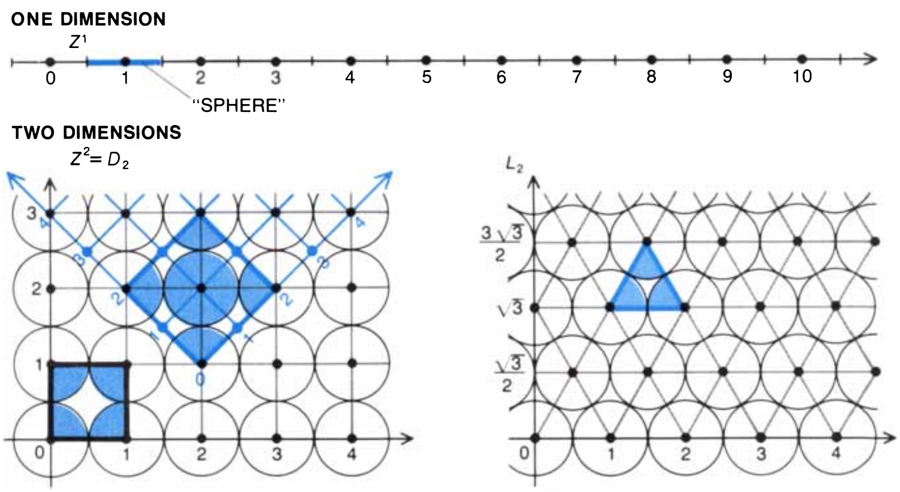

A
0°
B
90°
C
180°
D
270°
E
90°
F
270°
D1
270°
0°
D2
270°
90°
D3
270°
180°
D4
270°
D3a
270°
180°
D3b
270°
180°

many crystalline materials can be described, at least to a first approximation, as the effects of various forces on a huge assembly of closely packed spheres. Equally important is the application of the sphere-packing model to the properties of powders and porous materials.

Although experimental studies of sphere packings are important for an understanding of certain physical systems, there are also compelling reasons to carry out mathematical studies of ideally dense packings. For example, the fact that no sphere packing has been proved to be maximally dense suggests the mathematical understanding of ordinary, three-dimensional Euclidean space is far from complete. Moreover, for the mathematician the concept of a sphere and the problem of packing spheres can be generalized to include mathematical objects called n-dimensional spheres, whose algebraic form closely resembles the algebraic description of the sphere in ordinary space. The study of sphere packings in n dimensions has been recognized for some time as being mathematically equivalent to the design of a finite set of digitally encoded messages that do not waste power or cause confusion in transmission. Furthermore, in recent years the search for optimal sphere packings in spaces of 24 or more dimensions has led to major discoveries in the branch of mathematics called group theory.

There are two other problems, closely related to the sphere-packing problem, that are important in geometry. One is known as the "kissing number" problem: How many spheres can be arranged around a central sphere in such a way that all the surrounding spheres just "kiss," or touch, the central sphere? The kissing-number problem in three dimensions was the subject of a famous dispute in 1694 between Isaac Newton and the Scottish astronomer David Gregory. Newton maintained the kissing number is 12, and for the face-centered-cubic packing of spheres described above the kissing number is indeed 12 [*see illustration on preceding page*]. Gregory probably argued that an additional sphere could be squeezed in, although he was not able to prove it.

According to H. S. M. Coxeter of the University of Toronto, Gregory may have imagined that the 12 outer spheres could be rolled around the central sphere in such a way that all the gaps could be concentrated in one direction, thereby leaving room for 13 surrounding spheres. Actually it can readily be shown that the solid angle, measured from the center of the central sphere, that is subtended by one of the surrounding spheres is less than 1/13 of the total solid angle. The total volume of the space around the central sphere is therefore sufficient to accommodate the volume of 13 surrounding spheres. Inevitably, however, part of the total solid angle at the central sphere is subtended by the gaps that arise when the surrounding spheres are packed together. It was not until 1874 that the question of the kissing number was settled; R. Hoppe showed Newton was right.



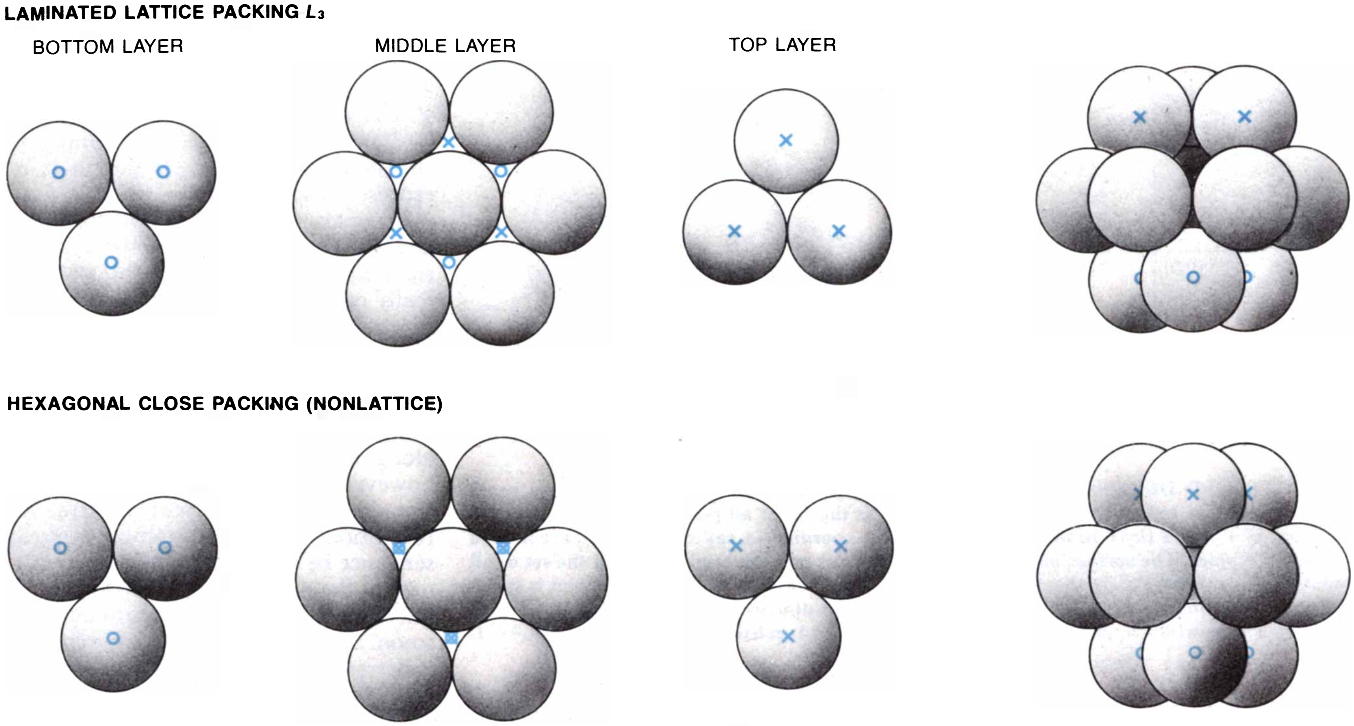
"SPHERE" PACKINGS can be carried out in one and two dimensions as well as in three. In one dimension the spheres are line segments of unit length, centered at integer points. The spheres cover 100 percent of the line and each sphere kisses two others. This packing is obviously as dense as possible: it is called Z^1. In the plane the spheres are circles and there are three packings of interest. In the Z^2 packing the spheres are centered at every point in the plane whose coordinates are integers; in the D_2 packing the spheres are centered at alternating points of the Z^2 packing, in checkerboard fashion. If the coordinate axes of the Z^2 packing are rescaled and rotated 45 degrees, the result is the D_2 packing. Hence the two packings are equivalent. Their density is the same as the part of the area of one square that is covered by circles or parts of circles (*shading in color*); it is $\pi/4$, or about .7854 (*left*). The densest packing of circles in two dimensions is the hexagonal lattice packing L_2. Parts of circles cover $\pi\sqrt{3}/6$ of each equilateral triangle; the density of the packing is therefore equal to about .9069 (*right*).

The second important problem related to sphere packing is called the covering problem: What is the least dense arrangement of identical spheres with the property that any point is inside or on the boundary of at least one sphere? Unlike the first packing problem, in which the spheres cannot overlap, the solution to the covering problem inevitably calls for overlapping spheres. One way to cover an entire volume with spheres is to inflate all the spheres in a sphere packing until they encompass all the deep holes in the original packing. In general, however, inflating the spheres in the densest nonoverlapping configuration does not lead to the best solution of the covering problem. In three dimensions, for example, it is believed the best covering is given by spheres arranged at the vertexes of what is known as a body-centered-cubic lattice. If nonoverlapping spheres are centered at these points, however, the packing that results is not as dense as other known sphere packings, such as the face-centered-cubic packing. Moreover, the conjecture that body-centered-cubic packings solve the covering problem has not been proved.

What makes it so hard to solve the sphere-packing problem in three dimensions? Consider the face-centered-cubic packing again. Since the problem states only that the density of the packing must be maximized, the number of spheres in the packing must be regarded as unlimited. If, say, a million spheres are removed, the packing density will remain effectively unchanged; in other words, the change in density can be made as small as one wants by increasing the number of spheres in the packing as a whole. For each sphere removed from the face-centered-cubic packing, however, a distinct packing is created. Thus a proof that the face-centered-cubic packing is optimal (as many people believe it is) is at the same time a proof that an infinite number of distinct packings are optimal as well.

There is another way to appreciate the complexities of the sphere-packing problem in ordinary space. In three dimensions at most four spheres can touch one another at a time. (This property must be clearly distinguished from the configuration required by the kissing-number problem. In that problem only one central sphere must be touched by surrounding spheres; the surrounding spheres need not touch one another.) When four spheres are mutually in contact, their centers form the vertexes of a regular tetrahedron, or triangular pyramid. Since the four spheres cannot move closer together, the densest configuration of four spheres possible in space is the tetrahedral configuration.

Suppose new spheres are added to a tetrahedral configuration one at a time so as to form a new tetrahedral configuration whenever possible. If this proce-



CONSTRUCTION of the laminated lattice packing L_3, which is equivalent to the face-centered-cubic packing D_3, is done by fitting together layers of spheres whose centers are arranged according to the hexagonal lattice packing L_2. If the spheres are packed so that the spheres in the third hexagonal layer are directly above the spheres in the first layer, the packing is called the hexagonal close packing. The hexagonal close packing is just as dense as L_3, but the centers of the spheres in the packing do not satisfy the definition of a lattice.

dure could be continued indefinitely, the resulting arrangement of spheres would have the greatest packing density possible, since every tetrahedral configuration would be packed as densely as possible. Rogers' upper bound for the densest sphere packing possible is based on precisely this argument; a straightforward computation in spherical trigonometry shows that about .7796 of the volume of a tetrahedron is filled by the four nonoverlapping spheres that can be centered at its vertexes. (The exact result is $\sqrt{2}$ [3 arccos (1/3) $- \pi$].) Unfortunately, however, tetrahedrons do not fit perfectly together to fill space. The strategy of packing spheres into tetrahedral arrangements whenever possible, which might be called a greedy algorithm, eventually forces a bad move. At a certain stage in the procedure the growing configuration of spheres presents a surface that cannot accrete more spheres without wasting interior space. Hence although the greedy algorithm generates an optimal sphere packing over the short range of, say, a few spherical diameters, it turns out that the algorithm gives rise to a packing that is less dense on a global scale than the face-centered-cubic packing.

In order to make further progress on the three sphere-packing problems mathematicians have found it convenient to supplement geometric intuition with an analytic representation of the spheres in terms of their rectangular coordinates. It is well known that any point in the plane can be specified by two coordinates, a horizontal coordinate x and a vertical coordinate y; the point is generally written as the ordered pair (x,y). For example, the point (3,4) refers to the point in the plane three units to the right of the origin along the x axis and four units above the origin along the y axis.

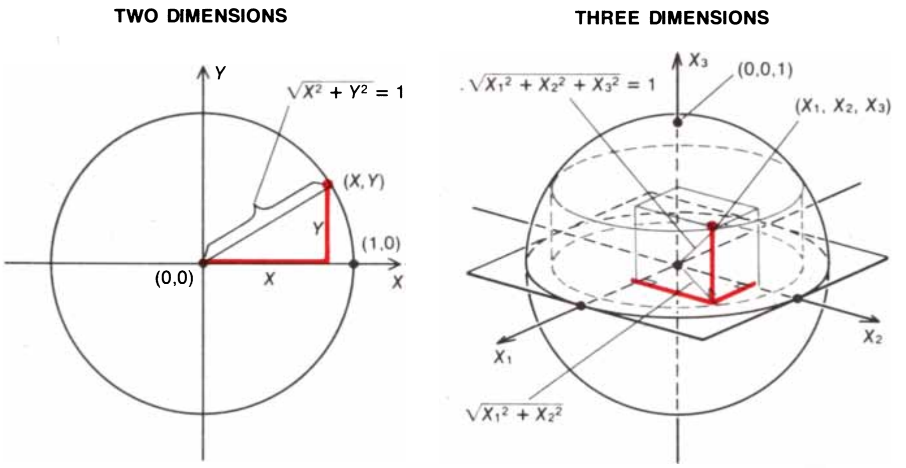
The distance between, say, the point (3,4) in the plane and any other point (x,y) can be calculated from the Pythagorean relation among the sides of a right triangle: The square of the distance between the two points is equal to the square of the distance between them along the x axis, $(x-3)^2$, plus the square of the distance between them along the y axis, $(y-4)^2$. Since a circle is by definition the set of all points in the plane equidistant from a central point (a,b), any point (x,y) on the circumference must satisfy the equation $(x-a)^2+(y-b)^2=R^2$, where R is the radius of the circle. If the radius of the circle is equal to 1 and the center is at the origin (0,0), the equation is much simpler: all points (x,y) on the circumference must satisfy the equation $x^2+y^2=1$.

Similarly, any point in three-dimensional space is specified by three coordinates, x, y and z; more suggestively, the point can be written (x_1,x_2,x_3). The surface of a sphere of radius 1 centered at the origin is made up of all the points (x_1,x_2,x_3) such that $x_1^2+x_2^2+x_3^2=1$. the equation arises, much as it does in two dimensions, from the geometric definition of a sphere and from two applications of the Pythagorean relation.

In more than three dimensions geometric intuition is of little value and one must begin to think exclusively in terms of coordinates. For example, a "point" in four-dimensional space is a mathematical object that requires four distinct numbers in order to be specified unambiguously; such a point is written (x_1,x_2,x_3,x_4). If a person's height, weight, age and income are sufficient to unambiguously pick out the name of that person from a list, the four quantities can be regarded as specifying a point in a four-dimensional space.

A four-dimensional sphere is defined by analogy with the definitions for the circle and the sphere in two and three dimensions. All the points (x_1,x_2,x_3,x_4) on the "surface" of the sphere are found at some distance R from a central point (a_1,a_2,a_3,a_4). The sum of the squares of the distance along each independent coordinate axis between any point (x_1,x_2,x_3,x_4) on the sphere and the central point (a_1,a_2,a_3,a_4) must always be equal to R^2.

There has been a great deal of nonsense written in science fiction and elsewhere about the mysteries of the fourth dimension. In mathematical discussions one must not assume, as the physicist does, that the fourth dimension represents time. Furthermore, one must avoid the temptation to reify the somewhat metaphorical terms "surface,"



POINT IN TWO DIMENSIONS is specified by assigning a value to two coordinates, say x and y. A circle of radius 1 around the origin (0,0) is the set of all points (x,y) that satisfy the equation $x^2+y^2=1$ (*left*). In three dimensions three coordinates, say x_1, x_2 and x_3, are needed to specify a point. The surface of a sphere of radius 1 around the origin (0,0,0) is the set of all points (x_1,x_2,x_3) that satisfy the equation $x_1^2+x_2^2+x_3^2=1$. In n dimensions a point is specified by n coordinates, $x_1,x_2,\ldots,x_n$. The surface of an n-dimensional sphere of radius 1 centered on the origin $(0,0,\ldots,0)$ is the set of points $(x_1,x_2,\ldots,x_n)$ such that $x_1^2+x_2^2+\ldots+x_n^2=1$.

"point," "sphere" and so on, which are applied to mathematical objects whose algebraic properties are otherwise perfectly straightforward. The terms are justified because the objects they refer to are constructed by analogy with the algebraic properties of ordinary circles and spheres; it must not be supposed, however, that the objects represent real geometric objects in some universe wider than our own. To repeat, in mathematics four-dimensional space consists of points with four coordinates instead of three (and the same holds for any number of dimensions).

Once the concept of a sphere is understood as a relation among coordinates, there are several problems, seemingly unrelated to sphere packing, that can be reduced to one of the three sphere-packing problems. The principal applications of high-dimensional sphere packing are to problems in digital communications, particularly in the construction of signals for use on a noisy channel and in the design of analogue-to-digital converters. It must be admitted, however, that most discussions of these applications have been theoretical; only quite recently have sphere packings been used to design practical systems. As communications systems become more sophisticated further applications can be expected.

In digital communications one of the major aims of the designer is to construct a list of distinct coded symbols, or code words, that can be transmitted with maximum reliability and minimum power. Each code word might be represented as, say, an eight-digit symbol, each digit of which can take on one of five distinct values: 0, 1/2, 1, −1/2 or −1. At first it would seem that the system would provide for 5^8, or 390,625, different code words, but the difference between many pairs of these code words is so small that such a system would be highly subject to random errors in transmission or to electrical interference. For example, the difference between the code word (1,1,1,1,1,1,1,1) and the code word (1,1,1,1,1,1,1,1/2) is far too small. If both code words were in use, it is likely they would often be confused. Another way to put the same point is that if the difference between two code words were as small as the difference between (1,1,1,1,1,1,1,1) and (1,1,1,1,1,1,1,1/2), a very large amount of power would be needed to guarantee that these two code words could be distinguished in the presence of background noise.

There is a general mathematical relation between the distinguishability of code words and the power needed to transmit them reliably. The relation was first formulated in 1948 by Claude E. Shannon, then of the Bell Telephone Laboratories, in his paper *A Mathematical Theory of Communication*. As my distinguished colleague David Slepian has said, "probably no single work in this century has more profoundly altered man's understanding of communication" than Shannon's paper. What Shannon showed is that given some fixed, finite amount of power there always exists a system of code words that can be transmitted essentially without error. The only qualification is that the rate at which the code words are transmitted cannot exceed a critical threshold called the capacity of the transmission channel. Unfortunately Shannon's theorem is nonconstructive; it proves that such systems for encoding a signal exist, but it gives no hint of how they might be designed. Although many signaling systems have been constructed, schemes that perform as well as Shannon's theorem promises have still not been found.

One way to design a signaling system that comes close to meeting the standards of Shannon's theorem is to represent each signal as a point in n-dimensional space. For example, consider any sequence of eight numbers in the signaling system described above. Physically each of the numbers corresponds to a voltage level on a transmission line, and so each code word can be plotted on a two-dimensional graph as a series of eight distinct pulses whose height is specified for each of eight intervals along the time axis. Mathematically, however, a single point in eight-dimensional space can represent the same information: let the first number in each sequence be the value of the first coordinate of the point, the second number be the value of the second coordinate and so on. Since a point in eight dimensions is determined by fixing the values of all eight coordinates, every code word in the system can be represented as a distinct point in eight-dimensional space.

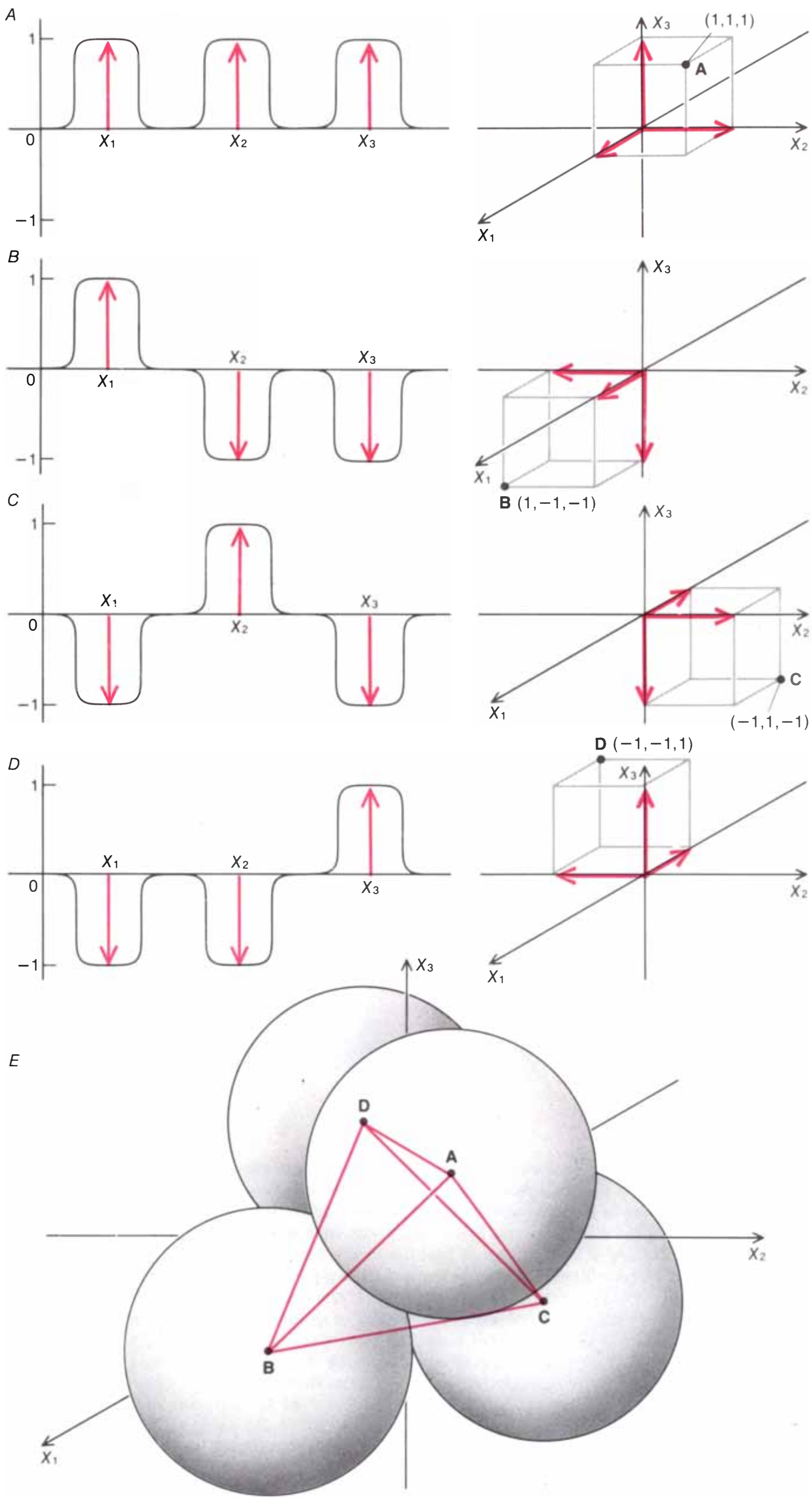
When the code words are represented as points, two important features of a system of code words can be given a geometric interpretation. First, remember that the code words must be reliably distinguishable from one another. In eight-dimensional space this suggests that the points representing the possible code words must be separated by a certain minimum "distance." How well does the Pythagorean distance between two points in eight dimensions measure the distinguishability between two code words? In order to determine the Pythagorean distance the difference between the two values of each coordinate of the two points must be squared. Small differences between coordinate values (that is, differences less than 1) are thereby reduced, whereas large differences (that is, differences greater than 1) are magnified. Because a small difference in the voltage levels between two signals is much more likely to lead to confusion between code words than a large difference, the Pythagorean distance is a reasonable measure of distinguishability.

For example, consider the two code words (1,1,1,1,1,1,1,1) and (1/2,1/2,1/2, 1/2,1/2,1/2,1/2,1/2). The square of the distance between the two points is the sum of eight squares, each of the form $(1-1/2)^2$, and so the distance between them is $\sqrt{2}$. According to the Pythagorean measure of distance, the two code words (1,1,1,1,1,1,1,1) and (0,0,1,1,1,1, 1,1), which differ only in the first two coordinates, are also $\sqrt{2}$ units apart, and so they are just as easy to distinguish as (1,1,1,1,1,1,1,1) and (1/2,1/2, 1/2,1/2,1/2,1/2,1/2,1/2).

A second important feature of any system of code words is that the power needed to transmit them must be minimized. It follows from the elementary definitions of power and voltage that power is directly proportional to the square of voltage; for a simple circuit the power is the square of the voltage divided by the resistance in the circuit. The total power necessary to transmit an eight-digit code word is consequently the sum of the squares of all eight digits that make up the code word. This sum is the square of the distance between the point representing the code word in eight-dimensional space and the origin (0,0,0,0,0,0,0,0).

The design of a signaling system that is reliable and makes efficient use of power can therefore be reduced to the geometric problem of placing points inside a region of space while constraining them not to be too close together. If the points must be at least a distance of, say, $\sqrt{2}$ apart, the problem is equivalent to the problem of finding the densest packing of spheres whose radius is half that distance, $\sqrt{2}/2$. A closely related problem is to find a set of code words that all have the same energy. This problem is equivalent to the problem of placing as many points as possible on the surface of an *n*-dimensional sphere while constraining them not to be too close together. That problem is in turn a generalized version of the kissing-number problem.

It turns out that in a space of eight dimensions there is an extremely dense packing called the E_8 packing; it was discovered in the last third of the 19th century by the Russian mathematicians Alexander N. Korkin and E. I. Zolotareff and by the English lawyer and amateur mathematician Thorold Gosset. The centers of the spheres in the E_8 packing are all the points whose coordinates are equal to whole numbers or to whole numbers plus a half. The sum of the coordinates for each point must be an even number. There are 240 such points whose distance is $\sqrt{2}$ from the origin: 112 points of the from $(\pm 1, \pm 1, 0,0,0,0,0,0)$, where the two 1's and any combination of signs can appear in any positions, and 128 points of the form $(\pm 1/2, \pm 1/2, \pm 1/2, \pm 1/2,$



DESIGN OF A CODE for the efficient transmission of information is closely related to the sphere-packing problem. The code is to be a finite set of signals called code words that are easily distinguished from one another and do not waste electric power. If each code word is a sequence of, say, three discrete voltage levels, each sequence can be represented as a point in three-dimensional space: the first coordinate of the point is the numerical value of the first voltage level, the second coordinate is the value of the second voltage level and so on (*A–D*). The transmission power required for each voltage pulse is proportional to the square of the voltage, and so the total power needed to transmit one code word is the sum of the squares of the three discrete voltages associated with the code word. The sum is equal to the square of the distance from the origin to the point in three-dimensional space that represents the code word. Thus the problem of minimizing transmission power is equivalent to the problem of placing all the points that represent code words as close to the origin as possible. On the other hand, the need to distinguish code words from one another can be treated as the requirement that the points in space representing the code words not be closer together than some minimum distance *d*. Meeting the two requirements simultaneously is geometrically equivalent to packing hard, nonoverlapping spheres of radius *d*/2 around the origin as densely as possible (*E*).

$\pm 1/2, \pm 1/2, \pm 1/2, \pm 1/2$), where the number of minus signs is even.

The E_8 packing could become the basis of a practical and efficient signaling scheme. If the scheme were to include exactly 240 code words, the 240 points in the E_8 packing that are equidistant from the origin could be chosen as the code words. In most practical systems, however, the number of distinct code words needed is some integral power of 2. For example, in medium-distance telephony the digital transmission system known as pulse-code modulation is now in wide service. The voltage of the voice signal is measured every 1/8,000 second, and the measured samples are quantized, or replaced, by one of 2^8, or 256, levels. Each quantized voltage level is then expressed as an eight-digit binary number; such binary numbers make up what is called the source code.

It has been known since the work of Harry Nyquist of Bell Laboratories in the 1920's that the voice signal can be reconstructed from the sample values alone. The voice signal can be thought of as a quantity, such as air pressure or voltage, that varies continuously with time. Early in the 19th century Jean Baptiste Joseph Fourier proved that the graph of any such quantity can be approximated to any desired degree of accuracy by superposing sine and cosine curves of appropriate amplitude and frequency. The sine and cosine curves (which, for an arbitrary curve, may be infinite in number) are called the Fourier components of the curve on the graph.

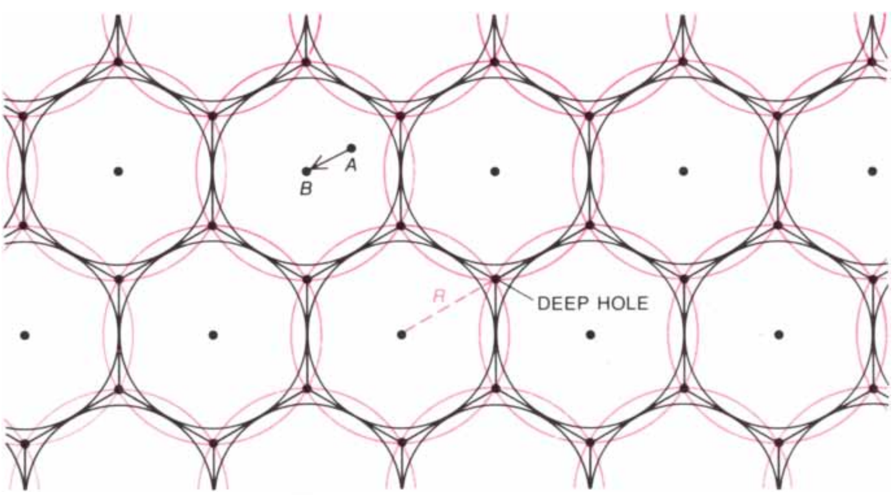
Suppose the graph of a quantity can be exactly generated by superposing a finite number of Fourier components whose frequencies are no higher than some frequency W cycles per second. Then what Nyquist showed is that the graph can also be perfectly reconstructed solely from the values it takes every $1/2W$ second. For example, a voice signal that has no frequency components higher than 4,000 cycles per second can be perfectly reconstructed from samples of the signal taken every 1/8,000 second. It is therefore enough to sample the voice signal and transmit only the sample values, which are represented by the code words in the source code, instead of transmitting the entire voice signal. If the sample values were transmitted without being rounded off, Nyquist's theorem shows that an entire second of speech signal, which corresponds to 8,000 sample values, can be represented as a single point in a space of 8,000 dimensions. This shows the power of mathematics.

For efficient transmission the source-code numbers that represent the sample values must be further encoded by another code, called a channel code; it is the channel code that pertains to the sphere-packing problem. An excellent way to derive a channel code from the E_8 sphere packing is to encode each pair of successive eight-digit binary numbers in the source code. Each 16-digit binary number that results is then assigned to the center point of one of 2^{16}, or 65,536, spheres in the E_8 packing. A good channel code can then be constructed by choosing the 65,536 center points that lie closest to the origin. At the receiving end of the telephone line the code words that correspond to the coordinates of each center point are converted back into the binary numbers of the source code, and the voice signal is reconstructed from the binary numbers.

There is a second major application of sphere packing in digital communications that I shall describe briefly. Remember that in deriving the binary numbers in the source code from the voice signal of the telephone it is necessary to quantize the precise intensity of the signal to one of 256 levels. The real world is full of awkward numbers such as .7913 . . . , but the world of computers and digital systems must ultimately deal only with round numbers such as 0 and 1. Any device that rounds off a continuously variable quantity to some set of discrete values is called an analogue-to-digital converter, or quantizer.

Quantization can be carried out in two or more dimensions as well as along a single coordinate axis. Imagine that the plane is divided into regions, not necessarily congruent, and imagine that within each region one point has been marked. Any such array of points and regions can function as a two-dimensional quantizer; the input to the quantizer is a pair of real numbers that specify some arbitrary point, and the output is the preselected quantizing point that lies in the same region of the plane as the arbitrary point. Thus any point in the plane is rounded off to one of the quantizing points. The process compresses input data; a single indexing number for the quantizing point can be transmitted in place of the precise coordinate values of the data point.

Quantizing introduces errors, and so one tries to choose the quantizing points in such a way as to minimize the average error. For example, if the input to the quantizer is uniformly distributed, or in other words if every input value is equally probable, it is straightforward to calculate the average error for a number of quantizing schemes. If a single coordinate axis is divided into equal seqments of unit length and the quantizing points are centered on each segment, the average error is 1/12, or about .0833. The same data can be quantized in two dimensions: the data points are paired, and each pair is considered a point in the plane. If the plane is then divided into squares and the quantizing points are centered on each square, the average error is still 1/12. On the other hand, if the plane is divided into regular



QUANTIZING DATA from a continuously variable source is closely related to the problem of covering space with the least dense arrangement of overlapping spheres. For example, in order to quantize data in two dimensions the input data are paired and each pair is treated as the coordinates of a point in the plane. Each data point, such as the point *A*, is then rounded off to a quantizing point, such as the point *B*, that lies in the same preselected region of the plane as the data point. The problem is to choose the quantizing points and the partition of the plane in a way that minimizes the average quantizing error. If the data points are uniformly distributed and the quantizing points are chosen as the centers of squares, the average error is 1/12. A better way of quantizing can be derived from the best covering of the plane by circles. That covering is generated from the hexagonal packing of circles (*black circles*) by increasing the radius of each circle by just enough to include every point in the plane inside or on the circumference of at least one circle (*colored circles*). The radius *R* of the covering circles is the distance from the center of a circle to the nearest "deep hole" in the packing. If the deep holes are connected by the appropriate straight lines, the plane is partitioned into regular hexagons; for uniformly distributed data points the average quantizing error for quantizing points that are centered on regular hexagons is $5\sqrt{3}/108$, or about .0802, which is slightly less than 1/12.

hexagons having the same area as the squares, and if each quantizing point is at the center of a hexagon, the average quantization error can be reduced to $5\sqrt{3}/108$, or about .0802.

Remarkably, it turns out that a corresponding improvement can always be made even if the data points are not evenly distributed. In a 1963 doctoral dissertation at Stanford University, P. L. Zador showed that it is always possible to reduce the average error by quantizing in a space of higher dimensions. It is more efficient to wait until several data points have been collected and then quantize them all at the same time by regarding them as a point in n-dimensional space than it is to quantize them one at a time along a single axis. In quantizing it pays to procrastinate.

Unfortunately Zador's result, like Shannon's theorem, is nonconstructive. The problem of finding good multidimensional quantizers even for uniformly distributed data is still unsolved. There are several sphere packings, however, that appear to give rise to excellent quantizing schemes. Consider the packing of spheres in two dimensions, that is, the packing of circles. It has been known since 1940 that circles can be packed at maximum density if they are arranged so that each circle is surrounded by six others [*see illustration on opposite page*].

Imagine now that each circle in the packing is bounded by a thin, flexible membrane and that the interior of the circle is inflated. As the circles get bigger the membranes press against one another to fill the remaining space in the plane; if the inflation is uniform throughout the plane, each circle will expand to form a regular hexagon. As I have mentioned, basing the quantization of uniformly distributed data on hexagonal regions leads to the minimum average error. A similar expansion of the eight-dimensional spheres in the E_8 packing also leads to low quantization error, which is even smaller than the error in two dimensions. The general problem of quantization, which requires that space be broken up, or covered, by a discrete partition, is closely related to the problem of finding the best covering of space by spheres.

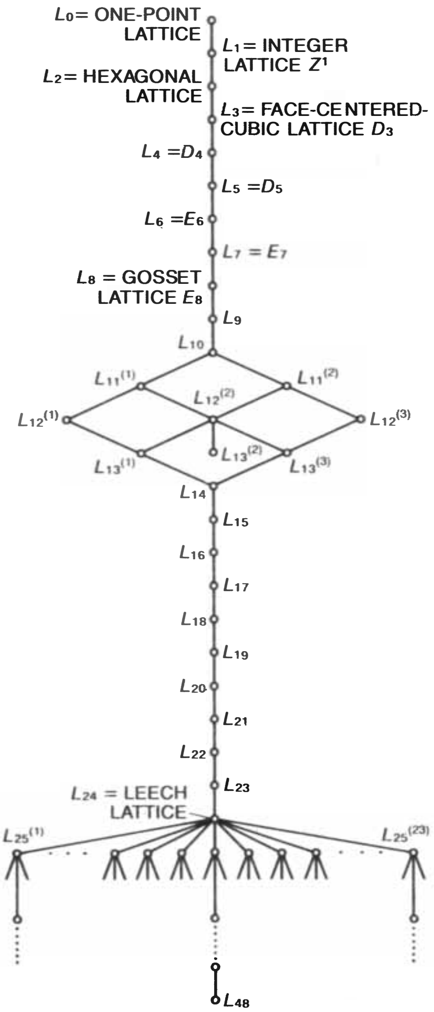
The search for dense sphere packings in multidimensional spaces is greatly simplified by focusing attention on certain kinds of packings called lattice packings that have a highly regular configuration. Consider the hexagonal packing of circles I have just described: notice that the centers of any two adjacent surrounding circles and the center of the circle in the middle form an equilateral triangle. In order to calculate the density of the packing it is sufficient to determine the proportion of each triangle that is covered by circles or parts of circles. Because the triangles tessellate, or completely fill, the plane and because the configuration of circles within a triangle is always the same, the circle density in one triangle is equal to the density of the packing throughout the plane. By applying elementary geometry one can show the density is equal to $\pi\sqrt{3}/6$, or approximately .9069.

The foregoing calculation would not have been possible if no repeating unit that fills the plane could have been found. It is nonetheless easy to imagine that there exist highly irregular sphere packings that have no repeating units. Such packings are much harder to study than those that repeat: not only is the density of an irregular packing difficult or impossible to determine but also even the coordinates of all the center points may not be specifiable. The definition of a lattice packing guarantees that these disadvantages can be avoided. A sphere packing is said to be a lattice packing if whenever there are two spheres, one that has its center at the point $(u_1, u_2, \ldots, u_n)$ and the other at the point $(v_1, v_2, \ldots, v_n)$, there are also spheres in the packing with centers at all points of the form $(au_1 + bv_1, au_2 + bv_2, \ldots, au_n + bv_n)$, where a and b are any whole numbers. The center coordinates of the latter spheres are said to be generated from the center coordinates of the first two spheres.

The simplest lattice packing is the cubic lattice packing, in which the center coordinates of each sphere in the packing are integers; the cubic lattice in an arbitrary number of dimensions n is designated Z^n. The one-dimensional "cubic" lattice Z^1 is made up of discrete line segments, each of unit length, that are centered at the integer points on the line. The "spheres," or line segments, cover 100 percent of the line, and each sphere touches two others; hence Z^1 solves the sphere-packing problem and the kissing-number problem in one dimension.

In two dimensions, however, the square lattice Z^2 is not the densest packing. Its packing density is $\pi/4$, or about .7854, far less than the density of the hexagonal packing of circles [*see illustration on page 118*]. In a like manner the sphere-packing density of the cubic lattice Z^3 is relatively low: it is $\pi/6$, or about .5236. A much denser family of lattice packings can be generated from cubic lattices if the spheres are centered at alternate points of the lattice in checkerboard fashion. To construct the new family of lattices color the points of a cubic lattice alternately red and black and make the centers of the spheres coincide with the black points. Equivalently, the centers of the new packing are the points with integer coordinates that add up to an even number. In an arbitrary number of dimensions n this packing is designated D_n. In D_3, for example, the origin (0,0,0) and the point (1,1,0) are legal centers of spheres, but the point (1,0,0) is not because $1 + 0 + 0$ is an odd number.

The sequence of lattice packings D_3, D_4, D_5 and so on is of considerable importance for the sphere-packing problem. The packing D_3 is the face-centered-cubic lattice. A model constructed of Ping-Pong balls shows that one repeating unit cell of the lattice is a cube two units on a side having a sphere at its center; the radius of each sphere is $\sqrt{2}/2$. The density of the packing can be cal-



LAMINATED lattice packings of spheres in n dimensions are built up by packing together layers of a suitable laminated lattice packing in the next smaller dimension $n - 1$. For example, the two-dimensional hexagonal packing L_2 can be built up by stacking rows of circles whose centers are arranged according to the Z^1 packing. Similarly, layers of spheres, each layer arranged according to the hexagonal lattice packing, can be stacked to yield the lattice packing L_3, the densest packing known in three dimensions. John Horton Conway of the University of Cambridge and the author have continued the construction and found all the laminated lattices in dimensions up to 25. The laminated lattices in dimensions 1 to 10 and in dimensions 14 to 24 are unique; there are two laminated lattices in dimension 11, three each in dimensions 12 and 13, 23 in dimension 25 and at least 75,000 in dimension 26.

culated by finding the fraction of the volume of the cube that is filled by spheres; it is equal to $\sqrt{2}\pi/6$, or about .7405. Although it is possible that denser sphere packings exist in three dimensions, Carl Friedrich Gauss proved in 1831 that D_3 is the densest three-dimensional lattice packing. It is also known that D_4 and D_5 are the densest lattice packings in four and five dimensions.

Above five dimensions, however, D_n is not the densest lattice packing, and by the time one reaches D_8 there are huge gaps between the spheres. The gaps are so large that it is possible to slide another copy of D_8 into the holes without overlapping the spheres; the result turns out to be the E_8 lattice. In 1934 H. F. Blichfeldt of Stanford University proved that E_8 is the densest lattice packing in eight dimensions, and he showed that certain cross sections of E_8 called E_6 and E_7 are the densest lattice packings in six and seven dimensions. In these dimensions no denser packings, which would necessarily be nonlattice packings, have been found since.

In 1965 John Leech, then at the University of Glasgow, constructed a remarkable sphere packing in 24-dimensional space; his construction is outlined in the illustration below. The study of the Leech lattice, as the packing is usually called, has led to a deeper understanding of the properties of other higher-dimensional lattices and to important results in the theory of groups. The lattice is almost certainly the densest sphere packing possible in 24 dimensions. C. A. Rogers, arguing as he did for sphere packing in three dimensions, gave bounds for the maximum density of packings in any n-dimensional space; his bound for any 24-dimensional sphere packing is only slightly greater than the density of the Leech lattice. Each sphere in the lattice touches 196,560 others, and in 1979 A. M. Odlyzko of Bell Laboratories and I proved that this number solves the kissing-number problem in 24 dimensions. The same method of proof also solved the kissing-number problem in eight dimensions; there the answer is 240, the number of spheres that touch one sphere in the E_8 lattice. These two results were found independently by V. I. Levenshtein of the L. V. Keldysh Institute of Applied Mathematics in Moscow. Incidentally, the problem remains unsolved in all other dimensions except one, two and three (where the answers are respectively two, six and 12).

The Leech lattice has been indispensable to group theorists for the construction of certain finite simple groups. These groups are the building blocks of all groups having a finite number of elements; their classification, which has just been completed, has preoccupied many mathematicians for more than 50 years. The simple groups play much the same role in group theory that the prime numbers play in number theory and the chemical elements play in chemistry. Several important simple groups have now been constructed by considering the set of all the rigid rotations and reflections of the Leech lattice that leave the central sphere fixed and permute the surrounding spheres. This set of operations is called the symmetry group of the packing; the analogous symmetry group of the face-centered-cubic packing in three dimensions is shown in the illustration on pages 116 and 117.

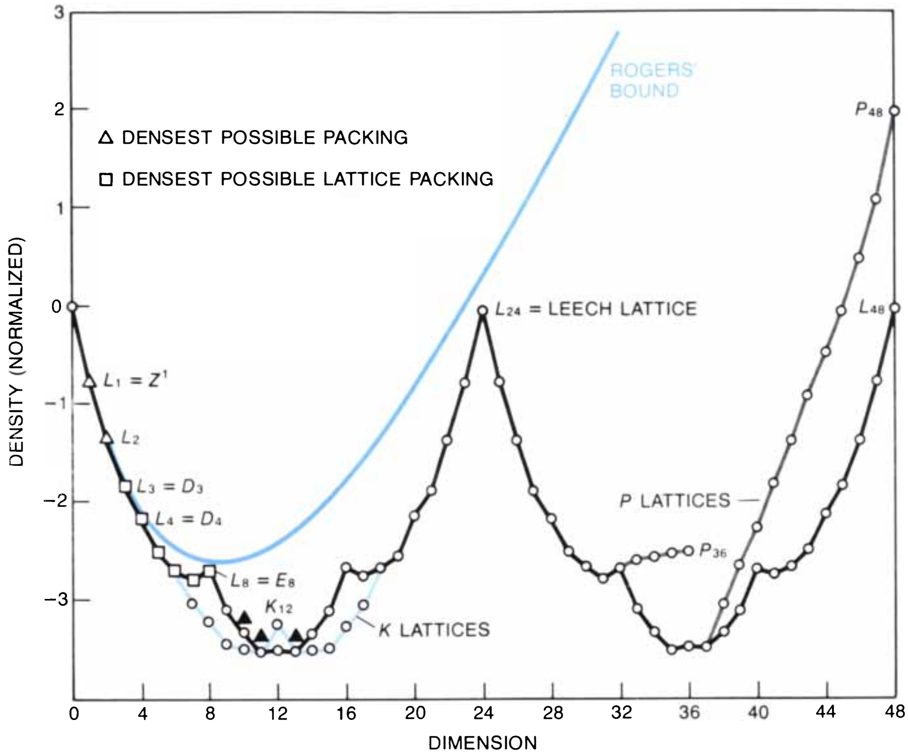
The symmetry group of the Leech lattice was found in 1968 by John Horton Conway of the University of Cambridge. Its order, or the number of elements in the group, is immense, although not particularly large in the context of group theory: it is $2^{22} \times 3^9 \times 5^4 \times 7^2 \times 11 \times 13 \times 23$, or 8,315,553,613,086,720,000. From this group, which is not a simple group, Conway constructed three previously unknown simple groups whose order in each case exactly divides the order of the symmetry group of the Leech lattice. In 1981 the Leech lattice enabled Robert L. Griess, Jr., of the University of Michigan to construct one of the last finite simple groups to be found. It is rather larger than Conway's groups and has been nicknamed the monster: the number of elements in the group is $2^{46} \times 3^{20} \times 5^9 \times 7^6 \times 11^2 \times 13^3 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times \times 71$, or 808,017,424,794,512,875,886,-459,904,961,710,757,005,754,368,000,-000,000. Griess's construction is not at all straightforward, however, and one of the fascinations of the Leech lattice is that one feels there should be a more direct connection between it and the monster simple group.

The Leech lattice is such a dense packing that its influence is felt in all lower dimensions. It is not surprising that a slice of a good packing gives a good packing in a space of one lower dimension; for example, one slice through D_3 exposes a surface of hexagonally packed spheres. Suitable cross sections of the Leech lattice, however, give rise to the densest known packings in all dimensions less than 24 except 10, 11 and 13. For example, one eight-dimensional slice through the Leech lattice exposes the E_8 lattice.

Since dense packings can be built from the Leech lattice, so to speak, from the top down, it is tempting to ask how the Leech lattice might be built from the bottom up, that is, from dense packings in lower dimensions. It turns out its construction can be done in a particularly simple way. Begin with the densest packing possible in one dimension, Z^1. At the center of each one-dimensional sphere in Z^1 construct a two-dimensional sphere whose radius is 1/2. Now construct another layer of two-dimensional spheres identical with the first layer and fit it into the holes in the first layer as tightly as possible. If an infinite number of layers are packed together in this way, the resulting two-dimensional lattice is the dense, hexagonal packing; because of this construction it can be

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	1	1	0	1	1
0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	1
0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	1
0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1	1	1
0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	0	1	1
0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	0	0	1
0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	0	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	1
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	1
0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	1	0	1
0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	1	1
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0

CONSTRUCTION OF THE LEECH LATTICE, the densest known packing of spheres in 24-dimensional space, is based on the 24-digit binary sequences shown. The set of all possible sums of the 12 binary sequences, where the addition is carried out modulo 2, is made up of 2^{12}, or 4,096, binary sequences called code words. (In addition modulo 2 the sum of 1 and 1 is equal to 0; the digit ordinarily carried in binary arithmetic is ignored.) The 2^{12} code words make up an efficient code for transmitting information that was devised by Marcel J. E. Golay at the U.S. Army Signal Corps Engineering Laboratories in 1949. The centers of the spheres in the Leech lattice all have the form $2C + 4X$ or $I + 2C + 4Y$, where C is a code word of the Golay code, I is the point (1,1,...,1) in 24 dimensions and X and Y range over all points in 24 dimensions whose coordinates are all integers. The sum of the coordinates of each point X must be even and the sum of the coordinates of each point Y must be odd. Each sphere has radius $2\sqrt{2}$, and the spheres closest to the origin have centers such as $(\pm4,\pm4,0,0,\ldots,0)$, $(\pm2,\pm2,\pm2,\pm2,\pm2,\pm2,\pm2,\pm2,0,0,\ldots,0)$ and $(\mp3,\pm1,\pm1,\ldots,\pm1)$. Each sphere touches 196,560 others.



DENSEST KNOWN PACKINGS of spheres in spaces of up to 48 dimensions are plotted according to a method suggested by John Leech; the "normalized" density of the packing depends on the dimension in which the packing is done. Its definition is based on the fact that the density of the 24-dimensional Leech lattice, divided by the volume of a 24-dimensional sphere of unit radius, is equal to 1. (The volume of an n-dimensional sphere of radius 1 is equal to $\pi^{n/2}/(1 \times 2 \times 3 \times \ldots \times n/2)$ if n is even, or to $2(2\pi)^{(n-1)/2}/(1 \times 3 \times 5 \times \ldots \times n)$ if n is odd.) That quotient in an arbitrary space of dimension n is called the center density D. The normalized density shown on the graph is equal to $\mathrm{Log}_2 D + n(24-n)/96$; for laminated lattice packings the graph is symmetrical about the normalized density of the Leech lattice. The graph shows that L_3, L_8 and L_{24} are quite close to the least upper bound known for the density of any sphere packing. The laminated lattice packings are the densest known packings in all dimensions up to 32 except dimensions 10 to 13. There is an alternative sequence of lattices called the K_n sequence, which starts at L_6 and rejoins the laminated sequence at L_{18}. The K_n packings are denser than the L_n packings in dimensions 11, 12 and 13. The P_n packings are also lattice packings, but the densest known packings in 10, 11 and 13 dimensions are nonlattice packings; all these packings are constructed from codes for the digital transmission of information.

called the laminated packing in two dimensions, abbreviated L_2.

I have already suggested how one goes to three dimensions: a ball of radius 1/2 is emplaced at the center of each circle in L_2, and identical layers of balls are fitted into the holes of the first layer in such a way as to form a lattice. Since this packing is equivalent to D_3, the laminating procedure gives the best-known packing in three dimensions as well as in two. If the procedure is continued in higher dimensions, one dimension at a time, the resulting lattice packings are extremely dense. It has been known for a long time that L_4 and L_5 are equivalent to D_4 and D_5, and that L_6, L_7 and L_8 are equivalent respectively to E_6, E_7 and E_8. Hence the laminated lattices are the densest lattice packings possible in up to eight dimensions.

Recently Conway and I continued the laminating process a bit further: we found all the laminated lattices in up to 25 dimensions and examples of them for each dimension up to 48. It turns out that although there is a unique laminated lattice in every dimension up to 10, there are two distinct ways to stack layers of the L_{10} lattice. The two 11-dimensional lattices have the same density but different kissing numbers. There are three 12-dimensional and three 13-dimensional laminated lattices but only one such lattice for each dimension from 14 through 24. The laminated lattice L_{24} is the Leech lattice. It is only in spaces having 11, 12 and 13 dimensions where cross sections of the Leech lattice are known that have a higher density than the laminated lattices in those dimensions.

The laminated lattices are built, layer on layer, by fitting each new layer as snugly as possible into the deep holes of the preceding one. Hence the investigation of laminated lattices is closely related to the covering problem: enlarge every sphere in any packing just enough to encompass the deep holes, and by definition the packing must cover the space. In 1966 Leech conjectured that if each of the spheres in the Leech lattice is enlarged by a factor of $\sqrt{2}$, the spheres would cover all the points in the space; such a covering might well be the best-possible covering of 24-dimensional space. Leech's conjecture, however, was quite difficult to prove, primarily because of the complexity of the deep holes in the Leech lattice.

In the two-dimensional lattice L_2 it is clear there is only one kind of deep hole; the maximum distance from any point in the lattice is always a point bounded by three circles, and so there is nothing to distinguish one deep hole from any other. In the Leech lattice, however, Conway, Richard A. Parker and I classified 23 distinct kinds of deep hole before we were able to prove Leech's conjecture. Not surprisingly there turn out to be 23 distinct ways to stack copies of the Leech lattice together in 25-dimensional space: there are 23 different 25-dimensional laminated lattices. In 26 dimensions the laminated lattices number at least 75,000.

In dimensions higher than 25 much less is known. Hermann Minkowski demonstrated in 1905 that there are lattice packings in any dimension n whose density is greater than 2^{-n}. His argument, like others I have mentioned, is nonconstructive. In low dimensions the estimate is rather crude: when n is equal to 24, for example, it states merely that there exist packings whose density exceeds about 6×10^{-8}, whereas the Leech lattice is known to be more than 32,000 times denser. In the other direction Blichfeldt showed in 1914 that for arbitrarily large numbers n the density cannot exceed about $2^{-.5n}$. In spite of much effort there was essentially no improvement on this bound until the 1970's. G. A. Kabatiansky, Levenshtein and V. M. Sidel'nikov in Moscow then showed that the density in very high dimensions n cannot exceed about $2^{-.599n}$.

E. S. Barnes of the University of Adelaide, A. Bos of N. V. Philips' Gloeilampenfabrieken in Eindhoven, the Netherlands, Conway, Leech and I have constructed a number of explicit packings in high dimensions, but none of them is as dense as Minkowski's theorem promises. Recently Barnes and I constructed lattice packings from the Leech lattice in dimensions up to 100,000. The density of these packings is roughly $2^{-1.25n}$, which at first glance seems almost as dense as the bound in Minkowski's theorem. Indeed, in dimension 65,536 our packings are about $10^{40,000}$ times denser than any lattice packing previously known. Unfortunately the exponential form of the bound can obscure the fact that we still fall rather short of the promised goal. The theorem guarantees there remain packings to be discovered that are $10^{4,000}$ times denser than the ones we have found.

The Invention of the Balloon and the Birth of Modern Chemistry

The first manned balloon flights in France 200 years ago were inspired by basic research into the nature of gases by some of the leading chemical investigators of the day

by Arthur F. Scott

The first free flight by human beings was a balloon ascension from the gardens of the Château de la Muette in the western outskirts of Paris on November 21, 1783. The passengers were Pilatre de Rozier, the young director of a museum of science in Paris, and the Marquis d'Arlandes, an army officer with good connections at the court of Louis XVI. In a hot-air balloon designed by the brothers Joseph-Michel and Jacques-Étienne Montgolfier the two passengers stayed aloft for some 25 minutes and came down unharmed in open country near the road to Fontainebleau, having traveled about five miles.

The flight was remarkable in its own right, but it also epitomized a major achievement in chemistry, namely the fall of the phlogiston theory of chemical composition under the impact of the discovery that gases are distinguishable by weight. The names of four preeminent chemists—Joseph Black, Henry Cavendish, Joseph Priestley and Antoine Lavoisier—are entwined in the records of the first balloon flights, manned and unmanned. Their work opened the way to the first clear understanding of the chemical nature of matter.

The Montgolfier brothers lived in Annonay, a town near Lyons. They had been fascinated with the idea of flight, and it occurred to them that a paper bag filled with smoke from a fire would rise in the air. Late in 1782 they carried out two preliminary experiments, which satisfied them that a bigger bag filled from a sufficiently large fire could be made to float. The brothers first demonstrated their idea publicly in Annonay on June 4, 1783. The balloon bag was a spherical sack made out of linen and lined with paper. It was 36 feet in diameter and weighed about 500 pounds. The balloon was inflated over a fire fed with small bundles of chopped straw. When it was released, it rose to a considerable height and came down in about 10 minutes, having traveled a mile and a half. Great excitement ensued, and news of the experiment spread quickly throughout France and the rest of Europe.

A different group carried out the next balloon trial in Paris two months later. The experiment was supervised by a physicist, Jacques Charles. Exploiting his knowledge of recent discoveries in the study of gases, Charles had decided to inflate the balloon with hydrogen. Since hydrogen would easily escape through a paper liner, the balloon bag was made out of thin silk fabric and coated with a solution of rubber. The hydrogen was obtained from the action of sulfuric acid on iron filings. Inflating the bag to its final diameter of 13 feet took several days and consumed nearly 500 pounds of acid and 1,000 pounds of iron. An immense crowd watched the ascension from the Champs de Mars on August 27. The balloon stayed up for 45 minutes, finally landing in a field near Gonesse, about 15 miles away, where it so terrified the inhabitants that they tore it to shreds.

Some three weeks later the Montgolfier experiment was repeated in Versailles, this time with Louis XVI and his court in attendance. Filling the fire balloon was a simple task compared with the preparation of the hydrogen balloon, and within 10 minutes the balloon was ready to take off. For this demonstration the balloon was fitted with a small cage carrying a sheep, a rooster and a duck. The balloon itself was not the plain bag of the first ascent but was brightly painted in oil colors. The flight ended in a wood about two miles away. The first aerial travelers were uninjured.

Once the feasibility of balloon flight had been established the practical possibilities were quickly tested. In October de Rozier was lifted to 80 feet in a tethered Montgolfier balloon and remained aloft for more than four minutes. A month later de Rozier and d'Arlandes made their historic free flight across Paris. Not to be outdone, Charles embarked with a passenger from Paris on December 1 in a hydrogen balloon. This flight lasted for two hours and took them some 27 miles to the small town of Nesle. There the passenger was let off and Charles continued alone, ascending to an altitude of about two miles. In a span of six months man had taken to the air and had learned how to fly.

In the years following 1783 there were many balloon flights all over Europe. Particularly noteworthy was an ascent by Joseph Montgolfier, the only one of the two brothers to have actually flown. He made his flight on January 10, 1784, from Lyons. The balloon, named *Flesselles,* was the largest one up to that date, standing more than 180 feet high and measuring 100 feet around. It was inflated from a straw fire in 17 minutes and rose to a height of 3,000 feet, carrying seven people.

Success followed success in ballooning. In August, 1784, a French chemist, Guyton de Moreau, accompanied by Abbé Bertrand, made a flight to more than 10,000 feet to collect data on the temperature and pressure of the atmosphere. In January of the following year the first crossing of the English Channel (from Dover to Calais) was achieved by Jean Pierre Blanchard, a French aeronaut, and John Jeffries, an American physician.

Following the first flight at Annonay

FIRST PUBLIC DEMONSTRATION of the flight of a hot-air balloon is captured in this somewhat fanciful contemporary print. The test was conducted by the brothers Joseph and Étienne Montgolfier near their home in Annonay in France on June 4, 1783. The balloon, a spherical sack of linen lined with paper, measured 36 feet across and weighed about 500 pounds. It was inflated over a fire fed with bundles of straw. The flight lasted for 10 minutes.

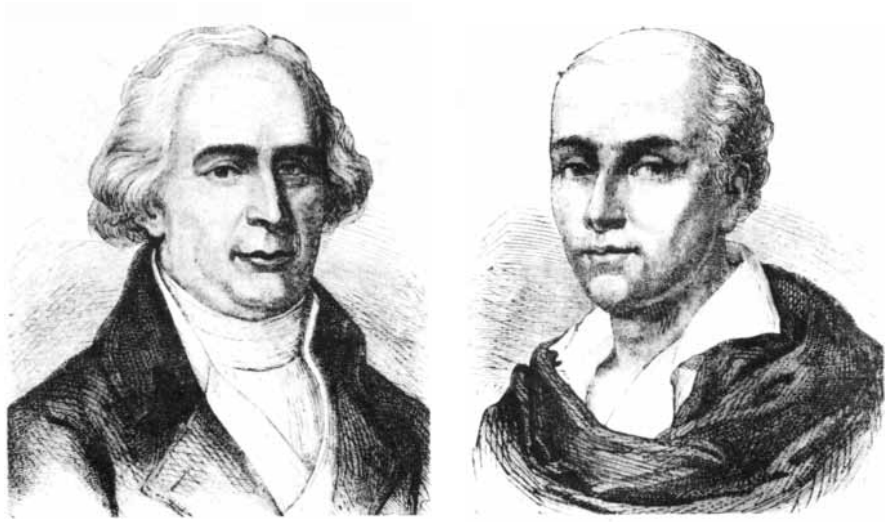
FIRST FREE FLIGHT of a manned hot-air balloon took place over Paris on November 21, 1783. The elaborately decorated balloon, designed by the Montgolfier brothers, was 46 feet across and 70 feet high. Including its two passengers, Pilatre de Rozier and the Marquis d'Arlandes, the rig weighed a total of 1,600 pounds. The view is said to be from the terrace of Benjamin Franklin's house at Passy.

the French Academy of Sciences, at the request of the government, appointed a commission to report on that experiment and to plan future experiments. Lavoisier, the French chemist whose scientific discoveries were among those underlying the first balloons, was one of the commissioners and played a leading part in the work of the commission. The French government evidently regarded the balloon invention as a significant one, since it covered the expenses of some of the later flights directed by the commission.

The reaction of British scientists to the balloon was less enthusiastic. In November, 1783, King George III and the court at Windsor were treated to an exhibition of a hydrogen balloon. Impressed, the king wrote to Sir Joseph Banks, the president of the Royal Society of London, offering to subsidize further experimentation. The reply came back that since "no good whatsoever" could be expected from such experiments, the society was not interested.

The potential value of the balloon in warfare, however, was recognized quickly. A pamphlet on the subject was published in the same month as the demonstration at Windsor. Benjamin Franklin expressed the situation clearly in a letter dated soon afterward:

"The invention of the balloon appears, as you observe, to be a discovery of great importance. Convincing sovereigns of the folly of wars may perhaps be one effect of it, since it will be impossible for the most potent of them to guard his dominions. Five thousand balloons, capable of raising two men each, could not cost more than five ships of the line, and where is there a prince who could afford to cover his country with troops for its defense as that ten thousand men descending from the clouds might not in many places do an infinite amount of damage before a force could be brought together to repel them?"



MONTGOLFIER BROTHERS, Joseph-Michel (*left*) and Jacques-Étienne (*right*), are pictured in these 19th-century line engravings. At the time of their first public balloon flights Joseph was 43 and Étienne 38. The image of Étienne is copied from a portrait by his daughter.

The spectacular developments in balloon flight were an inevitable consequence of a drastic change in the human understanding of matter. At the time the only science worthy of the name was mechanics, in particular celestial mechanics, the study of the motion of the heavenly bodies. Chemistry was just being freed from the dogma of alchemy, and biology and the other natural sciences were still in the early observational stage. It was a time when a scholar might still master all science and was properly referred to as a natural philosopher. Four of these natural philosophers greatly influenced the invention of the balloon: Black, Cavendish, Priestley and Lavoisier, all of whom would today be called chemists.

To anyone familiar with modern scientific concepts the primitive state of chemistry in the early 18th century is bewildering. The alchemical idea that all matter was compounded from four terrestrial elements—air, earth, fire and water—was still widely held. This simple notion had first appeared in Aristotle's *Natural Science* some 20 centuries earlier and had led, among other things, to the belief that different kinds of matter could be transmuted into one another. One result was the illusory quest for the philosopher's stone, which was said to be capable of turning iron and lead into gold. Another descendant of Aristotle's ideas was also in vogue at that time, namely the phlogiston theory, which was to cloud and confuse the thinking of chemistry for most of the 18th century.

The phlogiston theory was developed to explain the nature of fire. From man's earliest prehistoric experience it had been recognized that certain substances burn and others do not. The alchemists had enlarged on this observation, noting that with intense enough heat even the common metals would burn, leaving a calx, or ash, that could not be burned. Why was this so?

The explanation put forward by Georg Stahl in the first decades of the 18th century was based on an earlier idea of his teacher, Johann Becher. Becher had expanded the Aristotelian elements to include *terra pinguis* (fatty earth), which was supposed to be released from a substance when it burned. Stahl took the idea further, asserting that a metal was nothing more than a compound of its calx and fatty earth. Viewed from this perspective, burning was the release of the stored fatty earth, which Stahl renamed phlogiston, from the Greek for "flammable."

The phlogiston theory was as versatile as it was successful. It explained, for example, why a calx heated with charcoal would revert to the original metal: the charcoal, a flammable substance, is rich in phlogiston, whereas the calx, which does not burn, has none. Hence the charcoal transfers its phlogiston to the calx, thereby regenerating the metal and leaving charcoal ash. Successes such as this one propelled the theory into general acceptance, and the idea of phlogiston came to dominate chemical thought for most of the century that followed.

Black, Cavendish, Priestley and Lavoisier were all confirmed phlogistonists when they embarked on their chemical careers. Indeed, the first three were engaged chiefly in making experimental discoveries they often interpreted in terms of the phlogiston theory. Lavoisier, however, had the genius necessary to organize the new findings, along with older facts, into a system of chemistry that had no place for phlogiston. Ironically, although the work of Cavendish and Priestley contributed as much as anything to the overthrow of the phlogiston theory, neither man ever gave up the theory. Priestley remained so convinced in 1800, well after oxygen had become accepted as the cause of fire, that he called his final book *The Doctrine of Phlogiston Established.*

The first break with the chemical ideas of Aristotle stemmed from the experiments of Jan van Helmont, who made the results of his studies known early in the 17th century. Although the alchemists were aware earlier that gases were produced in reactions such as fermentation or the burning of charcoal, they regarded the gases as a form of ordinary air. By means of simple chemical experiments van Helmont produced gases he could distinguish from ordinary air. He gave them such graphic names (in translation) as windy gas, fat gas and smoky gas. Van Helmont made no chemical tests of these gases, nor

did he attempt to isolate them, but he is credited with introducing the word "gas" into the scientific vocabulary. The study and understanding of gases developed into a branch of chemistry known as pneumatic chemistry, and van Helmont is generally seen as its founder.

The field of pneumatic chemistry lay relatively dormant until the middle of the 18th century. Then it suddenly took on new life and new direction, with surprising results. In the words of Torbern Bergmann, the 18th-century Swedish chemist: "During the past 10 years chemistry has not only soared into regions of invisible aerial substances, but it has dared to explore the nature of these substances, and to search for their constituent principles." Indeed, by 1779, when those words were written, the chemical composition of eight gases was already known with certainty.

Although Black, Cavendish, Priestley and Lavoisier were not close associates, their scientific contributions built on one another and developed into what is now known as the scientific revolution. The chain of discoveries was initiated by Black. In the 1750's, while he was a medical student at Edinburgh, he undertook a careful examination of the gas released by the action of acids on solid magnesia (magnesium carbonate). His primary goal was to understand the antacid properties of magnesia, but in the process he established that the gas released was a chemical entity distinct from atmospheric air.

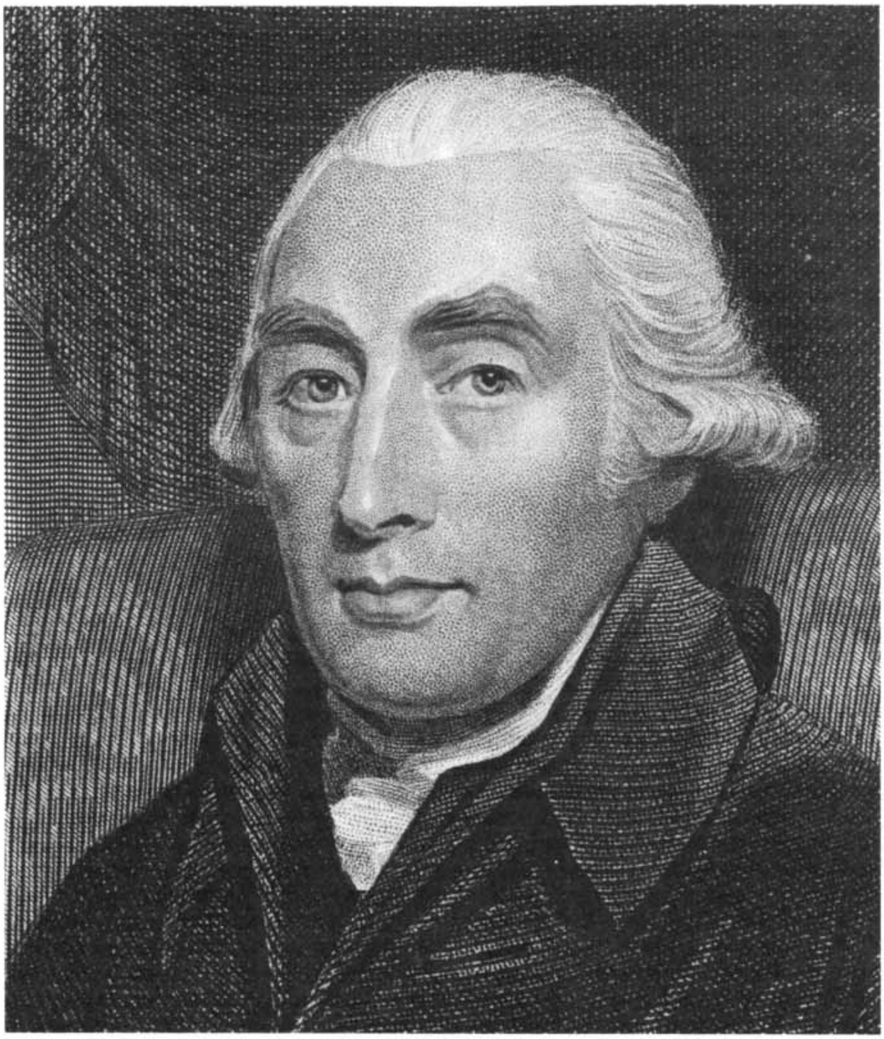
Black named this new substance "fixed air," since it seemed to be fixed or trapped inside the magnesia. At the time it was not yet understood that this gas was a compound of chemical elements, and several decades passed before it was renamed carbon dioxide after its atomic composition. By passing gas bubbles through limewater and looking for a milky precipitate Black was able to show that fixed air was released in the burning of charcoal, in respiration and in fermentation. He became one of the foremost chemical philosophers of his age, serving first as professor of chemistry at Glasgow and later returning to Edinburgh in the same capacity.

One of the first scientists to look at the properties of fixed air was Cavendish. He was the stereotype of the early natural philosopher: rich, eccentric and reclusive. Having inherited vast wealth (at one time he was among the richest men in England), Cavendish chose to live his life alone and to devote himself to experiment. In 1766 he published three papers titled *Experiments on Factitious Air.* By factitious air Cavendish meant any kind of gas that is "contained in other bodies...and is produced from them" by chemical manipulations. Before Cavendish only Black's fixed air was known to be factitious. Cavendish had followed Black's method of producing fixed air by adding acid to magnesia and had extended his work by trapping samples of the gas in animal bladders. By weighing the bladder filled first with atmospheric air and then filled with fixed air Cavendish discovered that fixed air was 1.47 times heavier than normal air.

Propelled by curiosity, Cavendish went further. What would happen if the magnesia in Black's experiment were replaced with a common metal such as iron? Once again gas bubbles were observed and once again Cavendish collected the gas in a bladder. This factitious gas, however, did not give rise to a precipitate in limewater and turned out to be 11 times lighter than air. Moreover, instead of extinguishing a flame as fixed air did, it produced an explosion when a flame was brought near it. Cavendish had clearly discovered a second factitious gas, which he appropriately named "inflammable air."

The work of Black and Cavendish firmly established that gases were separate chemical entities. It was no longer possible to think of air as one of the elementary constituents of matter. But what about earth, fire and water?

At about this time Lavoisier, a young French aristocrat, undertook a simple test that eliminated "earth" as an elementary substance. An old experiment, which had suggested that it might be, involved heating water for an extended period in a closed "pelican." (That was what the vessel, a retort, looked like.) In the end a small amount of solid was visible at the bottom of the vessel. This ex-



JOSEPH BLACK (1728–99) initiated the chain of discoveries that led to the invention of the balloon. In the 1750's, while he was still a medical student at Edinburgh, he identified the gas released by the action of acids on solid magnesia as a chemical entity distinct from atmospheric air. The new substance, which he called "fixed air," was subsequently renamed carbon dioxide. This steel engraving of Black is a copy of a painting made by Sir Henry Raeburn.

FIRST FREE FLIGHT of a manned hydrogen-filled balloon was launched from the Tuileries Gardens in Paris on December 1, 1783. The passengers were the physicist Jacques Charles and an assistant, M.-N. Robert. The flight lasted for about two hours, after which Robert got off and Charles continued alone, ascending to an altitude of some two miles.

periment had been interpreted as the conversion of water into "earth."

Lavoisier decided to test the accepted interpretation of this experiment with the chemical balance. He weighed the empty pelican and added purified water to it. After sealing the top he recorded the weight of the filled vessel and by subtraction ascertained the weight of the water. He then allowed the water to simmer for 101 days. At the end of the experiment he weighed the filled and empty pelican again. The combined weight had not changed, but a solid had appeared whose weight was equal to the loss in weight of the empty vessel. Obviously the "earth" had leached out of the glass pelican and had not come from the water. Combined with the work of Black and Cavendish, Lavoisier's experiment cast grave doubt on Aristotelian alchemical theory.

The final developments in the scientific revolution hinged on further discoveries in pneumatic chemistry, mainly by Priestley. He was a man of diverse talents and interests. Trained as a Nonconformist clergyman, he had several careers, including publicist, author (of 106 books) and chemist. In 1772 he published a paper titled *Observations of Different Kinds of Air*, in which he described the preparation of several new gases. In the next decade, which was the one just preceding the invention of the balloon, Priestley added eight more gases to his list of discoveries.

At the heart of Priestley's success lay a major improvement in the technique for recovering gases. Previously chemists had collected gases by bubbling them into a water-filled bell jar that had been carefully inverted and set in a pan of water. As gas collected at the top of the inverted jar, water was displaced from the jar into the pan. If a gas was soluble in water, however, it could not be collected by this technique. By the simple device of substituting liquid mercury for water, Priestley could collect and analyze many new gases.

Priestley's most important discovery came in 1774. With sunlight and a 12-inch burning lens he heated the red precipitate of mercury, a powdery substance long known to alchemists. Just as Black had observed in the burning of charcoal, a gas was produced, but it was not fixed air. Indeed, this gas had remarkable properties all its own: it made a candle burn brighter, and a mouse could live in it for nearly twice as long as it could in the same amount of atmospheric air.

Since Priestley adhered to the phlogiston theory of combustion, it was natural for him to interpret these properties in terms of phlogiston. Remember that phlogiston was supposed to leave a substance when it burned. Since it had to go somewhere, Priestley reasoned that it was going into the new gas. Thus the gas must be deficient in phlogiston, and Priestley called it "dephlogisticated air." This name was not destined to last for long. Soon, in the laboratory of Lavoisier, the gas would be renamed oxygen and would figure as the cornerstone of modern chemical theory.

It was at about this time that Lavoisier first began to seriously question the phlogiston theory. In 1772 he had prepared a memoir on the burning of sulfur and phosphorus in air. Again making use of the chemical balance, he determined that both substances gained weight in burning. He attributed the gain to their combining with air. In his memoir Lavoisier went on to speculate that "what is observed in the combustion of sulphur and phosphorus may well take place in the case of all substances... and I am persuaded that the increase in weight of metallic calces is due to the same cause." In agreement with his conjecture Lavoisier found on heating the calx of lead (lead oxide) with charcoal that "just as the calx changed into the metal, a large quantity of gas was liberated." These observations and speculations were in complete contradiction to the phlogiston theory, which taught that phlogiston escaped from a substance when it burned, with an accompanying loss of weight. The young Lavoisier, recognizing the heresy in these ideas, submitted his memoir as a sealed note to the French Academy, thereby ensuring his prior claim if such revolutionary thoughts were sustained by future work.

Lavoisier's mature thoughts on burning and the theory of combustion appeared in his famous memoir of 1783, *Réflexions sur le phlogistique* (Reflections on Phlogiston). There he summarized his many arguments against the validity of the phlogiston theory. The evidence was clear enough for him to write: "My only object in this memoir is to extend the theory of combustion

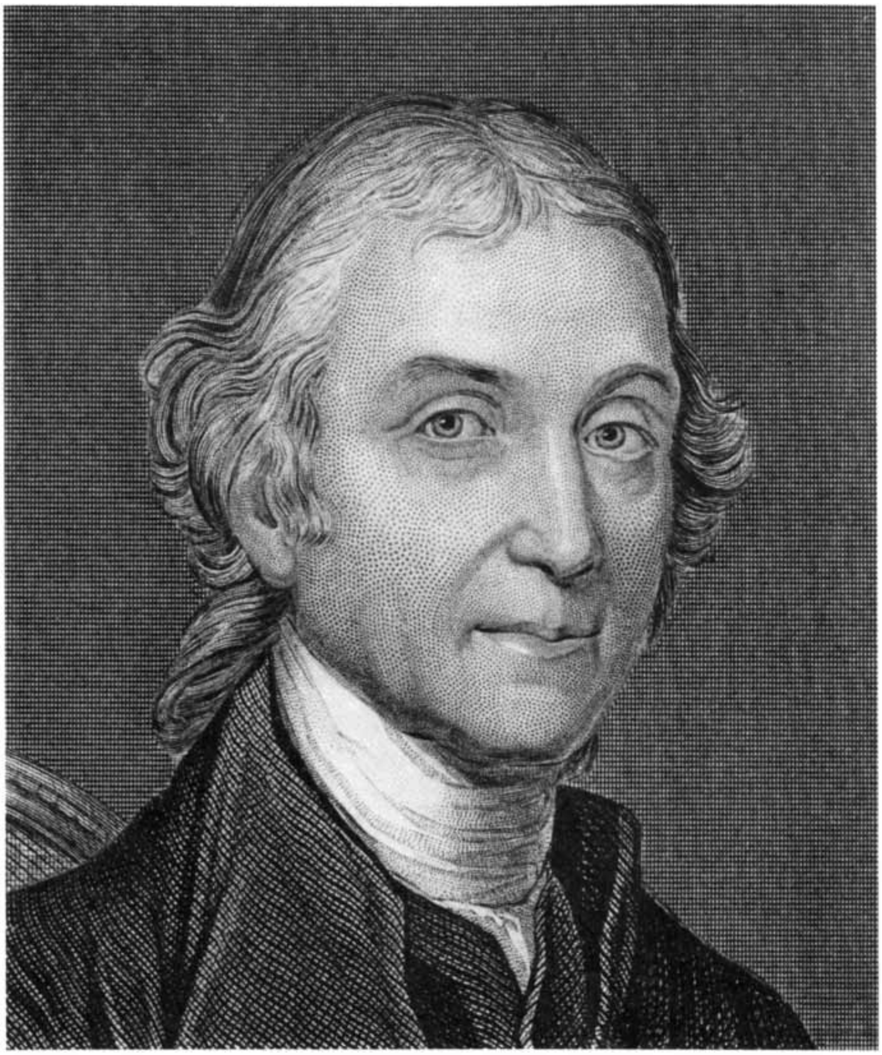
HENRY CAVENDISH (1731–1810) investigated the properties of both fixed air and another form of "factitious air" he called "inflammable air." The latter substance, which turned out to be 11 times lighter than atmospheric air, was later renamed hydrogen. This drawing by W. Alexander is the only known contemporaneous portrait of Cavendish. It is in the British Museum.

that I announced in 1777; to show that Stahl's phlogiston is imaginary and its existence in the metals, sulphur, phosphorus and all combustible bodies is a baseless supposition, and that all the facts of combustion and calcination are explained in a much simpler and easier way without it."

Lavoisier's explanation was indeed simple. It was not phlogiston that was lost during the combustion of a substance; it was oxygen from the air that combined with it. Lavoisier's hypothesis could explain all the facts about combustion. This included the quantitative ones, such as the fact that the weight gained by a substance in combustion was precisely equal to the weight of the oxygen gas that had disappeared.

By this time the results of the research in pneumatic chemistry were being widely disseminated. Great progress in the understanding of matter was being made, and gases with unusual new properties were being discovered. Based on his familiarity with the properties of Black's fixed air, Priestley had dissolved it in water and found that the resulting concoction was pleasant to drink. The new "soda water" became an immediate hit in European society. The Montgolfier brothers were also learning of the work with gases and translating it into their ideas about balloons. One more great stroke was required, leaving in its wake both the scientific revolution and the balloon.

The final piece of the puzzle was discovered in England. Cavendish had taken up Priestley's use of the electric spark to examine the new gases. Cavendish was particularly interested in the light gas he had earlier called inflammable air, and he was sparking mixtures of it and common air. The spark caused a bluish flame accompanied by contraction of the volume of the gas and the formation of a small amount of liquid, which he called a dew.



JOSEPH PRIESTLEY (1733–1804) succeeded in isolating eight more gases in the decade preceding the invention of the balloon. An adherent of the phlogiston theory of combustion, Priestley identified the most important of his discoveries—oxygen—as "dephlogisticated air."

It was the dew that attracted Cavendish's attention, and he set about designing a second experiment to collect larger quantities of it. The substance, he wrote, "had no taste or smell and ... left no sensible sediment when evaporated to dryness; neither did it yield any pungent smell during evaporation; in short, it seemed pure water." Indeed, further examination proved it to be just that.

This famous experiment was completed in 1781 but not reported to the Royal Society until 1784. Formal publication was delayed primarily because Cavendish wanted first to investigate his observation that when dephlogisticated air was substituted for common air in this experiment, the water produced was acidic. The cause of the acidity he found to be nitric acid, the composition of which he was able to establish for the first time.

Cavendish had shown that water was produced from inflammable air and dephlogisticated air when a mixture of the two gases was ignited. The finding did not mean to him that water is a compound of hydrogen and oxygen. Not to a confirmed phlogistonist! He wrote: "There seems the utmost reason to think that dephlogisticated air is only water deprived of phlogiston, and that inflammable air, as was before said, is either phlogisticated water, or else pure phlogiston, but in all probability the former." In other words, Cavendish believed water existed as water in each of the two "airs" and that the reaction between them released it. In the process the phlogiston was transferred from the phlogiston-rich inflammable air to the phlogiston-poor dephlogisticated air.

On November 12, 1783, Lavoisier read at a public meeting of the academy a memoir with the long title *Mémoire dans lequel on a pour objet de prouver que l'eau n'est point une substance simple, un élément proprement dit, mais qu'elle est susceptible de décomposition et de recomposition* (On the Nature of Water and on Experiments that appear to prove that this Substance is not properly speaking an Element, but can be decomposed and recombined). Although the methods and results were inferior to those of Cavendish, Lavoisier was bold enough to take the leap and assert that water was a compound of oxygen and hydrogen. Lavoisier had also devised a clever experiment showing that water could be broken down into its constituent elements. His idea exploited the fact that when steam reacts with red-hot iron (the iron was actually a gun barrel), the water decomposes to form hydrogen and iron oxide. Only when he found, in a side experiment, that copper at red heat did not react with water was he actually able to carry out the experiment he was seeking.

The successful experiment was set up in a copper tube containing small pieces of iron. A weighed sample of water was dropped into the red-hot tube. Any

Embark on a wondrous journey across the universe of science

SCIENTIFIC AMERICAN LIBRARY

An incomparable collection of enduring value

Out of *Scientific American's* collaboration with scientists in every field of endeavor, and after nearly a decade of planning, the SCIENTIFIC AMERICAN LIBRARY now offers a limitless resource for those intrigued by the wonder of science.

The SCIENTIFIC AMERICAN LIBRARY acquaints you with a world of understanding and shares with you the major ongoing advances at the frontiers of scientific investigation. Written by scientists who have made significant contributions, each volume rewards you not merely with information but with true understanding, as well as a command of the new concepts with which research proceeds.

As you receive your volumes you will see how they combine to make up a reference library of enduring value, one that embraces the unity of scientific understanding.

Share in this exciting enterprise with us. Simply fill in the attached card or coupon and we will send your splendid premier volume, POWERS OF TEN, to examine *free* for 15 days. You will be billed just $19.95 plus postage and handling for each volume, to be paid only *after* you have decided to add the volume to your personal library.

While you may cancel at any time with absolutely no notice, membership will bring you a cumulative library of scope and focus unparalleled in the literature of science.

YOUR PREMIER VOLUME

POWERS OF TEN

Philip and Phylis Morrison and the Office of Charles and Ray Eames

The essential survey of the universe from its largest dimension to its smallest

"If I were marooned on a desert island with just this one felicitous, densely informative book on all of science, I would count myself a lucky man. Having read it through, nothing on earth could prevent me from swimming off, any distance, to get another one like it."

LEWIS THOMAS, M. D.,
author of *The Lives of a Cell*

An atlas of the universe and a guide to its exploration, POWERS OF TEN takes you on an extraordinary adventure in magnitudes. This spectacular journey, with its sweeping changes in perspective, proceeds in ordered steps from the largest known dimension (10^{25} meters) to the smallest (10^{-16} meters).

You will find yourself abroad in a cosmos that in surprise and splendor beggars the past creations of man's innocent imagination. In vivid color illustrations, POWERS OF TEN shows you the universe scaled from the human dimension. You will know better not only where you are but who you are.

The SCIENTIFIC AMERICAN LIBRARY offers POWERS OF TEN as its premier volume because it sets out the agenda for the books that follow.

MUSICAL SOUND

John R. Pierce

John Pierce shares with you his joy in the discovery that understanding from physics and mathematics enlarges and enhances one's appreciation of music. He traces the connection between Pythagoras and Galileo to the current musical innovators — Pierre Boulez, John Cage, and Edgard Varese — now working at the frontiers of sound and music.

John Pierce imbues readers with his confidence that rational inquiry into the highly subjective experience of music will open up new realms to the enjoyment of that experience.

Two 33⅓-rpm sound sheets demonstrate acoustic illusions.

SUBATOMIC PARTICLES

Steven Weinberg

As everyone knows, the atom, once considered irreducible, is composed of more fundamental particles: electrons, protons, and neutrons. But how do we know this?

In THE DISCOVERY OF SUBATOMIC PARTICLES we come to know the answer. With the author we reenact the experiments that disassembled the atom. We are there as J. J. Thomson, Ernest Rutherford, and James Chadwick discover, in turn, the electron, the proton, and the neutron. We gain a confident understanding of how the laws of physics govern each event we encounter.

HUMAN DIVERSITY

Richard C. Lewontin

The relevance of science to human values finds compelling demonstration in our understanding of the genetics of our species. In HUMAN DIVERSITY, Richard Lewontin shows that each of us differs from all others because of the interaction of genetic and environmental differences coupled with chance events that occur during our development.

Lewontin makes clear that the human genetic endowment confers a plasticity of psychic and physical development; each of us has the potential to develop any of a great diversity of identities within the human ambit.

FOSSILS

George Gaylord Simpson

In FOSSILS, noted paleontologist George Gaylord Simpson sets out the splendor of evidence upon which the theory and facts of evolution rest. He explains how "the primary record of the history of life is written in the successive strata of rocks as in the pages of a book. Fossils may be called the writing on those pages."

In his lifework and in this book, Simpson presents a picture of evolution very different from that of "nature red in tooth and claw" depicted in the first crude Darwinian popularizations and persistent still in the popular imagination.

THE SOLAR SYSTEM

Roman Smoluchowski

Standing on the moon, men have seen the Earth "small and blue and beautiful." Beyond doubt, the space expeditions of the U. S. and the U. S. S. R. must be reckoned as the most successful, as well as the most extravagant, ventures in the popularization of science.

In THE SOLAR SYSTEM, Roman Smoluchowski assembles a wealth of images from satellites and earthbound telescopes to give us a magnificent picture of the Earth and its cosmic neighborhood.

ON SIZE AND LIFE

Thomas A. McMahon and John Tyler Bonner

A biologist and an engineer bring their disciplines to bear in showing how nature obeys its own physical laws in fashioning living things. They demonstrate how size relates to shape and how this in turn controls the upper limit to the sizes of animals and plants.

Beyond the diversity of life, the authors show orderly ratios and relations that surprise and delight: the time for 50 percent of a mammal's growth is about three percent of a lifetime, and every mammal, no matter how long it can live, will take about 330 million breaths, sustaining 1.5 billion heartbeats.

PERCEPTION

Irvin Rock

The eye, we say, is a camera. Consider, however, that an object appears to us unchanging in size, shape, or orientation despite changes in our position with respect to it. How can we explain this constancy of perception when we know that the size, shape, and orientation of an image projected by the lens of the eye upon the retina must vary as we move?

In PERCEPTION, Irvin Rock explores the problems posed by both correct and illusory perception and examines the traditions of thought which inform contemporary investigations of the subject.

A GUIDED TOUR OF THE LIVING CELL

Christian de Duve

Trillions of cells compose the human body, each cell about four orders of magnitude smaller than the world in reach of our unaided senses. In A GUIDED TOUR OF THE LIVING CELL, Nobel Laureate Christian de Duve takes us on an engrossing, beautifully illustrated expedition through a world far removed from everyday experience yet vital to our very existence.

A bold but reassuring guide, de Duve puts the wisdom and knowledge gained from a lifetime of research at our service with his narrative, nontechnical, and often anecdotal style.

In addition to these volumes, leading scientists are at work on the exciting future books in the SCIENTIFIC AMERICAN LIBRARY. James Watson is writing on DNA, David Hubel on vision and the brain, Julian Schwinger on space and time. In the pages of the SCIENTIFIC AMERICAN LIBRARY, these Nobel Laureates and other distinguished scientists will explore the world as we know it.

Special Offer to Readers of *Scientific American*
Membership Reservation Form

To SCIENTIFIC AMERICAN LIBRARY
P. O. Box 646
Holmes, Pennsylvania 19043

Please enroll me as a member of the SCIENTIFIC AMERICAN LIBRARY. Send me the premier volume — POWERS OF TEN — by return mail. Thereafter, I am entitled to receive for inspection a new volume every two months for 15 day's *free examination.*

I need send no payment now. For each volume I keep I will be billed $19.95 plus postage and handling. This price is guaranteed for at least one year, regardless of any price increases asked of new members.

I understand that I may cancel my membership at any time.

NAME ______________________________
(Please print clearly)

ADDRESS ______________________________

CITY ______________ STATE ______________ ZIP ______________

Canadian residents: U. S. dollars please

SSA014

steam that remained undecomposed was condensed and weighed; the gaseous fraction (hydrogen) was then collected over water and measured; finally the iron in the copper tube was reweighed to determine its gain in weight. From this experiment Lavoisier concluded that water was made up of one part of hydrogen and $6\frac{1}{2}$ parts of oxygen by weight. (The correct value is 1 : 8.) These graphic demonstrations that water can be taken apart and put back together again fit so neatly into Lavoisier's conception of nature that they dealt a final blow to the phlogiston theory. More than that, the experiment marked the end of the four Aristotelian elements.

Buoyed by the success of his interpretation of this experiment, Lavoisier was in a position to construct a new, logical system of chemistry. He set it forth in his *Traité élémentaire de chimie* in 1789. In it both oxygen and hydrogen appear among Lavoisier's list of 33 elements, all but two of which can be found in the modern periodic table. The publication of the book signified the beginning of the scientific revolution and the birth of modern chemistry.

Keeping in mind the story of the initial balloon ascents, the effect of the parallel revolution in chemistry on ballooning and particularly on the hydrogen-filled balloons seems obvious. Yet the interconnection of the two endeavors is even deeper. Although the story of the hydrogen balloon began in Cavendish's laboratory, where he first prepared his "inflammable air" and established that it is much lighter than atmospheric air, it was Black who applied the discovery to demonstrate for the first time the possibility of producing a lighter-than-air object. The following story by Thomas Thomson, the distinguished chemist who succeeded Black at Glasgow, sets forth Black's simple achievement:

"Soon after the appearance of Mr. Cavendish's paper on hydrogen gas, in which he made an approximation to the specific gravity of that body, showing that it was at least ten times lighter than atmospheric air, Dr. Black invited a party of his friends to supper, informing them that he had a curiosity to show them. Dr. Hutton, Mr. Clarke of Elden, and Sir George Clarke of Pennicuik were of that number. When the company invited had assembled, he took them into a room. He had the allantois [the thin fetal membrane] of a calf filled with hydrogen gas, and upon setting it at liberty, it immediately ascended and adhered to the ceiling. The phenomenon was easily accounted for: it was taken for granted that a small black thread had been attached to the allantois, that this thread passed through the ceiling, and that some one in the apartment above, by pulling the thread, elevated it to the ceiling, and kept it in position. This explanation was so probable that it was acceded to by the whole company; though, like many other plausible theories, it turned out wholly unfounded; for when the allantois was brought down, no thread was found attached to it."

Years later, in 1784, Black wrote a letter giving an account of his thinking. "As you speak of the 'birth' of aerostatic experiments," he said, "I beg leave to communicate to you more fully my thoughts on that subject. In the first place, although what I have already informed you of is strictly true, I by no means set up my claim for merit in the invention of machines for general flight and excursions. The experiment with the bladder, which I proposed as a striking example of Mr. Cavendish's discovery, was so very obvious that any person might have thought of it; but I certainly never thought of making large artificial bladders, and making these lift heavy weights, and carry men up into the air. I have not the least suspicion that this was thought of anywhere before we began to hear of its being attempted in France, and I do not doubt that what has been published in the newspapers is perfectly true, viz. that Mons. Mongolfier [*sic*] had sometimes before conceived the idea of flying up into the air by means of a very large bag or balloon of common air, simply rarefied by the application of Fire of Flame.

"The idea being founded upon a principle which has long been known, and which has no connection with Mr. Cavendish's discovery, it is only surprising that Mons. Mongolfier should not have put it sooner in practice. I suppose, therefore, that though he might have formed the Project a long time before, he never was roused into an operation for making the trial until others began to think of flying by means of inflammable air. Who first thought of the method I cannot tell, for I confess I did not read the history of the Experiments; they never interested me in the least."

ANTOINE LAVOISIER (1743–94) is shown in this hand-colored 19th-century engraving conducting an experiment designed to determine the composition of water by igniting a mixture of hydrogen and oxygen with an electric spark. It was Lavoisier who finally disproved the phlogiston theory of combustion and established the true compound nature of water. He also played a leading role in planning and reporting on the early balloon experiments, serving on a special commission formed for that purpose in July, 1783, by the French Academy of Sciences.

What speculations actually led the Montgolfier brothers to undertake their experiments with the hot-air balloon? This question is more difficult to answer, and what can be said is more in the way of speculation. James Glaisher, writing in the 1878 edition of the *Encyclopaedia Britannica,* states: "The Montgolfier brothers imagined that the bag rose because of the levity of the smoke or other vapour given forth by the burning straw: and it was not till some time later that it was recognized that the ascending power was due merely to the lightness of heated air compared to an equal volume of air at a lower temperature." Evidently the Montgolfier brothers were under the impression that the vapor given off by the straw was inflammable air or something similar. Black, however, knew better. He made it clear in his letter that the hotter air was simply "rarefied."

There is other evidence that the Montgolfier brothers at the time of their experiment were laboring under a misapprehension regarding the nature of the smoke and vapors from fire. In the correspondence of Sir John Sinclair, a lawyer who was active in English politics at the time, one finds the following story. "Towards the conclusion of the year 1785, some circumstances occurred, which induced me to take a short excursion from London to Paris, and accidentally I went in company with three distinguished foreigners, namely, Argand, so well known for his improvements in the art of making lamps; Reveillon, the greatest manufacturer of paper hangings then known . . . and [Joseph] Montgolfier, so celebrated for his discovery of balloons. I was able to obtain much information from the conversation of these intelligent men; and I remember, in particular, that the latter gave an account of the origin of his discovery, of which the following is the substance.

"Montgolfier said that he and his brother were paper manufacturers in Languedoc, but he had always felt a strong attachment to chemical inquiries. They were thence led to procure all the information they could regarding those subjects. It seems that Montgolfier and his brother had talked over the possibility of being able to ascend themselves, or to send up large bodies from the earth, at a very early period, without, however, having made any experiments to prove whether the idea was practicable or not; but happening to read an account of some experiments made by Dr. Black, which explained the nature of the various kinds of airs or gases, and, in particular, their differences in point of weight, he immediately said to his brother, 'The possibility of effecting what we talked about some time ago seems to be proved by a foreign chemist.' The point which should be generally known is this, that had it not been for Dr. Black's discoveries, no experiment would probably have been tried by the two Montgolfiers. This I can assert upon the evidence of the older Montgolfier, who was one of the most candid and able men I have met with, and who always mentioned Dr. Black with the respect to which he was so peculiarly entitled."

How we designed the Hy
for pe

Our first step was to remember how a person is designed.

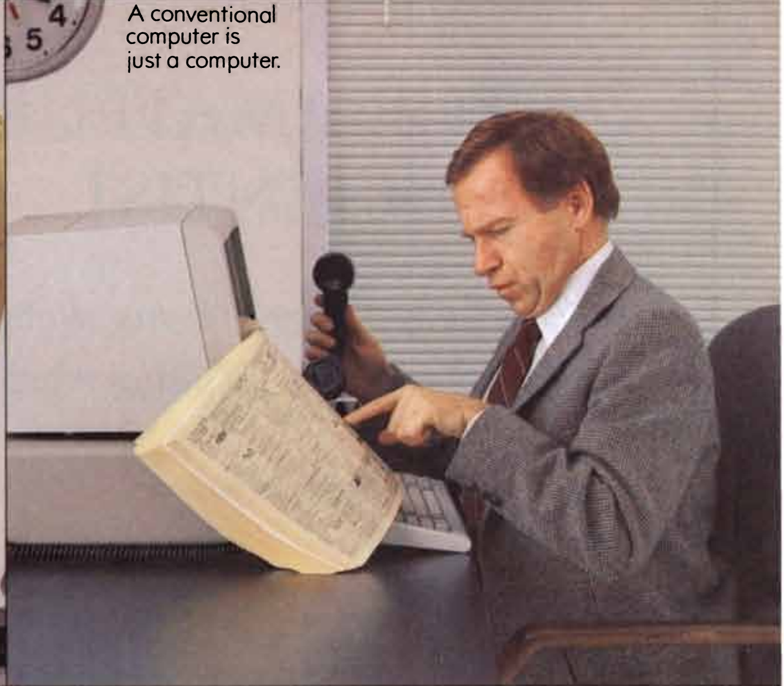
A person isn't designed to carry a nearly 30 lb. object that bangs against the hip and cramps the fingers. So the Hyperion is a manageable 18 lbs. and travels in an easy-to-carry tote bag.

A person isn't designed to work in the shadow of a computer. So the Hyperion keeps a low profile.

A person isn't designed to worry about an expensive computer being dropped by baggage handlers. So the Hyperion fits under an airplane seat. Easily.

A person isn't designed with a phone attached to his ear. So the Hyperion has a telecommunications hook-up that finds numbers, dials, and makes a connection

IBM is a registered trademark of International Business Machines Corp



erion personal computer
sons.

without a person ever picking up the receiver.

And no person, no matter what design, wants to work harder than needed. So the Hyperion has all the power needed, and uses IBM® compatible software — the best and easiest to use software written.

But no matter what we say about the Hyperion, the only way to see the difference is to come in and pick one up. In person.

For the dealer nearest you call 1-800-323-1717, Ext. 246. In Illinois 1-800-942-8881, Ext. 246. In Canada call 1-800-267-1002 or 1-613-226-7255.

Hyperion™

You'll never look at computers the same way again.

Hyperion is a registered trademark of Bytec-Comterm Inc.

THE AMATEUR SCIENTIST

The spectra of streetlights illuminate basic principles of quantum mechanics

by Jearl Walker

Spectroscopy has been the physicists' Rosetta stone for understanding atoms and molecules. It began as a puzzle of colors, deepened into a mystery of energies and wavelengths and finally blossomed through three stages of progression into quantum mechanics. Now applied in almost every field of science and technology, it also affords the amateur scientist a glimpse into atomic structure. Several simple experiments can be done to explore transitions in atoms giving rise to spectra in the visible range. For example, the spectra of sodium and mercury can be observed from laboratory lamps and from two common types of streetlight, mercury vapor and sodium vapor.

A spectrometer has long been a standard but expensive apparatus in introductory physics and chemistry classes. Usually it consists of three elements. A slit and a lens form a plane wave of the light from a source. The light is then dispersed into its component colors by a prism or a diffraction grating. Finally a small telescope enables the observer to view the dispersion in detail. The spectrum can be photographed by attaching a camera to the telescope by means of an adapter.

Simpler and cheaper ways of doing amateur spectroscopy have recently been devised. Leonard H. Greenberg and Thomas Balez of the University of Saskatchewan have described experiments in which an inexpensive diffraction grating was attached to the front of the lens of an instant-copy camera. It was a replica transmission grating with 15,000 lines per inch. The light sources were laboratory lamps filled with mercury, neon, helium and krypton.

Richard L. Bowman of Kent State University suggested that this simple spectrometer be directed at streetlights in order to study the emission spectra of their gases. He also recommended recording the spectra on black-and-white film so that the negative of a photograph can be mounted in a standard projector. The projection expands the details of the spectrum, allowing a group of students to calibrate the features of the spectra with a meter stick.

Richard Breslow of the University of Connecticut pointed out that the spectrum of a sodium streetlight has a curious feature. A dark band appears in the yellow part of the spectrum where the spectrum of a laboratory sodium lamp is brightest. The dark band is easily seen with an inexpensive diffraction grating and the unaided eye.

Recently David A. Katz of the Community College of Philadelphia sent me his plans for analyzing several sources of light with a 35-millimeter camera or a video camera (which is useful for large classes). He employs two types of light source. One is a standard laboratory lamp with a gas-discharge tube. The other is an apparatus that he builds out of easily obtainable laboratory supplies. Compressed air is blown across a solution of a metal salt and water and into a tilted flask. This procedure fills the flask with airborne metal ions. A hose running from a side arm on the flask to a Meker burner leads the ions into the flame of the burner, where they are excited by collisions with other molecules in the hot environment. As they deexcite they emit their characteristic spectra.

The device that sprays ions into the flask can be made easily by anyone who has had a little experience with glassblowing. It consists of two glass tubes, the inner one between six and eight millimeters in diameter and the outer one between 12 and 14 millimeters. Extending downward from the nested tubes is a capillary tube with a bore of one millimeter. The idea is for the compressed air (from a bottle or a laboratory outlet) to blow through the nested tubes. The narrowing of the escape route exerts a suction on the capillary tube. If the capillary tube is immersed in a solution of a metal salt, the suction draws some of the metal ions up into the stream of air.

Katz suggests making the inner tube first. Flare one end outward uniformly and draw the other end down to a small diameter. Join the capillary tube to the side of the larger of the nested tubes. Insert the inner tube and join its outwardly flared end to one end of the larger tube. The joint must be airtight.

Draw the opposite end of the larger tube down to a small diameter. Make its tip slightly larger and more extended than the tip on the inner tube. With another glass tube form a hose connection that is to be joined to the sprayer. This tube should have a hose nipple to ensure that the rubber hose from the air supply fits snugly. Anneal the glass. Open the bottom end of the capillary tube and grind it flat.

The flask should be between 500 and 1,000 milliliters in volume. Attach to its side arm a rubber hose about 15 centimeters long and one centimeter in diameter. Fasten the rubber hose with tape to one of the air vents in the burner. Seal the other vents with tape so that the only supply of air comes from the hose. Natural gas is supplied to the burner in the normal way. The apparatus thus supplies to the flame in the burner air mixed with a metal ion.

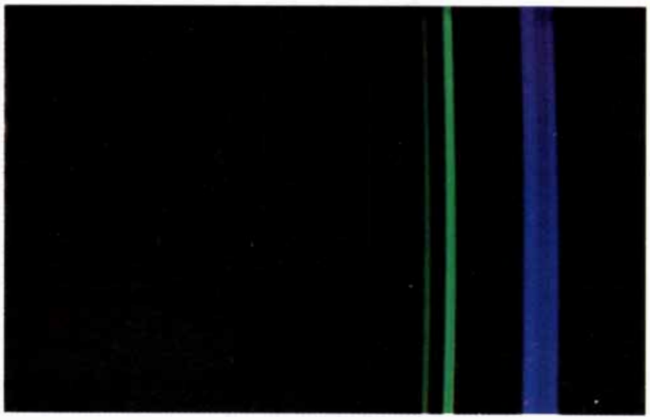
Katz buys 8½-by-11-inch sheets of diffraction grating of the transmission type from the Edmund Scientific Company (101 East Gloucester Pike, Barrington, N.J. 08007). Similar sheets are available from Jerryco, Inc. (601 Linden Place, Evanston, Ill. 60202). Katz glues the sheet to the hood of a camera lens. The hood is either an inexpensive plastic or rubber one or one that he forms out of cardboard. Its outer end is coated with glue and then put on a grating. After the glue dries Katz snips the excess grating from the edge of the hood.

The camera is mounted on a tripod or some other stable stand and positioned one or two feet from a light source. A black background will eliminate reflections from other objects in the room. The camera does not face the source directly but is off-center by about 30 degrees. (If it faced the source, the grating would transmit light to the film without dispersion.) The film is exposed with the first-order diffraction pattern produced by the grating: the light is spread across the film as a function of its wavelength.

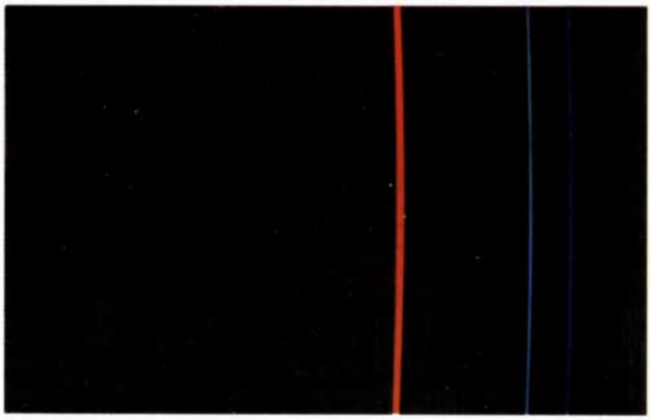
The aperture on the camera should be wide. Only experience can determine the proper exposure time, but it probably will be from 1/4 second to 10 seconds depending on the brightness of the source and the speed of the film. Katz suggests Ektachrome or high-speed Ektachrome. Instead one might replace the 35-mm. camera with a video camera, so that the spectrum can be transmitted to a television screen or a video projector.

I studied several atomic spectra by taping an inexpensive diffraction grating (13,400 lines per inch) in front of a lens mounted on a 35-mm. single-lens reflex camera. My color film had an ASA rating of 400, but you might try the newer ASA 1000 color film. When I stood near a source, the spectral lines in the first-

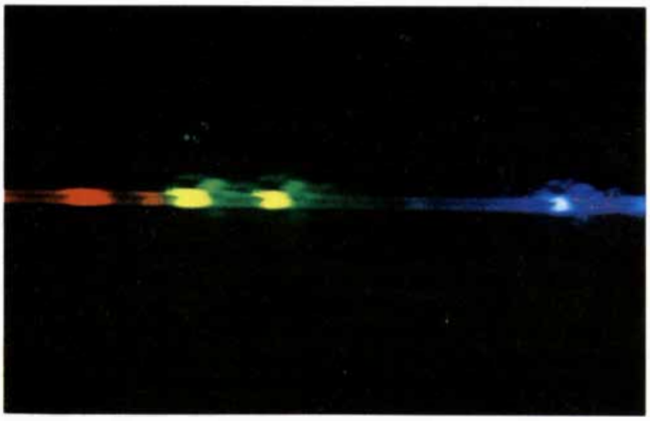
A street scene with several spectra



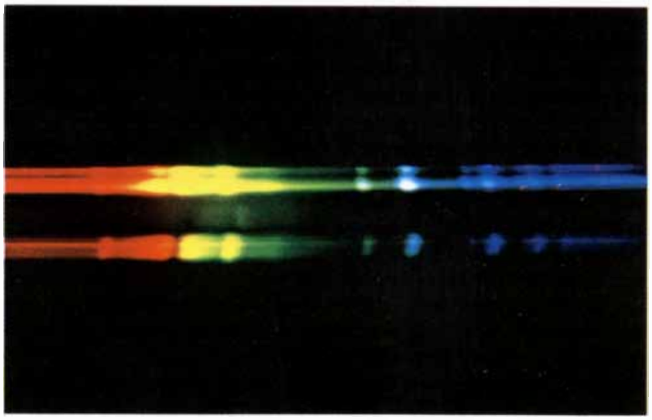
David A. Katz's spectrum of mercury



A hydrogen spectrum by Katz

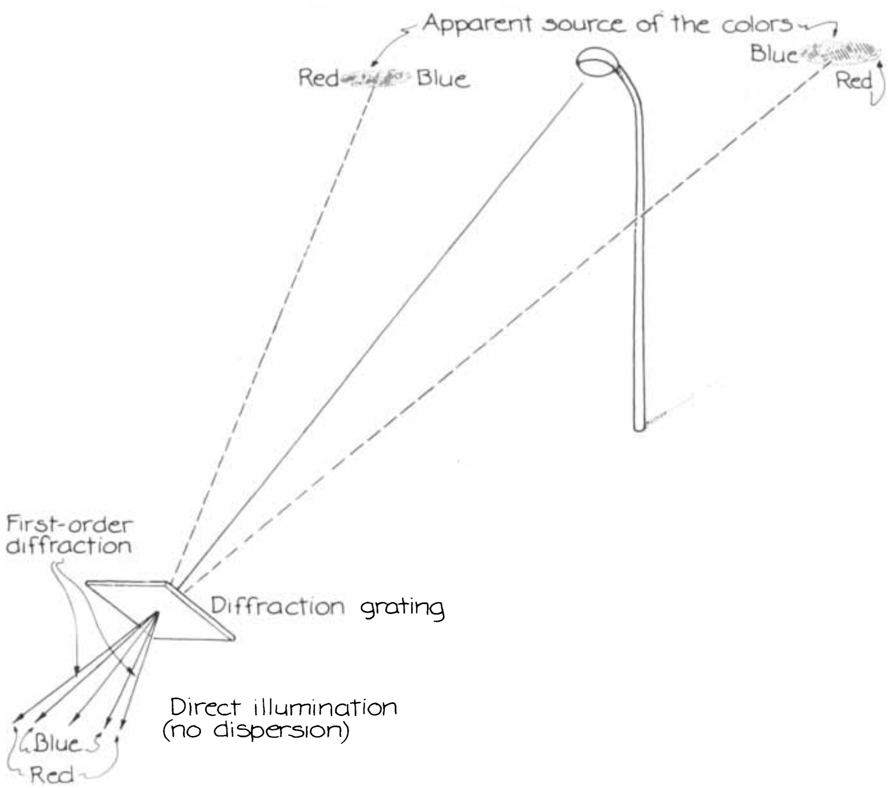


The spectrum of a mercury streetlight

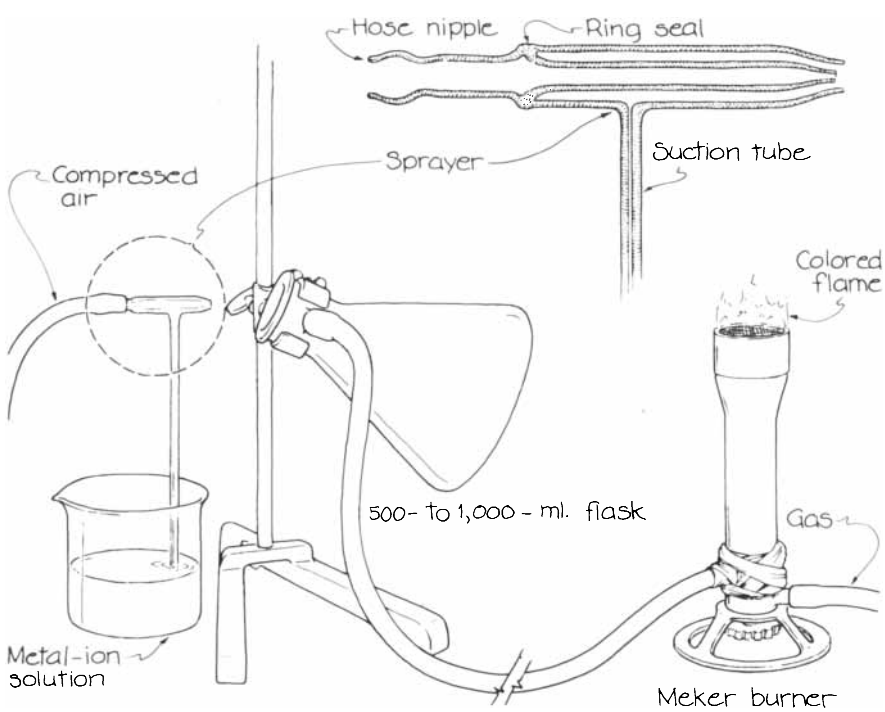


The spectra of several sodium streetlights order diffraction tended to overlap because the source was fairly large in my field of view. I backed away to narrow the lines. The lines were then dimmer. I set the camera at its widest aperture and made photographs at many exposure times between 1/4 second and 25 seconds, hoping to hit on the proper exposure. (An instant-copy camera would have greatly decreased my guesswork.)

I first photographed laboratory lamps of hydrogen, sodium and mercury. Next I photographed a mercury-vapor streetlight. The lamp emitted light across the full visible spectrum, but against that background I was able to see several bright lines. I saw deep blue (violet), faint green, bright green, bright yellow, faint yellow (or dull red) and bright red. (The photographs reproduced on the preceding page do not show the detail in the originals or visible when one looks directly at the light source through the grating.) Most of these emissions are from deexciting mercury atoms. The red emission is from a phosphor in the lamp that enhances its power at the red end of the spectrum.



How the spectrum from a streetlight is dispersed



Katz's setup for metal-ion spectroscopy

A different set of colors was visible in the spectrum from sodium. The brightest emission from the laboratory lamp was in the yellow at a wavelength of about 589 nanometers. The emission from a sodium streetlight was different partly because the lamp of that light contains a small amount of xenon and mercury, which contribute to the spectrum. Another striking difference is the dark band noticed by Breslow in the midst of the yellow.

The difference in the yellow light from the two types of lamp comes from their gas pressure. The pressure in the laboratory lamp is about .02 atmosphere, that in the streetlight about 1.5 atmospheres. In both cases sodium atoms are excited by collisions and emit photons of bright yellow light with a wavelength of 589 nanometers. At the lower pressure the chances are good that a photon can escape from the lamp without being absorbed by another sodium atom in its way. At the higher pressure the photon is likely to be absorbed by another atom and so prevented from escaping. The phenomenon is known as self-absorption.

Yellow light is nonetheless visible in the part of the spectrum surrounding the dark band because of the frequent and violent collisions among the atoms. A collision can briefly alter the energy of the electrons in an atom. If during that period the atom deexcites, the light has a wavelength slightly longer or shorter than the normal yellow light. Such light cannot be absorbed by another sodium atom. Thus this yellow light escapes from the lamp without self-absorption.

When the sodium spectrum from a laboratory lamp is studied in a good spectrometer, the yellow emission is discovered to be at two wavelengths: 589 nanometers and 589.6 nanometers. This doublet is curious because it is evidence of a subtle play of quantum numbers for an electron in the sodium atom.

Collecting spectra without understanding their origins is like collecting butterflies without studying their biology. The origins of atomic spectra occupied the thoughts of some of the best physicists for two centuries. Understanding came slowly, often flying in the face of common sense. By the middle of the 19th century students of spectra had noticed a peculiarity of the light emitted when pure elements were burned. Whereas light from a burning coal in a fireplace shows the full visible spectrum, a pure element produces only certain colors. For example, burning sodium yields a distinct yellow light. All the elements investigated had their own unique set of spectral lines, a phenomenon that puzzled the early investigators.

In 1885 Johann Jakob Balmer stumbled on a concise formula that fitted the lines in the visible spectrum of hydrogen. The formula related the frequency (or the wavelength) of the lines to certain sets of integers. One integer was the number 2. The other was an integer larger than 2. Neither Balmer nor anyone else had the slightest idea why the formula worked.

In 1911 Ernest Rutherford demonstrated that an atom consists of a highly dense, positively charged core surrounded at some distance by negatively charged particles. It is now known that the core—the atomic nucleus—consists of protons and neutrons and that the surrounding particles are electrons. Light is emitted when an electron loses energy by orbiting closer to the nucleus, but why are only certain spectral lines seen? According to the understanding of electromagnetism in Rutherford's time, an electron losing energy would have to fall into the nucleus almost immediately. If that were the case, all atoms would collapse and the world as we know it could not exist.

The first break in solving the puzzle came in 1913, when Niels Bohr was considering how to apply the new idea of the "quantum of action" to the hydrogen atom. Max Planck had earlier set the stage with the introduction of what is now called Planck's constant; the purpose of the constant was to model the atomic emitters of light on the surface of a hot object such as a burning coal. The application of his model to the Rutherford atom was challenging but possible. An explanation of the spectra of the atoms, however, was still thought to be out of reach.

One day a friend suggested that Bohr examine Balmer's formula. Bohr said later, "As soon as I saw Balmer's formula the whole thing was immediately clear to me." The formula employed integers, suggesting that some feature of the electron in the hydrogen atom was restricted to certain values involving those integers. The electron could not merely fall into the nucleus, because that would violate the restriction. Bohr found that if he quantized the angular momentum of an electron in orbit around the nucleus, its energy was then limited to certain values set by an integer he called the quantum number, n. If the electron is in orbit close to the nucleus, it has one of the low energy values in the allowed set. A larger orbit demands a greater value of n. The electron can jump between the allowed energy values, thereby changing the size of its orbit, but apparently it cannot be at an intermediate value of energy or radius of orbit. Moreover, a downward jump into the nucleus itself is not allowed. Since there is a definite lower limit to the energy the electron has, there is also a limit to how small the orbit can be.

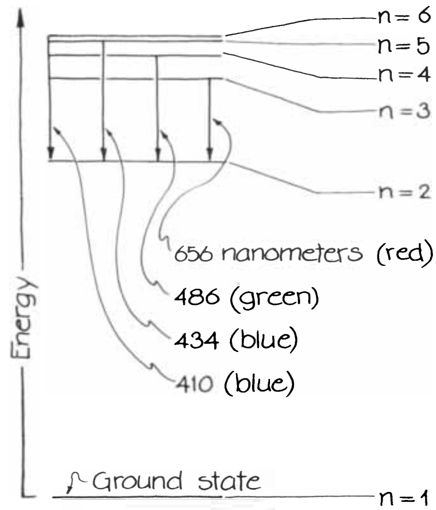
The electron can jump from a lower energy value to a greater one only by absorbing energy from outside the atom. That energy can be supplied by a collision with another particle or atom. It can also be supplied by the absorption of light. After being excited the electron soon deexcites by releasing energy. Either the atom has to run into something to rid itself of its energy or it has to emit light. When a hydrogen atom loses energy by emitting light, the emission forms the spectrum seen in the light produced by a hydrogen lamp.

The top illustration at the right gives the allowed energies for the hydrogen atom according to the Bohr model. The lowest level, which is termed the ground state, has a quantum number n of 1. The other levels are labeled with larger values of n. The Balmer formula (in its original form) covered the second level and higher ones. For example, if a hydrogen atom is excited to the third level by a collision, it soon deexcites to either of the two lower levels. Suppose it deexcites to the second level. Then it must rid itself of just the right amount of energy to make the jump.

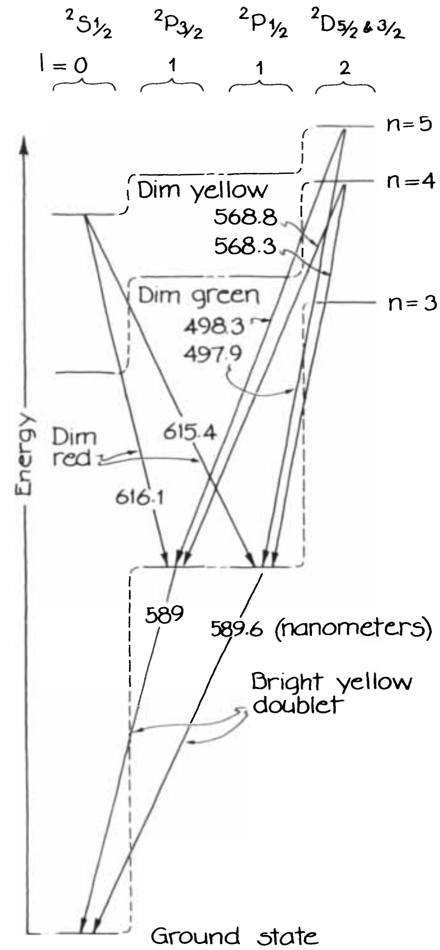
Prior to Bohr's work Albert Einstein expressed the fundamental relation between the energy and the frequency of light: the energy is equal to the product of the frequency and Planck's constant. For example, when the hydrogen deexcites from the third level to the second, it must emit light of a certain frequency in order to lose the right amount of energy. Since the frequency of light is inversely proportional to its wavelength, the wavelength is also fixed. The result of such a deexcitation appears as a red line in the spectrum at a wavelength of 656 nanometers.

In a lamp the atoms are constantly being excited by collisions to all the levels shown in the chart. The many downward transitions mean that different amounts of energy are released as light. Once the hydrogen atoms are somehow excited several different wavelengths of light emerge. The interesting point is that only certain values can be expected. In the visible part of the spectrum there are three emissions besides the red one, each associated with a jump downward to the second level. A green wavelength comes from the fourth level. A blue one comes from the fifth. A blue one of shorter wavelength comes from the sixth. All four emissions fall in the visible spectrum. Other downward jumps result in invisible emissions in the ultraviolet or infrared parts of the spectrum.

Bohr's model of the hydrogen atom was revolutionary in that it introduced the quantizing of motion and energy to the atom, but it was wrong. Indeed, the correspondence between its predictions for the spectrum of hydrogen and the observed spectrum was almost accidental. Bohr's model is based on the as-



An energy chart for hydrogen

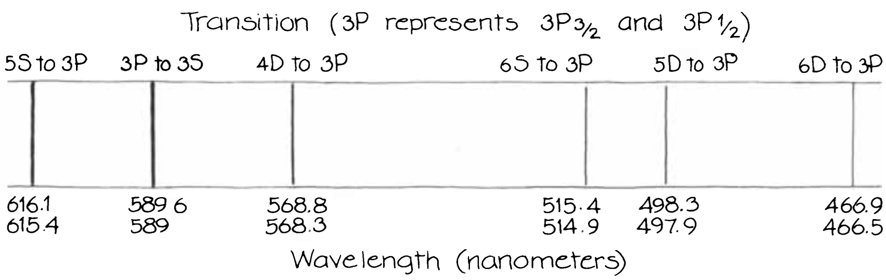


A sodium energy chart

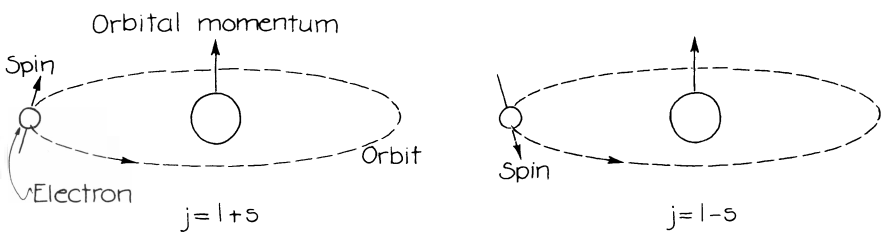
sumption that the electron goes around the nucleus in a circle, but there is no reason other types of orbit cannot exist.

Further work by Arnold Sommerfeld dealt with three-dimensional orbits. Although the model for the hydrogen atom then appeared to be fairly complete, extensions of the model to other types of atom failed. The model could not explain even the simple spectra arising from laboratory lamps and streetlights. Something was obviously wrong.

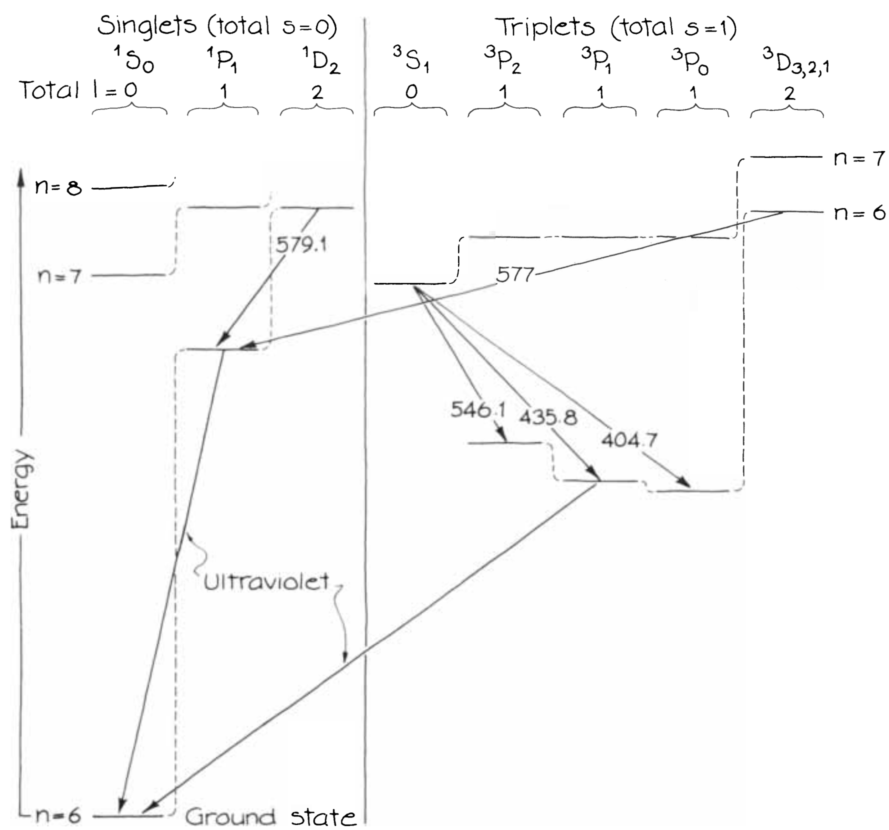
A true quantum-mechanical model of the hydrogen atom emerged in 1926 when Erwin Schrödinger wrote the wave equation that now bears his name. The equation was essentially a repetition of the formula pre-quantum-mechanical physicists had employed for waves of all kinds. Schrödinger's equation seemed strange because instead of a wave in water or some other medium it described a wave in what is called a scalar field. The equation was marvelously successful in fitting the spectra of atoms, but the physical meaning of the scalar field was so elusive that many physicists thought the theory was unnatural.



The sodium spectrum



Combinations of quantum numbers



An energy chart for mercury

New numbers popped out of the equation in orderly fashion. An electron in an atom has a principal quantum number n (somewhat similar to the n in the Bohr theory), a second one l that restricts the size of the angular momentum of the electron's orbit around the nucleus and a third one that restricts the orientation of that angular momentum. The first two help to make sense of the spectra from laboratory lamps and streetlights.

Once the value for n is determined, the value for l can be 0 or any integer up to $n - 1$. Now it is necessary to turn to another parameter that restricts an intrinsic angular momentum of electrons. Every electron behaves as though it were a miniature top spinning on an axis. This type of angular momentum, which is called spin, is given by the quantum number s.

I want to caution that these quantum numbers, pictures, energy charts and so on are all merely models associated with an atom. They are not reality of the normal kind. You will never directly observe an electron or sense the quantizing of its orbits and energies in an atom. Indeed, an atom is not simply a miniature solar system and spin is not simply the angular momentum of a tiny top. In the current view the electron is an infinitesimal point to which certain characteristics, among them charge and spin, are attributed.

The spectrum of sodium can serve to demonstrate how the quantum numbers figure in atomic spectra. A neutral sodium atom has 11 electrons. Ten of them play no immediate role in an understanding of the spectrum because they form what are called closed shells. The concept of a closed shell is based on the quantum-mechanical argument that no two electrons in an atom can have precisely the same set of quantum numbers.

The 11th electron is called a valence electron because it lies outside the closed shells. When a sodium atom is unexcited, the valence electron has an n of 3 and an l of 0. There are two more possible values for l (1 and 2), but for the electron to have them it must be given some energy. Instead it might be given enough energy to jump to a state with n of 4 or more. These possibilities for sodium are charted according to energy in the bottom illustration on the preceding page.

The ground state (shown at the lower left in the chart) can be labeled as $3^2S_{1/2}$. S is the traditional way of representing an l value of 0 for the valence electron. The 3 is the value of n. The superscript 2 comes from an equation involving the

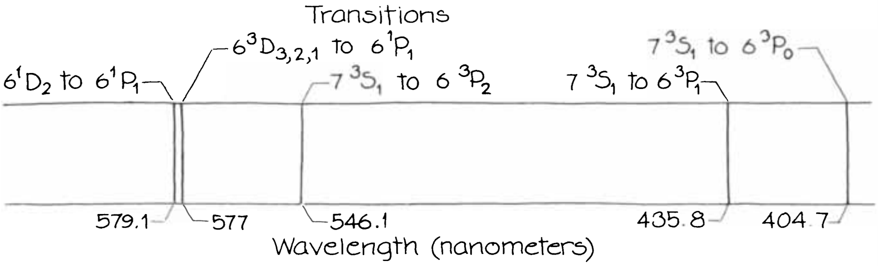
spin. The subscript brings in a new quantum number j, a combination of l and s. This new number is important because it restricts the total angular momentum of the electron, that is, the combination of its orbital angular momentum and its spin angular momentum. The values of j are computed by adding and subtracting the values of l and s. The largest j value equals the sum of l and s. The lowest value is the subtraction of l and s (ignoring any negative sign that might result). The intermediate values of j are found by subtracting 1 from the largest value until the smallest value is reached. For the ground state the calculation is simple. There l is 0 and s is 1/2. Hence j can only be 1/2, which appears as the subscript on the column label.

A sodium atom is more interesting if the valence electron is given some energy, so that (while it is still in the state with n of 3) it has an l of 1, leaving two possible values for j. If l and s are added, j is 3/2. If they are subtracted, j is 1/2. Physically the two situations involve the orientations of the two components of the electron's angular momentum. If the orbital and spin components are in the same direction, j is 3/2. If they are in opposite directions, j is 1/2.

The only reason for introducing the quantum number j is that the electron's energy (and thus the atom's) depends on it. The two orientations of the angular momentum (the orbital and spin components) require different energies for the electron. Less energy is required when they are in opposite directions. Slightly more is needed if they are in the same direction. The way to keep track of the situation is to consider the j values.

The two states appear in the energy chart as the lowest levels under the labels $P_{3/2}$ and $P_{1/2}$. As before, the subscript gives the value of j. The P is the traditional way of representing a value of 1 for l. These two states differ so slightly in energy that they appear to be at the same level in the energy chart. Still, the difference in energy is responsible for a closely spaced doublet of yellow lines in the sodium spectrum.

Suppose a sodium atom in a lamp is excited to the $P_{3/2}$ state. It quickly drops back to the ground state, emitting light to get rid of the appropriate amount of energy. The wavelength of that light is 589 nanometers, which is in the yellow part of the spectrum. If instead the atom is excited to the $P_{1/2}$ state, it deexcites to the ground state by emitting a slightly smaller amount of energy, because $P_{1/2}$ is a bit lower in energy. This time the wavelength of the light is 589.6 nanometers, which is the second line of the sodium doublet in the yellow. Thus the presence of the yellow doublet in the sodium spectrum demands the invention of the quantum number j. Otherwise the slightly different yellow wavelengths make no sense.



The spectrum of mercury

Several other excited states attained by atoms in the lamp are shown in the chart. Soon after excitation the atoms drop back in various ways to reach the ground state. As long as the lamp is lighted the atoms are reexcited. The result is a steady emission of light at the wavelengths arising from all the possible deexcitations. I have noted only a few of the possibilities in the energy chart and the drawing of the sodium spectrum.

For example, a red emission line results when the valence electron is first excited to the level with n of 5 and l of 0. For a reason I shall not go into the electron cannot merely drop back directly to the ground state. In the chart the deexcitation route must be both downward and to another column of levels. Hence the electron can deexcite by dropping to either the $P_{3/2}$ level or the $P_{1/2}$. In one case the electron emits a wavelength of 616.1 nanometers, in the other case one of 615.4 nanometers. The difference in these two wavelengths of red light is due entirely to the difference in energies of the two P states. Thus the two possible values of j for the P states end up creating another closely spaced doublet of spectral lines.

An analysis of the spectrum of mercury is similar to that of the spectrum of sodium. An electrically neutral mercury atom has five complete shells and two valence electrons. Although figuring out the quantum numbers for two electrons is more work than it is for one electron, the procedures are about the same. The energy levels for mercury are shown in the bottom illustration on page 142. As before, the state for an unexcited atom is shown at the bottom left. Some of the possible excited states are shown higher in the chart. Mercury atoms in a lamp are constantly being excited to those upper levels and then finding their way back to ground state by emitting light.

The major difference in analyzing the energy levels for mercury is that it is necessary to focus on the total l and the total s for the valence electrons. The totals are combined to find the total j, which determines the energy. To aid in such calculations the energy chart is divided into two sections labeled singlets and triplets. Under the singlet label are the energy levels for cases where the two electrons have their spins in opposite directions and so have a total s of 0. The total j for any of these levels is simply the value of the total l. For example, the lowest level under the label P in this section is for the case where one electron has an l of 0 and the other has an l of 1. The total l is 1. Since the total s is 0, the total j is also 1. In the triplet section the left side under the label S is also easy. Here the total l is 0 and the electrons have their spins in the same direction to give a total s of 1. Hence the total j can only be 1.

The fun begins in the other part of the triplet section. I shall concentrate on the lowest levels under the label P in that section. There the total l is 1 but the total j can have three possible values, each associated with a different energy. The sum of total l and total s gives a value for total j of 2. A subtraction gives a total j of 0. An intermediate value for total j is 1. The energies associated with those three values are shown in the energy chart.

Three of the wavelengths in the mercury spectrum depend on these three energy levels. Consider a mercury atom excited by a collision to the lowest S_1 level. It can deexcite by dropping to a P level in the same section of the chart. (As with the sodium chart the deexcitation route must be downward and into another column.) Since there are three choices of P levels, the atom can emit light at either 546.1, 435.8 or 404.7 nanometers. With a great many atoms being excited in the lamp some are always dropping out of the S_1 along the three routes, giving continuous emissions at the three wavelengths. Thus the presence of this triplet of wavelengths in the mercury spectrum can be understood only with the quantum number j.

This is only a start on exploring the spectra of ordinary lamps. You might enjoy working out other features of the sodium and mercury spectrum. Then you might try to analyze the spectrum of a fluorescent lamp, which has a thin coating of phosphor on its inner surface. There are also lamps in which the light is emitted by molecules rather than atoms. It is only fair to state, however, that unraveling molecular spectra can keep you occupied for quite a long time.

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

COMPUTER RECREATIONS

THE $3x + 1$ PROBLEM AND ITS GENERALIZATIONS. J. C. Lagarias. To appear in *The American Mathematical Monthly.*

LAUNCH UNDER ATTACK

THE ELECTROMAGNETIC PULSE AND ITS EFFECTS. Samuel Glasstone and Philip J. Dolan in *The Effects of Nuclear Weapons.* U.S. Department of Defense and U.S. Department of Energy, 1977.

LAUNCH UNDER ATTACK TO REDRESS MINUTEMAN VULNERABILITY? Richard Garwin in *International Security,* Vol. 4, No. 3, pages 117–139; Winter, 1979–80.

NUCLEAR DECAPITATION. John D. Steinbruner in *Foreign Policy,* No. 45, pages 16–28; Winter, 1981–82.

A COMPUTER SIMULATION OF LAUNCH PROCEDURES FOR A SQUADRON OF MINUTEMAN II ICBMs. Paul Morawski in *Simulation,* Vol. 41, No. 1, pages 16–24; July, 1983.

THE ATMOSPHERIC EFFECTS OF EL CHICHÓN

THE CLIMATIC SIGNIFICANCE OF EL CHICHÓN. C. B. Sear and P. M. Kelly in *Climate Monitor,* Vol. 11, pages 134–139; 1982.

STRATOSPHERIC SULFURIC ACID FRACTION AND MASS ESTIMATE FOR THE 1982 VOLCANIC ERUPTION OF EL CHICHON. D. J. Hofmann and J. M. Rosen in *Geophysical Research Letters,* Vol. 10, No. 4, pages 313–316; April, 1983.

STRATOSPHERIC AEROSOL MASS AND LATITUDINAL DISTRIBUTION OF THE EL CHICHON ERUPTION CLOUD FOR OCTOBER 1982. M. P. McCormick and T. J. Swissler in *Geophysical Research Letters,* Vol. 10, No. 9, pages 877–880; September, 1983.

CLIMATIC EFFECTS OF THE ERUPTION OF EL CHICHON. Special issue of *Geophysical Research Letters,* Vol. 10, No. 11, pages 989–1060; November, 1983.

HIGH-ENERGY COLLISIONS BETWEEN ATOMIC NUCLEI

THE FIRST THREE MINUTES: A MODERN VIEW OF THE ORIGIN OF THE UNIVERSE. Steven Weinberg. Basic Books, Inc., 1977.

NUCLEUS-NUCLEUS COLLISIONS: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON NUCLEUS-NUCLEUS COLLISIONS, MICHIGAN STATE UNIVERSITY, SEPTEMBER 26–OCTOBER 1, 1982. Edited by G. F. Bertsch, C. K. Gelbke and D. K. Scott. North-Holland Publishing Co., 1983.

ANOMALONS AS PINUETS BOUND TO NUCLEAR FRAGMENTS: A POSSIBLE EXPLANATION. Wm. C. McHarris and J. O. Rasmussen in *Physics Letters,* Vol. 120B, No. 1, 2, 3, pages 49–53; January 6, 1983.

RICE

TROPICAL CROPS: MONOCOTYLEDONS 1. J. W. Purseglove. John Wiley & Sons, Inc., 1972.

AFRICAN CEREALS: *ELEUSINE,* FONIO, BLACK FONIO, TEFF, *BRACHIARIA, PASPALLUM, PENNISETUM* AND AFRICAN RICE. R. Porteres in *Origins of African Plant Domestication,* edited by Jack R. Harlan, Jan M. J. de Wet and Ann B. L. Stemler. Mouton Publishers, 1976.

THE ORIGIN, EVOLUTION, CULTIVATION, DISSEMINATION, AND DIVERSIFICATION OF ASIAN AND AFRICAN RICES. T. T. Chang in *Euphytica,* Vol. 25, pages 425–441; 1976.

PRINCIPLES AND PRACTICES OF RICE PRODUCTION. Surajit K. De Datta. John Wiley & Sons, Inc., 1981.

UNISEXUAL LIZARDS

MEIOTIC MECHANISMS OF NATURALLY OCCURRING UNISEXUAL VERTEBRATES. Thomas Uzzell in *American Naturalist,* Vol. 104, No. 938, pages 433–445; 1970.

EVOLUTION OF PARTHENOGENETIC SPECIES OF REPTILES. Charles J. Cole in *Symposium on Intersexuality in the Animal Kingdom,* edited by Rudolf Reinboth. Springer-Verlag, 1975.

CHROMOSOME INHERITANCE IN PARTHENOGENETIC LIZARDS AND EVOLUTION OF ALLOPOLYPLOIDY IN REPTILES. Charles J. Cole in *Journal of Heredity,* Vol. 70, No. 2, pages 95–102; 1979.

PARTHENOGENETIC REPRODUCTION IN LIZARDS: HISTOLOGICAL EVIDENCE. Laurence M. Hardy and Charles J. Cole in *Journal of Morphology,* Vol. 170, No. 2, pages 215–237; 1981.

THE CONTROL OF RIBOSOME SYNTHESIS

RIBOSOMES. Masayasu Nomura in *Scientific American,* Vol. 221, No. 4, pages 28–35; October, 1969.

REGULATION OF BACTERIAL GROWTH, RNA, AND PROTEIN SYNTHESIS. D. P. Nierlich in *Annual Review of Microbiology,* Vol. 32, pages 393–432; 1978.

REGULATION OF RIBOSOMAL PROTEIN SYNTHESIS IN *ESCHERICHIA COLI* BY SELECTIVE mRNA INACTIVATION. Ann M. Fallon, C. Sue Jinks, Geneva D. Strycharz and Masayasu Nomura in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 76, No. 7, pages 3411–3415; July, 1979.

IN VITRO EXPRESSION OF *ESCHERICHIA COLI* RIBOSOMAL PROTEIN GENES: AUTOGENOUS INHIBITION OF TRANSLATION. John L. Yates, Ann E. Arfsten and Masayasu Nomura in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 77, No. 4, pages 1837–1841; April, 1980.

REGULATION OF RIBOSOME BIOSYNTHESIS IN *ESCHERICHIA COLI.* Masayasu Nomura, Sue Jinks-Robertson and Akiko Miura in *Interaction of Translational and Transcriptional Controls in the Regulation of Gene Expression,* edited by M. Grunberg-Manago and B. Safer. Elsevier Science Publishing Company, Inc., 1983.

THE PACKING OF SPHERES

PACKING AND COVERING. C. A. Rogers. Cambridge University Press, 1964.

PRINCIPLES OF COMMUNICATION ENGINEERING. John M. Wozencraft and Irwin Mark Jacobs. John Wiley & Sons, Inc., 1965.

TWELVE GEOMETRIC ESSAYS. H. S. M. Coxeter. Southern Illinois University Press, 1968.

LAGERUNGEN IN DER EBENE, AUF DER KUGEL UND IM RAUM. L. Fejes Tóth. Springer-Verlag, 1972.

THE LEECH LATTICE, SPHERE PACKINGS, AND RELATED TOPICS. J. H. Conway and N. J. A. Sloane. Springer-Verlag, 1984.

THE INVENTION OF THE BALLOON AND THE BIRTH OF MODERN CHEMISTRY

THE CHEMICAL REVOLUTION: A CONTRIBUTION TO SOCIAL TECHNOLOGY. Archibald Clow and Nan L. Clow. Batchworth Press, London, 1952.

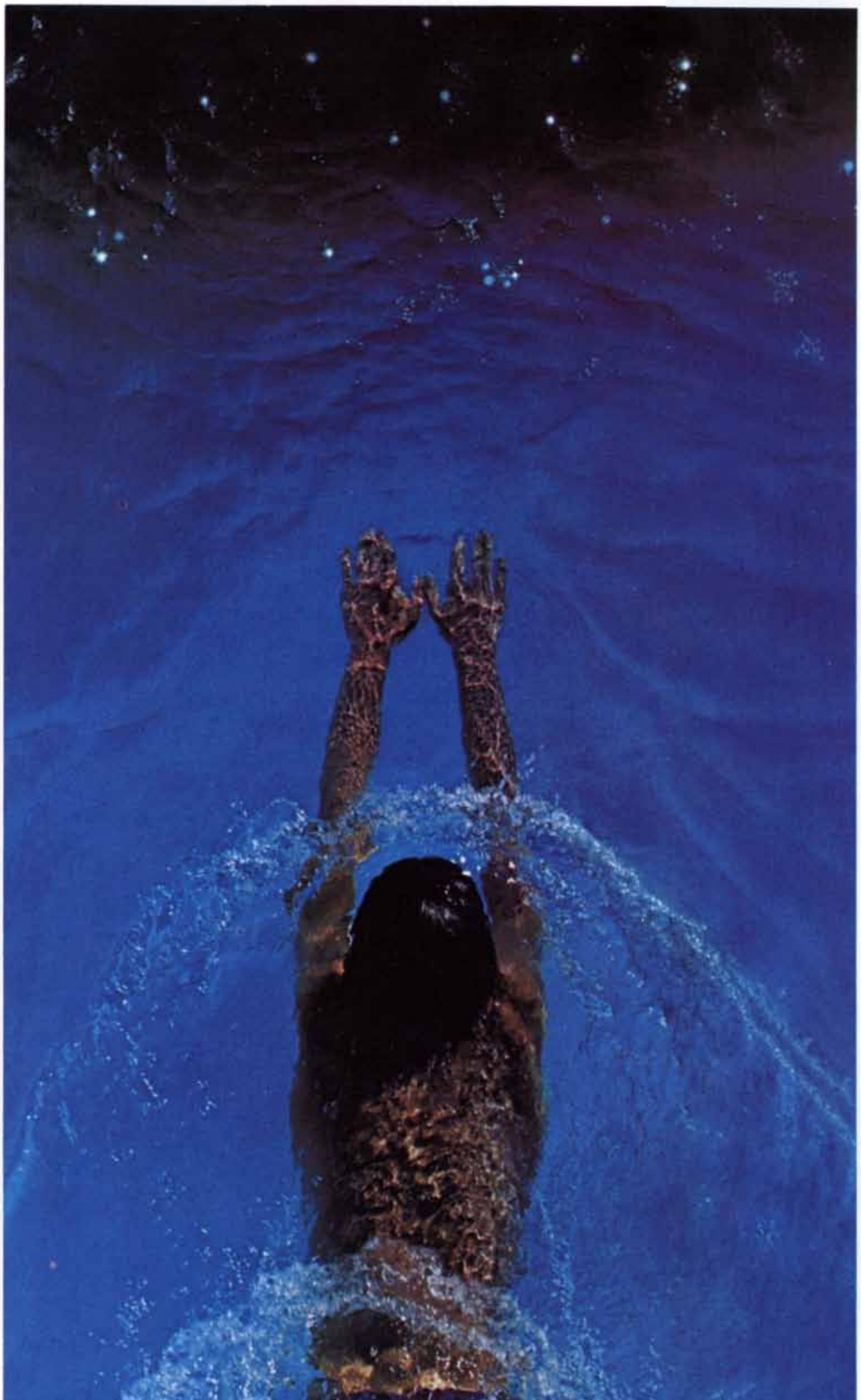
THE HISTORICAL BACKGROUND OF CHEMISTRY. Henry M. Leicester. Dover Publications, Inc., 1971.

A BOOK REVIEW OF AN EARLY (1781) SURVEY OF PROGRESS IN CHEMISTRY. Arthur F. Scott in *Survey of Progress in Chemistry,* Vol. 8, pages 253–277; 1977.

THE AMATEUR SCIENTIST

ATOMIC SPECTRA. H. G. Kuhn. Academic Press, 1963.

LIGHT AND COLOR IN NATURE AND ART. Samuel J. Williamson and Herman Z. Cummins. John Wiley & Sons, Inc., 1983.



Stretch.

It's the only way to see how far there really is. For thousands of young people, one place to stretch is in local swim clubs and the novice programs of United States Swimming. They get coaching, fun, encouragement, a little taste of competitive swimming and a big lesson that we can all afford to remember. They're learning that nothing is too new to try. And that's a lesson that reaches a long way past the edge of the pool.

PHILLIPS 66

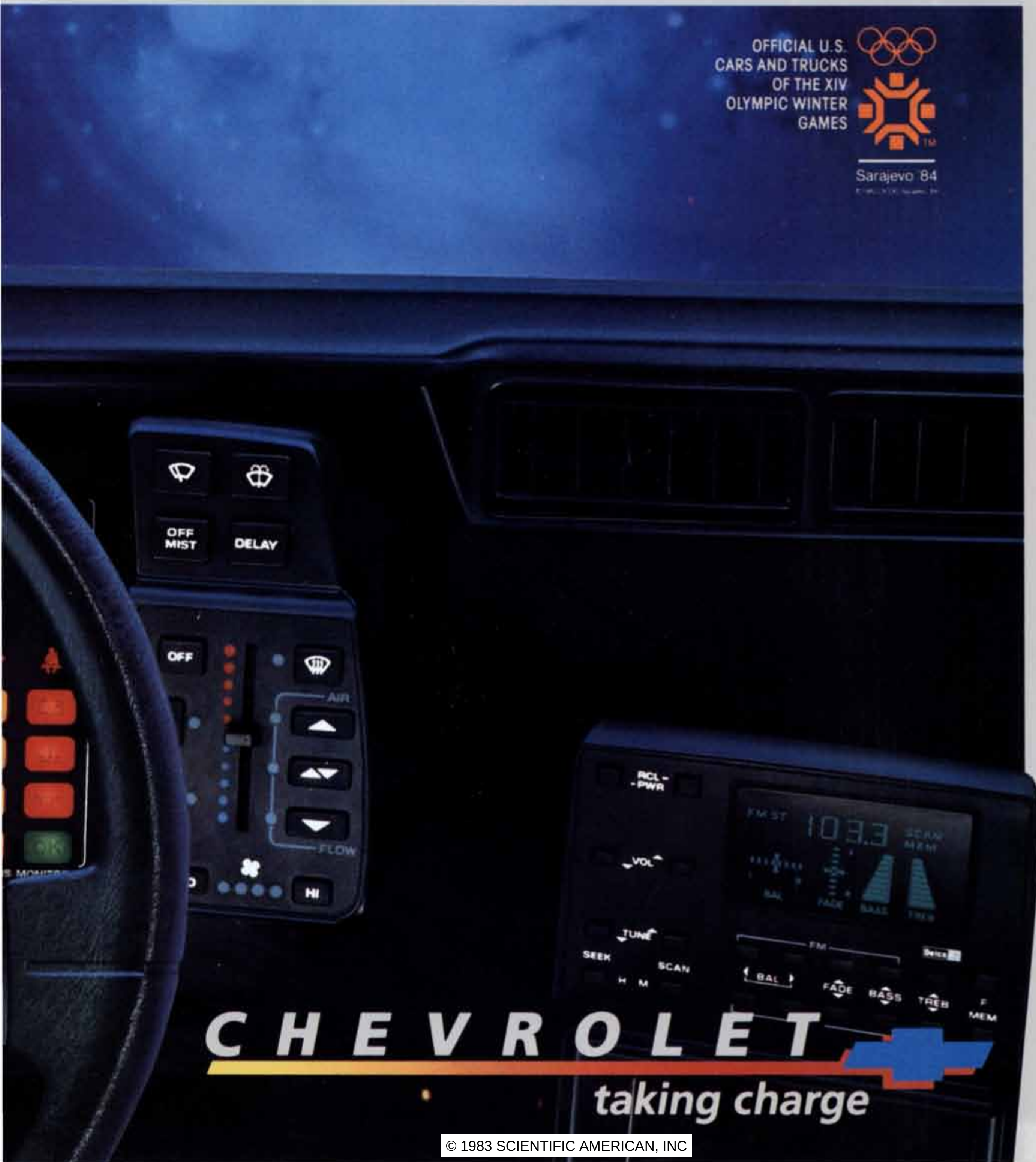
Phillips Petroleum has sponsored United States Swimming for eleven years, and United States Diving for five, helping with operating costs of national and international competitions leading to world championships. We continue to support them. Because their ideals touch the lives of everyone.

Today's Chevrolet

Welcome aboard

Camaro Berlinetta. Above, the night sky is black, punctuated by the flickering of stars and planets. Below, the white lines of the roadway snake deep into the unknown. Ahead, the blue-green glow of your instrument cluster advises you all systems are go on board Starship Camaro. This is not science fiction. This is Berlinetta.

Starship control central. Climb in. Buckle up. Adjust the retractable push-button instrument pods so your hands never have to leave the wheel in controlling vital functions. Turn the key and watch the system monitor perform seven preflight tests as the engine sparks to life. Blip the throttle and watch the vacuum-fluorescent tach dance to



OFFICIAL U.S. CARS AND TRUCKS OF THE XIV OLYMPIC WINTER GAMES
Sarajevo '84
OFF MIST
DELAY
OFF
AIR
FLOW
HI
RCL - PWR
VOL
TUNE
SEEK
SCAN
BAL
FADE
BASS
TREB
MEM
CHEVROLET
taking charge


Step ahead of your time to Panasonic Color TV with CompuFocus.™ Then try to tell the difference between the real scene and your screen.
Panasonic technology narrows the gap between image and reality. With CompuFocus. An ingenious optical system with image-purifying electronic filters. And revolutionary lens gun design. To give you a TV picture so incredibly lifelike you can hardly tell where image ends and reality begins.
And CompuFocus is just the beginning of exciting behind-the-screen Panasonic technology. The slender 19" (meas diag) CT-9053 (shown) has a cable-compatible 139 channel tuner. And a wireless remote that allows direct access to any channel. Or instant switching back and forth, between any two channels with the touch of a button.
You'll find the unsurpassed quality of CompuFocus optics in 13 different Panasonic Color TVs. Each one, a step ahead of its time. TV picture simulated.
Panasonic
just slightly ahead of our time.