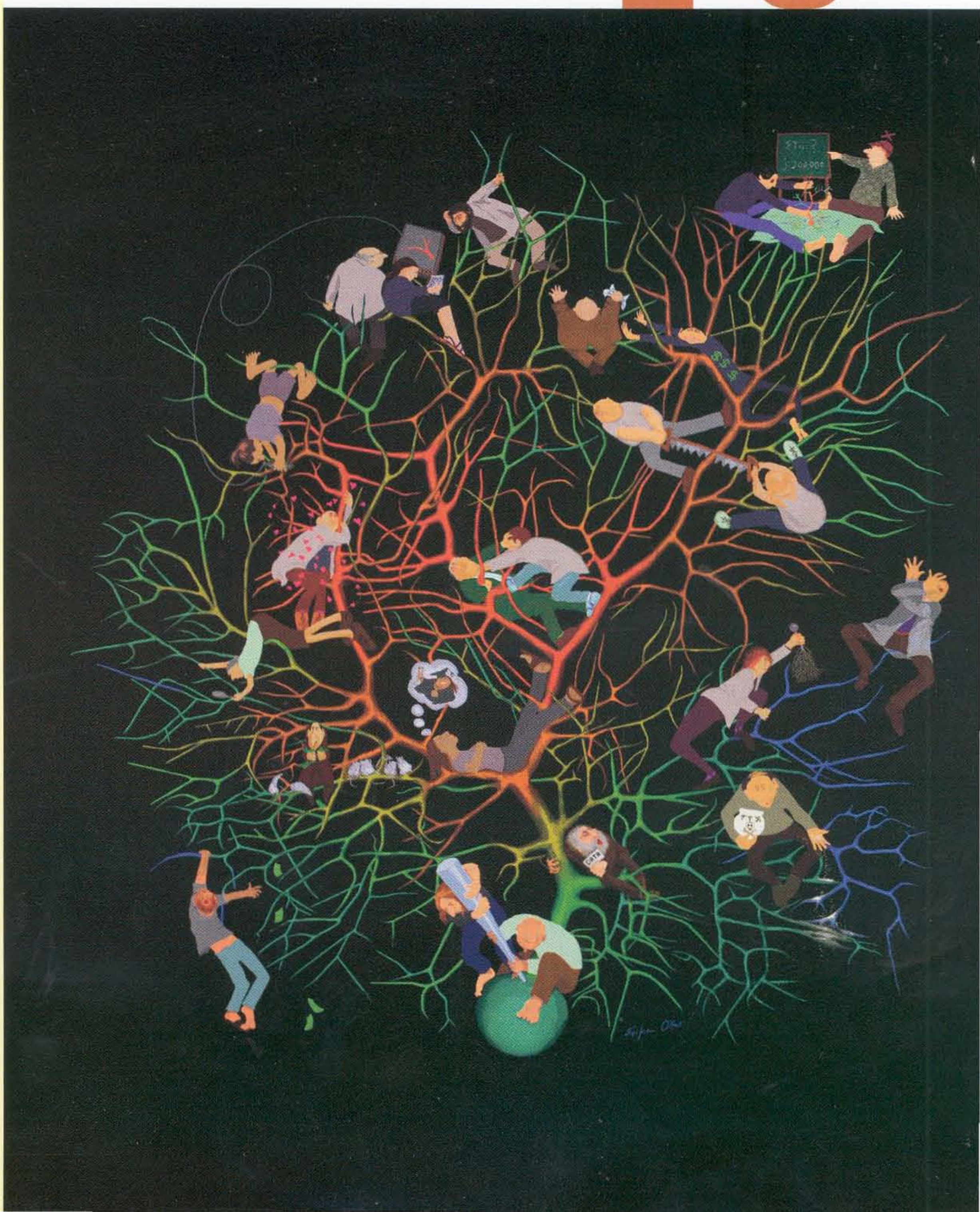
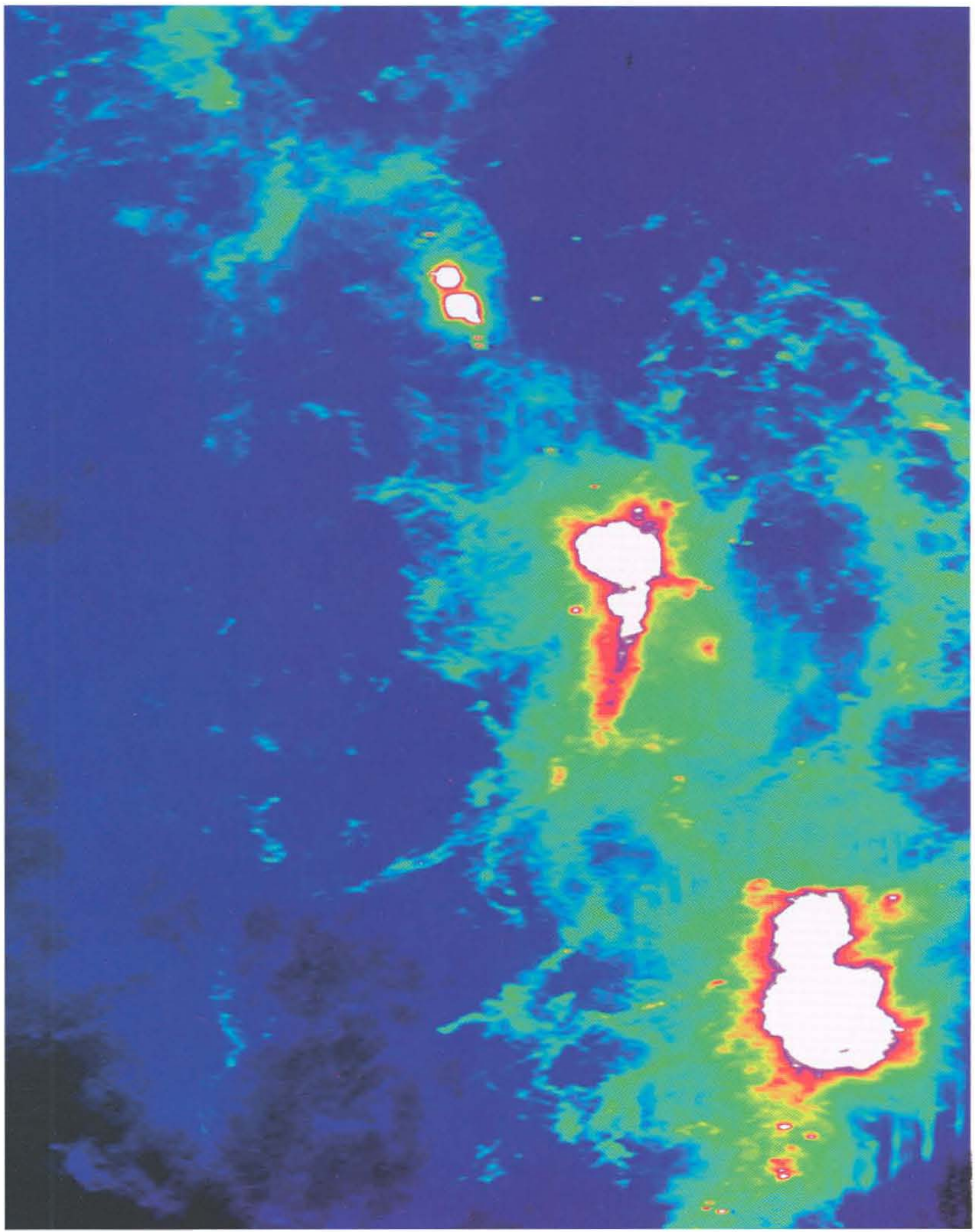
California Institute
of Technology

E&S
ENGINEERING AND SCIENCE

Volume LIX, *Number* 4,
1996

IN THIS ISSUE



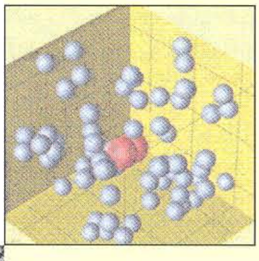


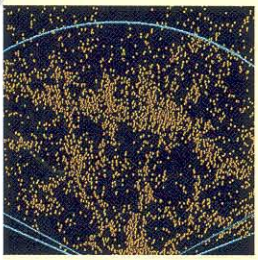
This high-resolution infrared image of the Orion nebula was made from Infrared Astronomical Satellite (IRAS) survey data by grad student Yu Cao on the Intel Paragon, Caltech's newest concurrent supercomputer. IRAS surveyed 96 percent of the sky in 1983, the most complete infrared sky survey ever made. Many data sets at other wavelengths and higher resolutions have become available since then, so Cao used the Paragon to wring more detail out of the IRAS data and create the IRAS Galaxy Atlas. This collection of high-resolution (approximately one arc minute) images has comparable resolution to the other data sets, allowing them to be compared directly.

ENGINEERING AND SCIENCE

California Institute of Technology

Volume LIX, Number 4, 1996





On the cover: This whimsical rendition of a Purkinje cell shows some of what's going on in computational neuroscience. At the bottom, for example, a microelectrode is being injected into the cell body; on a branch above a student dreams of graduating; farther up the trunk some experimentalists are sawing off the limb the theorists are sitting on. One thing the picture doesn't show is that supercomputers are changing the way neuroscience is done. For more on supercomputing in many fields of science, see the story on page 8.

Engineering & Science (ISSN 0013-7812) is published quarterly at the California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125. Annual subscription $10.00 domestic; $20.00 foreign air mail; single copies $3.00. Third class postage paid at Pasadena, California. Published by the California Institute of Technology and the Alumni Association. Telephone: 818-395-3630.

PICTURE CREDITS: Cover, 12 – Erika Oller; 2 – Cathy Hill; 4-7– NASA; 8 – JPL; 12 – Erik de Schutter; 14 – Qianli Liu; 15, 43, 44 – Bob Paz; 18 – Caltech Libraries, Jean Mawhin; 30, 42 – Steve Padin; 32, 35, 36, 40, 41– Tony Readhead; 33, 34 – Hubble Space Telescope; 36 – Lucent Technologies; 37 – W. H. Freeman and Co.; 38 – Palomar Observatory; 42 – Albert Stebbins; inside back cover – Herb Shoebridge

Random Walk

How Now Mad Cow?

The overall goal of the project was an objective description of various diseases of the brain and nervous system, including Alzheimer's disease, Parkinson's disease, Huntington's chorea, and schizophrenia, by the spinal-fluid population of particular proteins that might serve as markers for the diseases.

Two small spots, the modest signature of two proteins in a sea of others, have led to an unambiguous diagnostic test for Creutzfeldt-Jakob disease (CJD), a rare but inevitably fatal neurological disorder. Developed recently by researchers at Caltech and the National Institutes of Health, the test involving the same two proteins appears also to have great potential for diagnosing bovine spongiform encephalopathy (BSE). Better known as mad cow disease, BSE has ravaged Britain's cattle industry and created considerable political uproar in the European Union since the March announcement of the hypothesis that the disease had entered the human food chain. As many as 15 British cases of a new variant of CJD have been tentatively linked to the consumption of BSE-infected beef, and several hundred cows a week are still being diagnosed with BSE in Britain, despite measures to contain the disease.

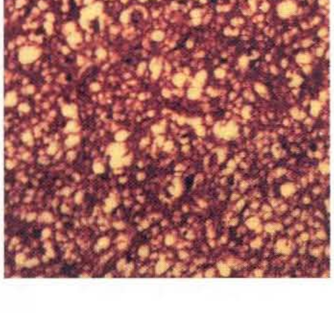
Although only suddenly a very hot property, the two proteins were originally discovered more than a decade ago. Michael Harrington, a member of Caltech's Beckman Institute, spotted them during an NIH project examining the spinal fluid of 541 subjects—about 100 of them normal and the rest suffering from a variety of neurological diseases, including 21 with CJD. Harrington was screening the fluid for about 100 proteins, using two-dimensional gel electrophoresis, a technique that applies an electric field to separate a complex mixture of proteins, which end up as distinct spots of different size and charge in different positions on the gel.

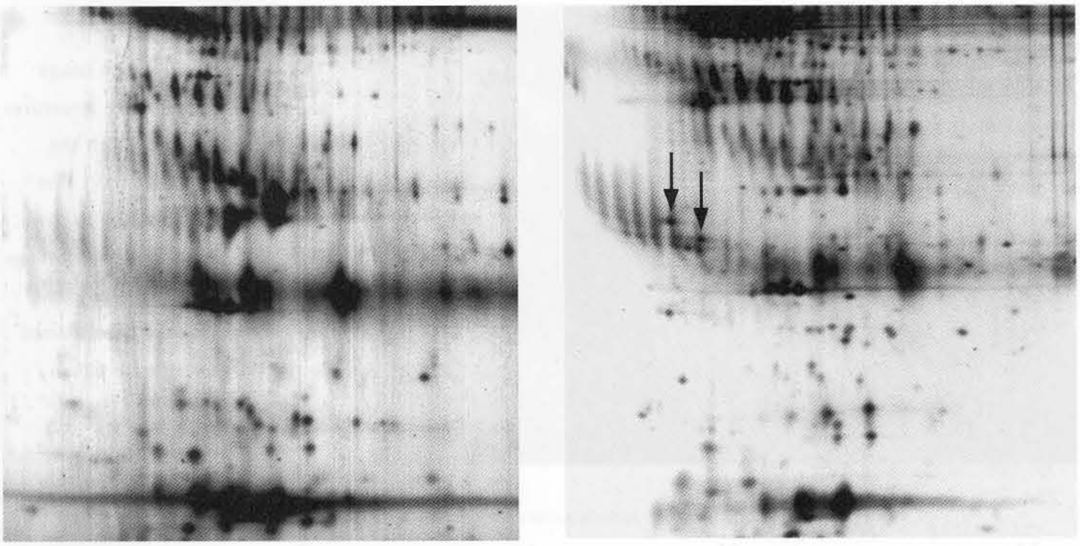
The overall goal of the project was an objective description of various diseases of the brain and nervous system, including Alzheimer's disease, Parkinson's disease, Huntington's chorea, and schizophrenia, by the spinal-fluid population of particular proteins that might serve as markers for the diseases. Harrington, who is a physician as well as a biologist, refers to the spinal fluid as the "urine of the brain" for its diagnostic versatility.

The two spots, known only as 130 and 131, turned up in this spinal "urine" of all the CJD patients and of only the CJD patients; that is, they described CJD with 100-percent specificity and selectivity. Harrington thought the characteristic spots would make a useful diagnostic marker for the disease, and in 1986 he published this in a paper.

Unfortunately, 10 years ago such a diagnostic test took three days of highly skilled labor, and since only about 200 CJD cases occur annually in the U.S., it didn't exactly catch on in everyday medical practice. Still at the NIH, Harrington did continue to follow up on

Brain tissue that resembles a sponge is characteristic of all the transmissible spongiform encephalopathies; this sample is from a person who died of Creutzfeldt-Jakob disease.





Two-dimensional gel electrophoresis, which separates a complex mixture of proteins by size and charge, turned up two spots, originally called proteins 130 and 131, that were present in the cerebrospinal fluid of patients with Creutzfeldt-Jakob disease (right, arrows) but not present in normal cerebrospinal fluid (left).

pathology samples sent to him from all over the world and diagnosed 260 cases of CJD with 99-percent accuracy (one false positive and one false negative). But for all their diagnostic accuracy, proteins 130 and 131 were still unidentified. From their position coordinates on the two-dimensional gel, you can assign a charge and a size—that's all. So Harrington decided to try to find out what they were, a task that became possible through the protein-sequencing techniques that had been developing in the years since his first CJD paper was published.

By 1988 Harrington was working at Caltech, where biologists had pioneered much of the work in protein sequencing. (Image-processing algorithms developed at Caltech's Jet Propulsion Laboratory also made comparison of spot patterns in the gel much easier; see *E&S*, Fall 1990.) But in order to sequence proteins, you first had to isolate them, and even this proved difficult for 130 and 131. Because they exist in such trace amounts in spinal fluid, it proved impossible to obtain a sufficient amount of CJD-spinal fluid to purify a sample large enough to sequence. No amount of spinal fluid that Harrington could obtain would yield enough of a protein sample.

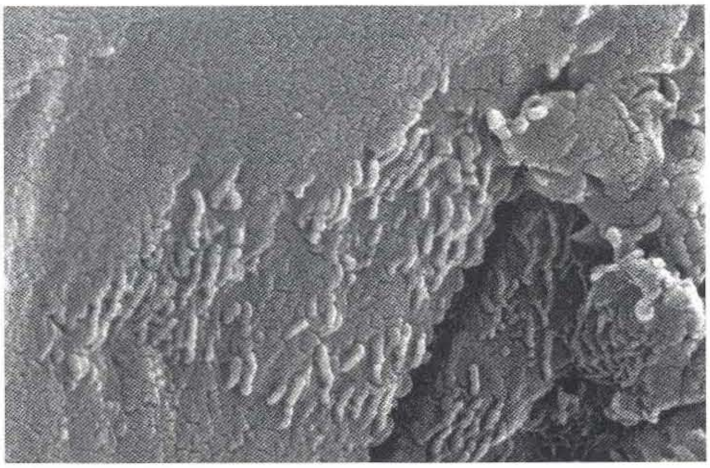
Then Harrington speculated that 130 and 131 might exist in normal brain tissue and only leak out into the spinal fluid from a brain damaged by CJD. This turned out to be correct, enabling Harrington, postdoc Kelvin Lee, and researchers at NIH to purify sufficient pro-tein for sequencing from a mere gram of normal brain tissue. From the peptide fragments that emerged from this process, a partial sequence was obtained—enough to feed into a protein database to look for a match. Protein 130 turned out to be a well-known neuronal pro-tein, 14-3-3, for which commercial antibodies were available. These antibodies should bind to the protein and light up, indicating its presence. When Harrington screened the spinal-fluid proteins from a CJD patient with the antibodies, both 130 and 131 lit up. Harrington and his colleagues now had a simple, rapid, clinical test for CJD; they published it in *The New England Journal of Medicine* at the end of September.

The implications of proteins 130 and 131 reach wider, however, than a diagnostic test for a rare disease, or even for a politically and economically significant one. The two turn out to belong to a highly conserved (in evolutionary terms) family of brain proteins that have already attracted the interest of biologists, says Harrington, because they appear to be involved in a wide variety of functions, including the transmission of signals—and in protein folding. The latter function is of particular interest to Harrington: CJD and BSE are thought to be caused, not by an infectious agent such as a bacterium or a virus, but by an abnormal prion—a protein that propagates out of control and accumulates in the brain. Unlike normal proteins, prions in CJD and BSE are conformationally abnormal. This leads Harrington to speculate that his two proteins that play a role in the conformational stabilization of other proteins might be more than just markers and could perhaps be implicated in the development of the disease. And by manipulating them, you might be able to stop the prion's progress.

In his continuing study of CJD, Harrington now has 15 protein markers for the disease, and markers that may turn out to be even more descriptive of other diseases such as Alzheimer's disease or Parkinson's disease. Harrington has also characterized about 2,000 normal spinal-fluid proteins in his attempt to create a complete profile of proteins in the nervous system, in which specific changes indicating pathology can be observed and therapies monitored. "We can build up a molecular picture of disease," Harrington believes, through proteins, which are more descriptive of disease than genes. He even envisions a massive "Human Proteome Project" to complement the Human Genome Project. "Proteins are what we're made of," he says. "It's how our genes have produced different proteins that make us what we are."

In the meantime, however, Britain's beef disaster may provide the first direct application of Harrington's work. During his research for the 1986 paper, he did look at scrapie (a disease of sheep that is presumed to be the precursor of BSE) to see if a form of proteins 130 and 131 occurred, because the CJD and scrapie brain pathology was similar. All of the sheep proteins, however, lodged at coordinates on the gel different from the human ones, rendering a comparison with a sheep form of 130 and 131 impossible at the time. But the antibody that binds to and lights up the two CJD-marker proteins now changes all that and has enabled Harrington to trace successfully the now-known marker proteins in several different transmissible spongiform encephalopathies that affect sheep, cattle, and chimpanzees.

Harrington has not yet been able to test spinal fluid from any mad cows, but "the data obtained from experimentally transmitted disease in cattle suggest that it might be a useful test for BSE," he says. In the meantime he's been hounded by the British media more tenaciously than any other Caltech scientist in recent memory. □



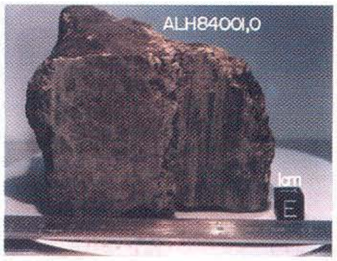
This electron microscope image shows the tiny tube-like structures (the largest being 1/100th the diameter of a human hair) that may be fossils of ancient bacteria. They were found within the orange carbonate globules (below) that are believed to have been formed in the Martian rock 3.6 billion years ago.

Microbes in Martian Meteorite?

Ancient bacterial life on Mars? Announcement of its possible manifestation leaked to the press a couple of weeks before the scientific paper actually appeared in the August 16 issue of *Science*. A team of scientists from NASA, McGill University, the University of Georgia, and Stanford had found several types of evidence for these bacteria in a 4.5-billion-year-old Martian meteorite, named ALH84001 for the Allen Hills region of Antarctica, where it was found after it was blasted off Mars by the impact of another meteorite millions of years ago. While any one bit of evidence might be explained by other means, the concatenation of the data led the authors to conclude that the most reasonable explanation was that it had been produced biologically about 3.6 billion years ago.

A NASA news conference trumpeted the results. The thrill of discovering that perhaps "we are not alone," even if our ostensible companions are microscopic organisms that ceased to exist 3.6 billion years ago, clearly touched a collective emotional nerve. Reactions ranged from jubilant hype to cold-fusion sneers; sci-fi fans cried "I told you so," and the devout made haste to reconcile it with the Almighty's plan; scientists, on the whole, took the paper seriously but maintained skepticism. *E&S* asked some of Caltech's scientists for their opinions.

Norman Horowitz (PhD '39), professor of biology, emeritus, and former chief of the bioscience section for JPL's Mariner and Viking missions (which concluded that life was impossible on the present Martian surface) found the paper technologically impressive, "a serious paper by competent investigators using the most advanced analytical methods." He notes that the authors "don't claim to have proof of Martian life, only that that is the best explanation for their findings," and that they are



The 4.5-billion-year-old rock ALH84001, believed to be part of a Martian meteorite, was found in Antarctica in 1984.

"The one aspect of their finding that I find interesting is that these hydrocarbons are the first organic matter that, as far as I know, has been identified with Mars. The Viking Lander found no organic matter in the surface of Mars."

cautious. "In my view there is much to be cautious about."

"In the first place, the idea of past life in a piece of igneous rock—a rock that crystallized from a melt, which is what the meteorite is—is unusual. Evidence of past life is normally found in sedimentary rock. The authors' position is that fissures in the rock became a home for microbes a billion years after the rock was formed.

"Then, the evidence they present is based almost entirely on inorganic chemistry. Carbonates, magnetite, and iron sulfide found in the fissures of the meteorite are of biological origin, they argue. Their argument is a sophisticated one, based on the particular minerals found, their morphology, and their association in the rock. I know little mineralogy, but I am skeptical of this argument because these substances, and/or their close chemical relatives, are common in the solar system. They are found in meteorites and on the surface of Mars, and they are formed by well-known chemical reactions. I am sure that a knowledgeable chemist could produce a credible nonbiological model for the occurrence of these same minerals in the Mars meteorite involving fewer assumptions than those made by McKay et al.

"The only organic substances that the authors report in the meteorite are polyaromatic hydrocarbons. They are aware that polyaromatics of nonbiological origin are found in carbonaceous meteorites and that they are readily produced in the laboratory. Their argument for a biological source of these meteoritic polyaromatics is based on their mass distribution and other characteristics, none of them biodiagnostic. The one aspect of their finding that I find interesting is that these hydrocarbons are the first organic matter that, as far as I know, has been identified with Mars. The Viking Lander found no organic matter in the surface of Mars.

"Finally, electron microscopy of the chips showed tiny structures that, except for their minute size, resemble ordinary rod-shaped bacteria. The authors suggest that these may be microfossils. No supporting evidence for this interpretation is given."

So Horowitz remains unpersuaded. What would it take to bring him around? He thinks that fossil evidence is required to make the case convincing. The distinctive chemical features of life—such as optical activity—could not survive for 3.6 billion years, and, if found, would indicate contamination.

Joseph Kirschvink (BS, MS '75), professor of geobiology, is an expert on biomineralization. In addition to making the initial prediction and discoveries of earthly magnetofossils, he has found magnetite in the brains of whales, tuna, and humans (the phenomenon of biologically produced magnetite was discovered at Caltech by the late Heinz Lowenstam in the early sixties and confirmed when magnetotactic bacteria were observed in 1975). Kirschvink was quoted in *Science* as believing that a Martian biogenic source for the meteorite's magnetite was "not unreasonable at all." He has written that the "putative Martian magnetofossils. . . look interesting, perhaps the most convincing of all the evidence marshaled in the paper."

Kirschvink claims that the hypothesis of biogenic Martian magnetite can be tested quite easily. Characteristic of terrestrial magnetofossils and true bacterial magnetosomes (groups of magnetic crystals), is the alignment of the magnetite crystals in linear chains. Experiments could show whether the Martian magnetite occurs in linear chains. Also, depending on what the strength of Mars's original magnetic field might have been (Kirschvink believes it was probably substantial) and whether it shifted in direction, and depending on the conditions under which the magnetite-

Feynman's Misplaced Lecture

In Feynman's Lost Lecture, which appeared in our last issue, David Goodstein recalled Richard Feynman giving a guest lecture (not the lost one, but the one in which Feynman spoke of "his" supernova) to Goodstein's freshman physics class shortly before Feynman died. David A. Edwards, BS '90, PhD '94, who was in the freshman physics class (and so was his wife) in December 1987, when this lecture allegedly took place, writes that it wasn't so—that neither Feynman nor Goodstein taught the class—and sent along his homework assignment (listing McKeown and Frautschi as instructors) to prove it. Goodstein researched the class records, and found that indeed his memory was faulty, and that Feynman had delivered his last guest lecture to Goodstein's class on March 13, 1987, just a month after the supernova, but 11 months before Feynman's death.

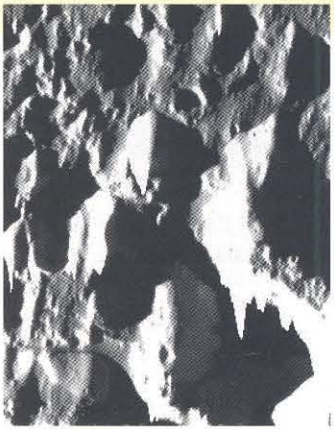
And a Misplaced Person

In the last issue of *E&S*, the caption on page 10 accidentally transposed the names of the two graduate students. Brett Doleman is, in fact, on the left, and Erik Severin is on the right.

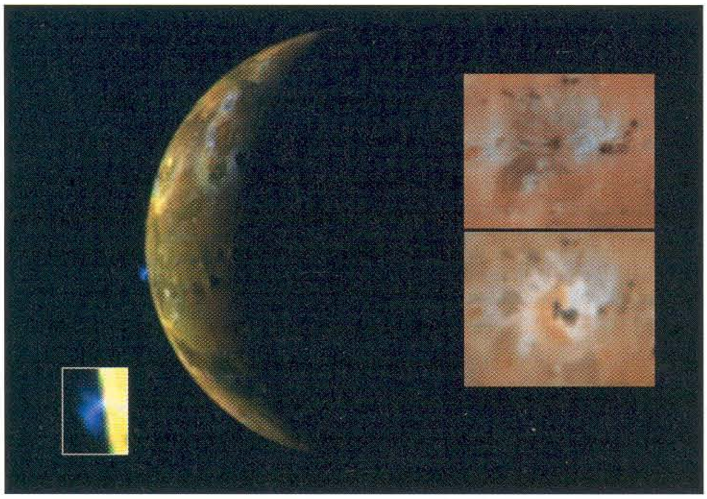
containing carbonate was formed, various experiments might be able to determine whether the magnetization direction of the magnetite in ALH84001 is different from what would be expected of its host rock.

Kirschvink's own lab is the only magnetically shielded, clean-lab facility in the world housing a superconducting magnetometer system, which, although designed for studies of biogenic magnetite in animal tissue, could also accommodate a piece of rock. And his 15-year-old SQuID (superconducting quantum interference device) moment magnetometer was recently rebuilt and equipped with new sensors that would easily be able to measure tiny amounts of Martian magnetite. Hojatollah Vali, of McGill University, one of the *Science* paper's coauthors, was a visiting associate in this lab in 1989, and Kirschvink has offered his facility for further tests to help determine whether life did indeed exist on ancient Mars. □

Galileo Shoots the Moons



On July 27 the Galileo orbiter spacecraft flew by Jupiter's moon Ganymede within 519 miles (70 times closer than Voyager 2), sending back stunning images of incredible sharpness. Ganymede, the largest moon in the solar system, is believed to be about half water ice and half rock. Galileo's experiments revealed that Ganymede has a magnetosphere, and its pictures show the moon's surface to be pock-marked by impacts with comets and asteroids and wrinkled and fractured by the same internal forces as Earth. While it was in the neighborhood, Galileo also snapped some shots of Europa from about 96,300 miles and Io from 604,000 miles. The spacecraft will pass closer to Europa in December.



Left: A new volcanic plume, colored blue by sulfur dioxide and extending 60 miles into space, erupts at Ra Patera on the constantly changing surface of Io. The inset at right shows how this volcano has changed since Voyager's visit in 1979 (top). The new plume has covered a surrounding area the size of New Jersey with the volcanic deposits seen by Galileo (bottom).

Caltech Almost Makes the Movies

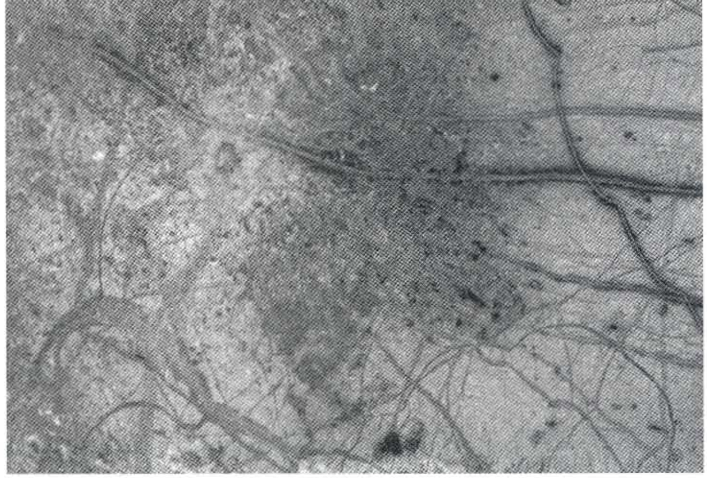
Sharp-eyed moviegoers amongst us may have picked up on the Caltech presence in last summer's brace of alien-invasion flicks. An early scene in *Independence Day* shows the military brass trying to figure out why all their satellites are going off-line. During the conversation, a badly pixellated, purple-and-red photo—obviously an infrared-telescope image—is handed round. The image shows a mysterious round object, the size of a minor planetoid, orbiting Earth. (This, of course, is the alien mother ship, but only the audience knows that then.) A couple of scenes later, a second photo shows a set of thin, flat objects (the just-launched invasion fleet) underneath the mother ship. It's only on the screen for a few seconds, but emblazoned across the bottom of the photo is "10-METER KECK TELESCOPE, MAUNA KEA, HAWAII." No other affiliation is given, however.

And the folks who took in *The Arrival* saw Caltech but didn't know it. Caltech's Owens Valley Radio Observatory played the part of Oro Valley Observatory in the film. Several exterior shots were filmed there, including those of the actors clambering about in the 130-foot dish. On the other hand, the scenes that purported to be of Caltech's Jet Propulsion Laboratory, where radio astronomer and SETI (Search for Extra-Terrestrial Intelligence) expert Zane Ziminski (Charlie Sheen) worked, were shot elsewhere. And to round out the "things are not what they seem" category, Tony Award winner Ron Silver plays Gordian (as in knot), the head of JPL who's also One Of Them.

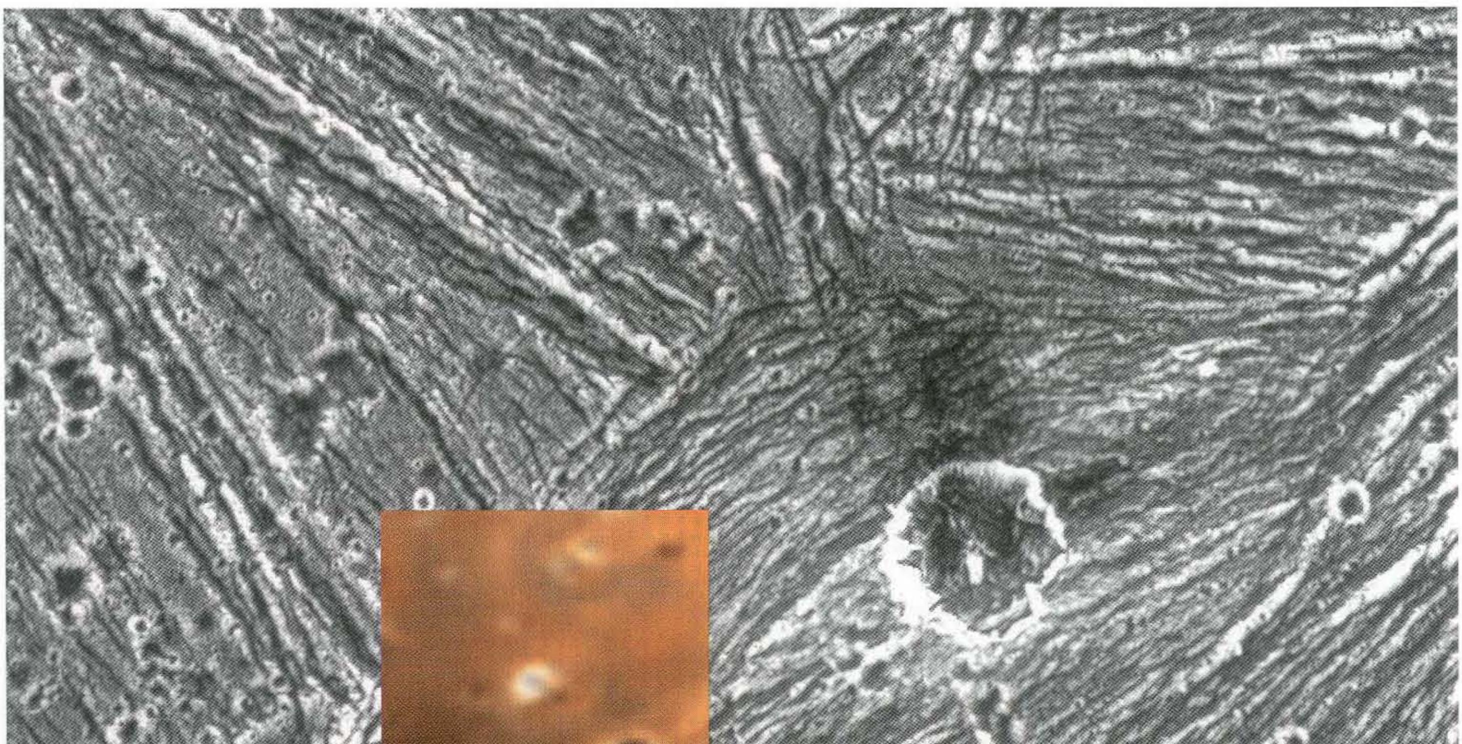
And released shortly before *E&S* went to press is *Infinity*, starring Matthew Broderick as a very young Richard Feynman and Patricia Arquette as Arline Greenbaum. Greenbaum, Feynman's first wife and the great love of his life, died of tuberculosis in an Albuquerque sanatorium while Feynman worked on the Manhattan Project at Los Alamos, a hundred miles of bad road away. Since Feynman didn't come to Caltech until 1950, Caltech isn't mentioned in this movie, either. □

Left: Ice hills in an unnamed region of Ganymede, their western sides lit up by the sun, were seen in 2000-times greater detail than any previous images. The smallest objects here are only 11 meters across.

Below: This terrain in the Uruk Sulcus region is typical of about half of Ganymede's surface. The ancient, pock-marked, cratered terrain to the left (north) is cut by younger striations on the right. Next to the large impact crater at lower right, some dark material has been ejected onto these linear ridges. The smallest features that can be seen here are about 74 meters across.

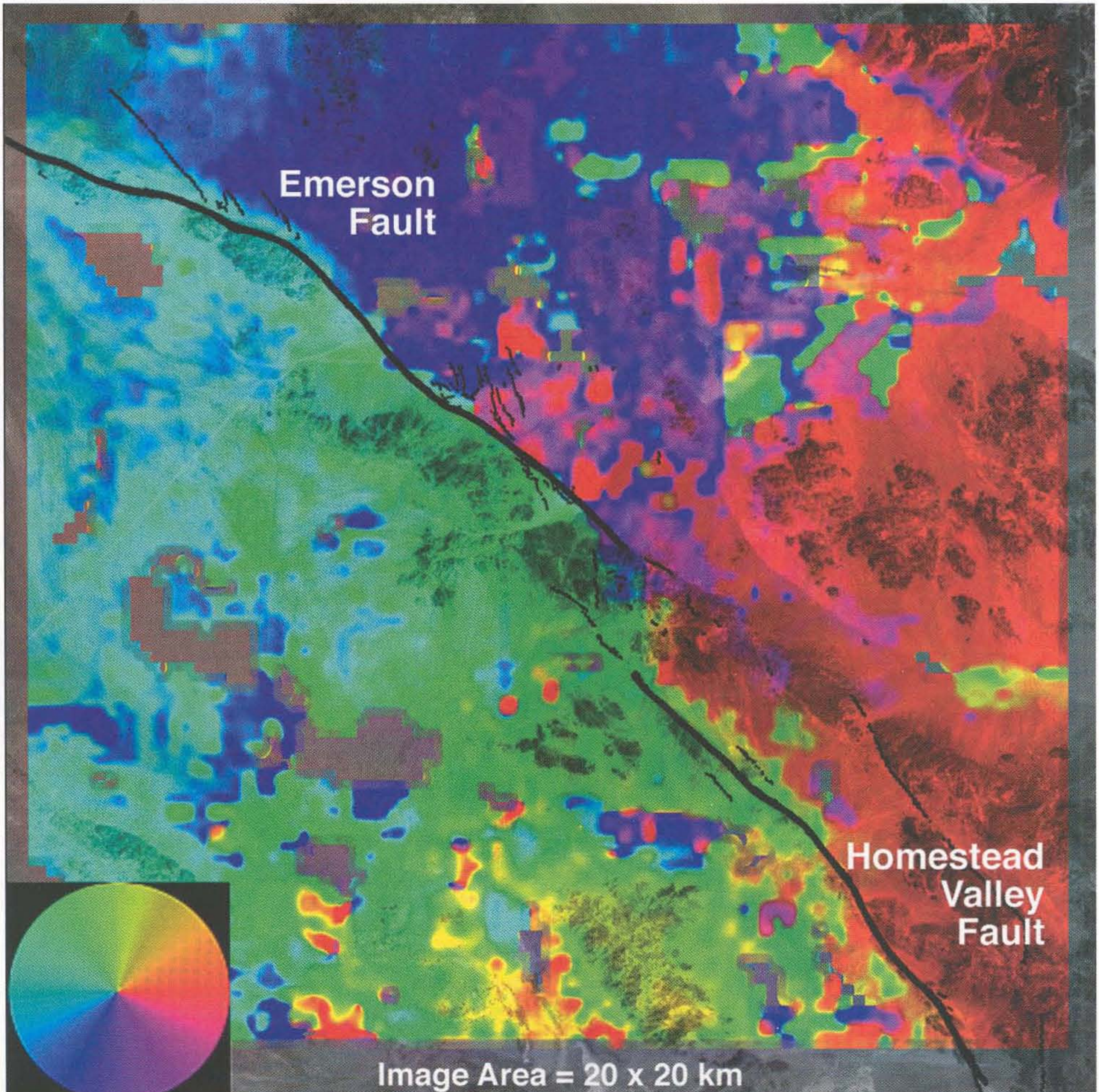


Above: On Jupiter's moon Europa, a new impact crater (left of center), about 30 km across, scattered light-colored debris over a wide area as it struck Europa's icy crust. The X-shaped pattern at right probably occurred as the icy crust fractured and then filled in with slush.



Galileo's view (bottom) of Euboea Fluctus on Io (seen here in simulated Voyager color for comparison) shows diffuse material that has been deposited around the volcano since Voyager 2 flew by in 1979 (top). It extends over a radius of 285 km.

"Faculty from many disciplines are discovering, pretty much on their own, that many very large, computationally intensive problems are now becoming approachable, but *only* through these big, concurrent machines."



The Nodes Know

by Douglas L. Smith

The colors in this picture show the displacements caused by the Landers earthquake, as revealed by analysis of satellite photographs taken before and after the quake. (North is at the top; the faults have been drawn in as black lines.) The ground moved in the direction shown in the color wheel in the lower left corner—for example, northeast motion would be represented as yellow. A new analytical method, practical only with a supercomputer, can retrieve motions that are much smaller than the size of an individual pixel. The result is pixel-by-pixel readings of ground motions—an impossible feat for a survey team.

In the late 1960s, computers of any size were scarce, and time on the big ones was a precious resource. That's when Professor of Chemical Physics Aron Kuppermann started knocking atoms together in an IBM 360. He was modeling the simplest possible chemical reaction—hydrogen exchange, in which a hydrogen atom slams into a hydrogen molecule and replaces one of its two atoms. Even so, the computer was barely up to the task—it could only handle the model as long as all three atoms were kept in a straight line. Caltech ushered in the 1970s by acquiring a bigger IBM, the 370, and Kuppermann was keen to try hydrogen exchange in three dimensions. But in those days, the Campus Computing Center charged for computer time, and the new machine went for $300 an hour. Kuppermann didn't have that kind of money, but discovered that "there was a place in town that had a duplicate of Caltech's machine—at the time essentially the world's most powerful—and this was the Worldwide Church of God. They used it to keep track of their donors." He also discovered that they didn't use it on weekends and religious holidays. He somehow talked the church leaders into letting him use the machine during that idle time, a feat he calls a minor miracle. But the ground rules were strict—concerned for the privacy of their donors, they wouldn't let him anywhere near the machine. So on Friday afternoons, he would drop off at a little cashier's window the box of punched cards that was the weekend's calculation. On Monday morning a printout—read to be sure it contained no donor information—would be handed back to him through the same window. "If they read in my box of cards and one got mangled, that was it—a weekend lost." Kuppermann logged more than a thousand hours on that computer while finishing the world's first complete three-dimensional quantum-mechanical calculation of a chemical reaction. "Without it, we couldn't have done the work."

Thirtysomething years later, computers are ubiquitous, but the *really* big ones are still relatively rare, and Caltech still has some of the most powerful computers in the world. Unlike those IBM mainframes, which executed one instruction at a time, today's zippiest machines are "concurrent" computers that execute many instructions simultaneously. Rather than having one large, fast brain, concurrent computers have lots and lots of small or medium-sized brains, called nodes, working in parallel. Each node chews on its own piece of the problem independently, passing information back and forth to its fellows as needed. This approach, developed in the early 1980s at Caltech and elsewhere (*E&S*, March 1984), is ideal for attacking problems where the same basic calculations have to be applied to a large set of elements—be they stars in a gravitational field, grid points in a wind-tunnel flow, or eigenvalues describing the quantum state of a nuclear particle. Caltech's two largest machines, the Intel Paragon and the Intel (Touchstone) Delta, have 512 nodes apiece. Each node is about 10 times as powerful as that old IBM 370. This bequeaths a huge increase in computational speed—the 370 could execute a million floating-decimal-point calculations per second; the Paragon does ten billion. Meanwhile, says Paul Messina, assistant vice president for scientific computing and director of the Center for Advanced Computing Research (CACR), "faculty from many disciplines are discovering, pretty much on their own, that many very large, computationally intensive problems are now becoming approachable, but *only* through these big, concurrent machines."

These faculty are in the vanguard of a quiet revolution in the way science works. Because concurrent machines can handle complex models jam-packed with minutiae, the monogamous relationship of theory and experiment has become a ménage à trois of theory, model, and experiment, where results from any one can spur developments in the other two. The availability of so much raw computational power has also led to "data mining," the sifting of huge mounds of information for correlations that would be missed if it weren't possible to crunch all the numbers six ways from Sunday. In the environmental, earth, and planetary sciences, for example, data sets from different satellites can be amalgamated for exploration. (This has been eased by the development of high-speed data links between computers, and Caltech has been involved in that, too, but that's another story.) Here are some dispatches from the front lines of that scientific revolution.

Professor of Physics Thomas Prince and his colleagues are working to integrate a supercomputer into a radio telescope—essentially making the computer just another panel on the instrument rack. Normally, the incoming data is piped through a bank of amps and filters to pump up, say, the blip-blip-blip of a distant pulsar while hopefully squelching the local interference from radars, TV stations, and what have you before being recorded on tape for later analysis. "But you're limited in how much data you can record," says Prince. "So you need to throw out a lot of information. Ideally, you'd like a digital receiver, but this would take as much computational capacity as the Delta." Such a receiver would amplify, filter, and analyze the signal in real time. The analysis requires small slices of time to measure the pulsing of the pulsar, as well as small slices of frequency over a very broad range so you can study the pulse's shape. All this generates data at a prodigious rate. To further complicate things, the signal is smeared out by its passage through the plasma clouds (clouds of charged particles) in our galaxy, and so a pulse's frequency components will arrive on earth at slightly different times. The analysis has to "de-disperse" the pulse to reconstruct its original form, but the average workstation can only handle a few seconds worth of signal at a time. It would be like trying to fill a demitasse cup with a fire hose.

Since both the Delta and the Paragon live on campus, hundreds of miles from the nearest radio telescope, Prince's group designed a custom chip to do the next best thing—digitize as much of the raw signal as possible and record it on a very fast tape deck. The faster the tape, the wider a chunk of the spectrum you get, and the more accurate the time information. Then the Paragon could pretend to be a digital recorder while listening to the tape. Datatape Incorporated donated the tape decks for the telescope and the computer to CACR, and collaborated in their installation. These babies whir at 50 megabytes (which would fill 36 high-density floppy disks) per second—the fastest recorders then available commercially. After being tested on the 40-meter dish at Caltech's Owens Valley Radio Observatory, the system got its first real workout on the 64-meter dish at Australia's Parkes Observatory in July 1995, recording 10 terabytes (the equivalent of all the Library of Congress's printed holdings) of data from an assortment of objects, including globular clusters such as 47 Tucanae. This took 100 cassettes somewhat larger than VHS tapes—a bargain considering that the equivalent stack of floppies would stand 14 miles high.

Prince's group focused on 47 Tucanae, which is known to have at least a dozen very fast pulsars, in hopes of finding a record-breaking rotation rate. (A pulsar is a rapidly spinning neutron star—the cinder left after a dying star explodes into a supernova—that emits a powerful radio beam along the axis of its magnetic field. If the star's magnetic and rotational axes don't line up, the beam sweeps through space like the beacon from a cosmic lighthouse. If the earth lies in the path of the beam, we see a radio pulse—hence the name—with every sweep.) The fastest known pulsar spins once every 1.5 milliseconds. This is approaching the speed limit, believed to be somewhat less than one millisecond, where matter on the surface of the neutron star (which is about 10 kilometers in radius) is whirling around the axis at nearly the speed of light. (And you thought the teacups at Disneyland made you dizzy!) The search came up empty, but the tapes contain lots of other interesting data that's still being analyzed.

At the opposite end of the cosmic distance scale, Professor of Theoretical Physics (and vice president and provost) Steven Koonin (BS '72) has been using the Delta, the Paragon, and other machines to look inside atomic nuclei. You can think of the energy levels within a nucleus as a ladder of infinite height. Each rung is wide enough to hold two nucleons—two protons, two neutrons, or one of each—and each particle seeks out the lowest unoccupied rung. But the particles continuously interact with one another through the strong nuclear force, kicking each other up and down the ladder, creating and filling vacancies in the low-energy rungs. Thus, to find the overall configuration of the nucleus, you set up a table, or matrix, that lists all the interactions between every possible configuration. Calculating the strengths of these interactions—diagonalizing the matrix, it's called—gives you all the possible energy states and their probabilities of occurring. This works

This took 100 cassettes somewhat larger than VHS tapes—a bargain considering that the equivalent stack of floppies would stand 14 miles high.

fine for light elements such as carbon, with 12 nucleons, but things rapidly get out of hand. The mere 28 nucleons in silicon have about 100,000 possible energy states; zinc, with 60 nucleons, about 30,000,000; and the rare earth dysprosium, bustling with 166 nucleons, has a staggering 10^{21} (one sextillion) possible states—a number that would make even a federal economist blanch.

Diagonalizing a dysprosium-sized matrix in its full glory remains out of the question, but a few years ago Koonin, postdoc Erik Ormand, and grad students Calvin Johnson and Gladys Lang realized that all they really needed was a statistically valid sample of the matrix's energy states. By using this technique, the heaviest nucleus calculated has jumped from 35 nucleons (chlorine) to 76 (germanium)—a very considerable leap, but one that leaves most of the periodic table yet to go. Meanwhile, Koonin's group has applied the method to two long-standing theoretical problems. One has to do with the Gamow-Teller process, in which a proton, upon swallowing an electron, gags, coughs, flips its spin, loses its charge, and turns into a neutron—a critical step in the "neutronization" of iron that leads to the production of heavier elements in the fiery belly of a supernova. The rates of those Gamow-Teller processes that have been measured experimentally are less than 30 percent of what was predicted by previous calculations, but these simulations are right on the money. The other problem concerns double-beta decay—an event so rare that it takes 10^{20} years to happen to the average nucleus. This is unfortunate for people watching for it, as the universe is only about 10^{10} years old, but it's a good thing for the rest of us—if the nuclei were decaying much faster, we wouldn't be here to watch them do it. For reasons too discursive to go into, the rate of double-beta decay is linked to the mass of the neutrino, which is the most evanescent of particles and is generally assumed to be massless. If the neutrino proves to have even an infini-tesimal mass, it will make the universe a much heavier place, and one more likely to eventually recollapse on itself in the so-called Big Crunch. Again, previous calculations of double-beta decay rates were seriously inconsistent with reality, but the calculations by Koonin's group matched the data nicely.

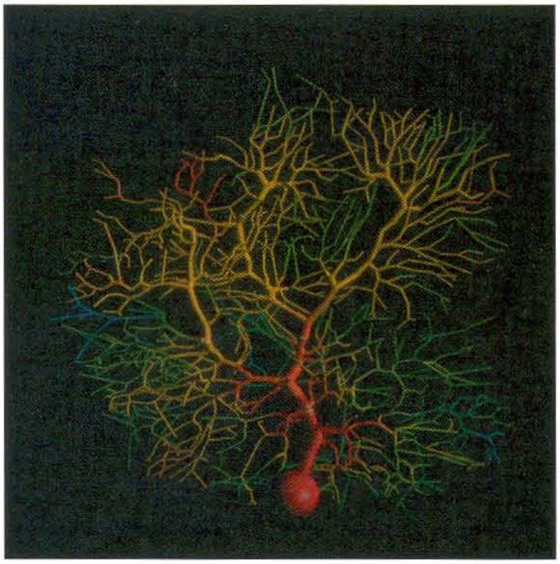
The group is now moving on to the bizarre properties that heavy nuclei exhibit when spun and heated. At room temperature and with a normal nuclear spin of "only" 10^{19} revolutions per minute, a nucleus can be spherical. But heated to several million electron volts (about 10^{10} degrees Centigrade) and cranked up to 10^{21} rpm, it can become football- or even cigar-shaped. It can even flatten, like a cigar that's been run over by a truck. Such nuclei have actually been observed experimentally, but the theory to describe them has been lagging behind. But now, says Koonin, we can rev them up in the computer and see how their structures change.

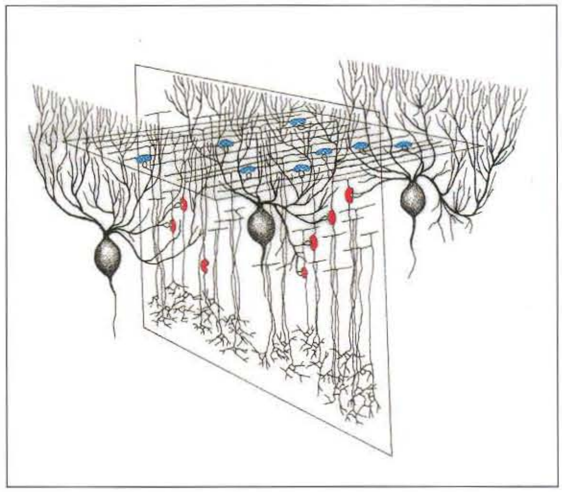
Paul Stolorz (PhD '87), technical group supervisor of the machine learning systems group at JPL, is also watching structures change, but his are geological. Stolorz and then-undergrad Chris Dean, along with JPL geologists Robert Crippen and Ronald Blom, have refined a method Crippen and Blom invented that compares a pair of before-and-after pictures and computes the motion of each pixel to within 10 percent of its width—a form of data mining. For example, pictures of the Mojave Desert taken by the French SPOT satellite before and after the Landers earthquake show no visible changes, because the quake's maximum offset was only six meters and each pixel covered a 10-by-10-meter chunk of desert. But the collaboration was able to measure displacements to within a meter by taking a block of pixels (100 by 100 in this case) from the "before" picture and sliding them around on the "after" picture to get the best match. This gave the motion of the block's central pixel very accurately, because the individual matching error for any pair of pixels got buried in the statistics of the whole block. The process was repeated for similar blocks centered on each of the 40,000 pixels in the overlapping image area. This comparison, and its preliminary image processing—subtracting out the motion of the spacecraft and compensating for the different look angles—took JPL's Cray T3D, which has 256 nodes, the equivalent of 24 hours of uninterrupted work. It would have taken months to do on a workstation.

The rare earth dysprosium, bustling with 166 nucleons, has a staggering 10^{21} (one sextillion) possible states—a number that would make even a federal economist blanch.

The collaboration is now reexamining old Mariner 9 and Viking Orbiter photos of Mars for possible evidence of sand-dune motion. They will also be comparing Voyager's images of Jupiter's moon Europa with the upcoming Galilean ones (the first of Galileo's three close flybys of Europa will be on December 19, 1996) for evidence of surface motions that would betray the presence of liquid water beneath its crust of fractured ice. (See page 7 of this issue.) But the method is applicable to any image composed of pixels, even medical MRI scans.

Stolorz's group, in collaboration with biochem-





Right: The colors in this model of a single Purkinje cell represent its electrical potential partway through the firing process. The redder a portion of the cell is, the more recently the impulse moved through it. The yellow and green portions of the dendrite tree are returning to their resting state, and the blue parts are fully recovered and ready to transmit the next impulse. Far right: The relationship between the Purkinje cells (the beet-like objects) and the granule cells below them. As the granule cells' axons ascend, they make connections (red) with nearby Purkinje cells. Once the axons split, the parallel fibers make additional connections (blue) to Purkinje cells.

ists at the University of Vienna, Los Alamos National Laboratory, the Santa Fe Institute, and the University of Illinois, is doing some biology already—calculating how a piece of RNA will fold, based on the sequence in which its letters are strung. RNA is a wonderfully versatile molecule. In higher organisms, RNA is the messenger molecule that takes DNA instructions to the cellular machinery, but in some viruses, including HIV, the RNA is all there is—it both stores the genetic instructions and carries them to the host cell's co-opted machinery. And in all organisms, RNA plays a variety of roles in the machinery itself.

Like DNA, there are four letters in the RNA code, and the letters pair up with one another in only one way. So the individual calculations of which letters will willingly pair up are very simple, but the number of overall calculations increases with the cube of the number of letters being considered, and the memory required to juggle the calculations increases with the fourth power. Other machines had hit the ceiling at about 3,000 letters, but the Delta plowed through all 10,000-odd letters in an HIV virus—the first time a predicted structure of such length had ever been computed. Stolorz's group has since folded all of the other 19 known strains of HIV to determine the similarities and differences between them. The structures that recur from mutation to mutation are obviously important to the virus's survival, and analyzing them might suggest how the virus could best be attacked. For example, one of the biggest hurdles to developing an HIV vaccine has been the rapidity with which the virus mutates, making vaccines developed against one strain ineffective against others. Locating features that remain constant from mutation to mutation might give the vaccine designers a fighting chance. And understanding how the other features differ could illuminate the various ways the virus attacks us. "These calculations are just best guesses," Stolorz emphasizes. "It's very difficult to verify these kinds of structures experimentally. But the things people do know are consistent with our calculations." And viruses three times the size of HIV, which would include Ebola's 20,000 letters, appear to be within the Delta's reach.

Associate Professor of Biology James Bower and his colleagues have cast their eyes rather higher up the evolutionary ladder—they are looking at mammalian nervous systems. Says Bower, "These are realistic models, based on anatomy and physiology, that require solving thousands of differential equations at once." The models can reflect many levels of detail: subcellular, single-cell, and multicell. On the subcellular level, for example, the equations describe the way ions diffuse across the cell membrane as the cell fires; how the proteins that make up the ion channels through that membrane are flexing to open and close them; and how the cascade of kinases—the signaling proteins that control the process—are behaving. Zooming the microscope out a bit, Bower's group has constructed detailed models of single neurons that allow the researchers to try to understand how the cell's complex structure relates to its function. And finally, cells can be linked together into networks, and even systems of networks, to explore how the brain's circuits work. These models were created using GENESIS (GEneral NEural SImulation System), a versatile software package Bower and his group developed a decade ago that is now used worldwide.

While GENESIS will run on any machine, even a workstation, only recently have supercomputers become powerful enough to exploit the software's potential. "A closed solution of an analytic problem means you know everything that's going on in the system," says Bower. "But in biology, you have thousands of variables, so all the solutions can't be known. Instead, understanding how nervous systems work will increasingly depend on numerical simulations that explore a complex set of parameters." In other words, you run the model over and over and over again while you twiddle various parameters to build up a picture of how the system behaves. The models are compact enough to put a copy (sometimes several copies) on each node of the Delta or the Paragon, but running them often enough (say 500 runs) to get meaningful statistics soaks up machine time—the equivalent of years on a workstation.

The models' compactness belies the level of detail they include. For example, the Bower group has been studying a model of a Purkinje cell—a kind of neuron that lives in the middle layer of the cerebellar cortex and is involved in sensorimotor coordination. The Purkinje cell receives inputs through its enormous dendrite, which looks a bit like the Engelmann oak by Millikan Library—a spreading mass of gnarled limbs and branches. The model, created by then-postdocs Dieter Jaeger, now at Emory, and Erik De Schutter, now at the University of Antwerp, reconstructs this tree in GENESIS by linking together 4,588 electrically distinct compartments, each of which has a characteristic set of up to 10 different types of ion channels. When the cell is stimulated, the flow of ion currents along the dendrite can be "seen" directly as each compartment's channels open and close in succession.

The model has suggested that the cell works in a fundamentally different manner than previously assumed. A Purkinje cell has some 175,000 sensory inputs, so it's literally bathed in continuous, random stimuli. Some are excitory and make the cell want to fire; others are inhibitory and tell the cell to mellow out. Excitory stimuli come from the granule cells, which the model suggests may have two very different effects on the Purkinje cells. The granule cells are located beneath the Purkinje cells, and lie in the densest cell layer in the mammalian brain (a staggering 6,000,000 cells per cubic millimeter). The granular cells' output fibers, called axons, ascend out of the granule-cell layer and make up to 100 connections with the immediately overlying Purkinje cells. Each axon then splits, forming two so-called parallel fibers that run in opposite directions to each other, but parallel with every other granule cell's parallel fibers. The parallel fibers pass through the Purkinje cells' dendrites like telegraph wires through overgrown trees along a badly maintained right-of-way, with each fiber making at most one connection with each Purkinje cell it encounters. For 100 years, it has been assumed that the connections between the parallel fibers and the dendrites were the important ones. But the simulations have suggested that the synchronous activation of the connections on the ascending part of the granule cell's axon may have a much more profound effect on the Purkinje cell than the sequentially activated parallel-fiber connections. Within months of Bower's lab obtaining this result, corroborating physiological data was independently published by another lab.

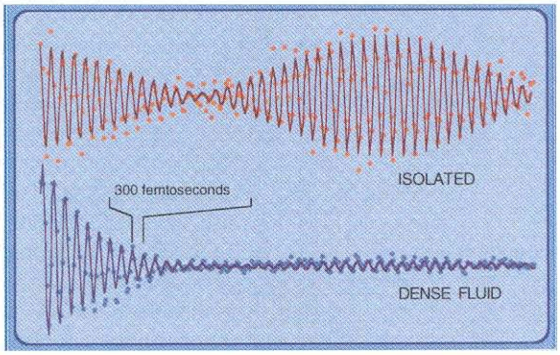
While some folks are modeling processes that occur in biological brains, others are modeling processes that are used to build silicon brains. Professor of Theoretical Chemistry B. Vincent McKoy and collaborators are studying the collisions of electrons and chlorofluorocarbon gases (CFCs), a process critical to manufacturing computer chips. When you hit a CFC molecule with an electron, you can break off a chlorine or fluorine atom—voracious fellows that eat silicon and practically everything else, and which are used to etch the electronic circuits onto the chip. But to keep the process under control and avoid frying the chip, the reaction has to be run at relatively low temperatures (say, 80° C). Nonequilibrium, low-temperature plasmas are well suited for this, because the electrons in the plasma are energetic enough to shatter a CFC molecule upon collision with it. There's surprisingly little experimental data on how CFCs disintegrate under these conditions, so designing the etching systems has relied extensively on intuition and experience. "We wanted to model the process in such a way that it relates electrons and CFCs to the composition of the plasma and to the etching results," says McKoy. "Then you can design the process more intelligently."

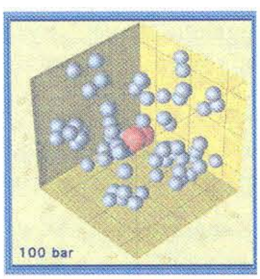
Of necessity, the model started from first principles, which in chemistry means the dreaded Schrödinger equation.

Of necessity, the model started from first principles, which in chemistry means the dreaded Schrödinger equation. This equation in principle provides a complete, quantum-mechanical description of the motion of any molecular system, but there's a catch—as soon as more than a few particles are involved, the equation gets fantastically complex and incredibly difficult to solve. There was a way out in this case, however—because the incoming electrons whiz by the atomic nuclei in 0.1 to 0.01 femtoseconds (a femtosecond is 10^{-15} seconds; it takes about one second for light to travel from the earth to the moon, while in one femtosecond light traverses 1/100th the thickness of an eyelash), the nuclei don't have time to react and could thus be treated as stationary objects. The calculation fires electrons at the CFC molecule over and over again at various speeds and angles, and tracks how the molecule's electron cloud quivers under the blows. Where the electrons end up then tells the researchers how the molecule will

Below: At low solvent pressures (top), the reacting molecule (red) can spin unimpeded by the solvent molecules. But as the pressure goes up (bottom), the reacting molecule gets hemmed in.

Right: The experimental version, done with iodine molecules dissolved in helium or argon. The plot shows the iodine–iodine bond stretching and shrinking, with a period of 300 femtoseconds. At high densities (bottom), the vibration dies out rapidly. But at low densities (top), the vibration persists long enough for the individual bonds to drift slowly out of, and back into, phase—just as two notes at very nearly the same pitch will "beat" against each other.





fall apart. Even so, as of a few years ago, the only electron-collision processes that had been modeled involved very simple gases such as nitrogen and carbon monoxide. More complex molecules just had too many electrons and too little symmetry (symmetrical electron clouds are easier to calculate).

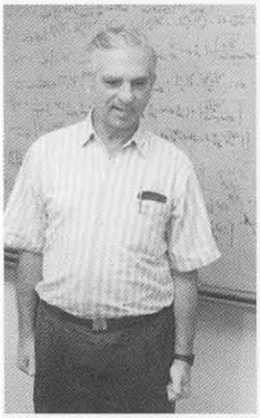
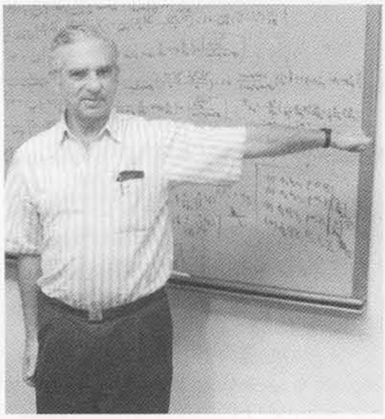
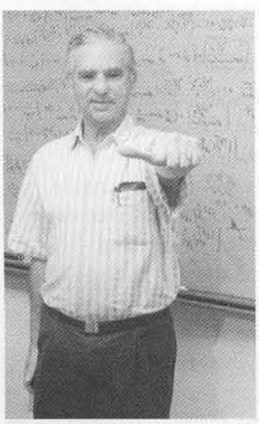
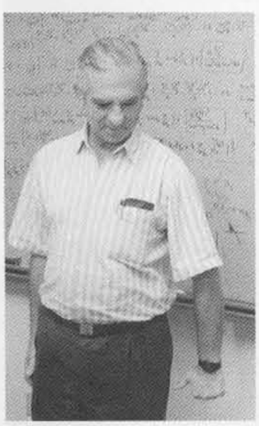
But with the advent of the Delta, McKoy, Senior Research Fellow Carl Winstead, and grad student Howard Pritchard (PhD '94) were able to calculate the electron-collision probabilities for boron trichloride, a common etchant, within a few months. The group has now moved on to bigger game, and is almost finished calculating C_2F_6 and C_4F_8 as part of the first year's work in a longer-term project to model the electron-collision probabilities of important etchant gases. This project, sponsored by SEMATECH (the national research consortium of semiconductor manufacturers), is one component of an ambitious plan by McKoy, Assistant Professor of Computer Science Stephen Taylor, and Sadasivan Shankar of Intel to simulate the low-temperature plasma etching process, in which the breakup of the etchant gas is merely the first step. Such simulations will provide the foundation for improved computer-aided design tools that could reduce the cost of developing the next generation of etching equipment.

Moving from the gas to the liquid phase, Ahmed Zewail, Pauling Professor of Chemical Physics and professor of physics, is examining chemical reactions in a solvent. It's long been known that the solvent molecules play an important part in determining the course of the reaction, but there were just too doggone many of them to simulate them in any detail until recently. Zewail, grad student Qianli Liu, and research fellow Chaozhi Wan took up the challenge in an attempt to explain some very odd results that Wan had been getting in experiments he had been doing with the group's femtosecond-spectroscopy apparatus (*E&S*, Spring 1988), which "freezes" a reacting system in slices of time a few femtoseconds wide and allows a "movie" to be made of a reaction as it occurs.

One expects a reaction rate to increase smoothly with the number of atoms per unit volume. After all, most reactions are molecular fender benders, happening when one atom collides with another, so it makes sense that they'd happen more frequently in crowds than out in the wide-open spaces. But Wan was finding instead that the rate remained constant over a pretty wide range of densities—from five to 15 atoms per cubic nanometer, in fact. (By comparison, nitrogen at room temperature and pressure has .05 atoms per cubic nanometer; methanol has 26; water 33.)

So Liu used the Delta to look at the 500 solvent molecules closest to the reaction, and calculated their motions in relation to the reacting atoms as the reaction proceeded. Because these simulations focused on the atomic nuclei, which move much more slowly than the electrons, the calculations could assume that the electrons were keeping up and sidestep the Schrödinger equation by using non-quantum-mechanical approximations for the atoms' motions.

The simulations, run at pressures ranging from zero up to liquidlike densities, showed that the reaction rate did in fact increase smoothly, but only up to the point where the molecules began to need more elbow room. In a very diffuse solvent, the simulation showed the reacting molecules were rotating very rapidly, as everyone knew. But what hadn't been appreciated was that the centrifugal force from this rotation stretched the bond that was eventually going to break, priming the reaction. The incoming atom didn't have to slam the rotating molecule too hard to knock the outbound atom free. But as the neighborhood got overcrowded with solvent molecules, the reacting molecule kept whacking into the solvent molecules and couldn't spin freely. With-

A macroscopic example of a geometrical phase effect. Kuppermann's thumb and wrist remain rigid in relation to his forearm—in other words, the thumb's position in a local coordinate system based on the forearm remains unchanged. Yet as Kuppermann rotates his arm, the local coordinate system rotates in relation to a global coordinate system, and the thumb rotates in relation to Kuppermann's body.

out the centrifugally induced bond stretching, the incoming atom had to hit the reacting molecule harder. These two effects—centrifugal stretching and collisional slowdown—neatly balanced each other over the density range of five to 15 atoms per cubic nanometer.

There's a popular saying among mathematicians that physics is just applied math, chemistry is just applied physics, and biology is just applied chemistry. Professor of Chemical Physics Aron Kuppermann must be a mathematician at heart, as he hopes to eventually derive a large amount of chemistry from the Schrödinger equation. Because molecule–molecule collisions are about 1,000 times slower than electron–electron collisions, Kuppermann can't ignore the nuclei's motions the way McKoy can, but he can still simplify the calculations by treating the electrons and the atomic nuclei independently because the electrons move so much faster. First, Kuppermann "freezes" the nuclei, and calculates the motions of the electrons for every possible nuclear configuration. This gives the so-called potential energy surface, which shows how much energy is stored in those configurations. The system will try to minimize its potential energy, so a second pass with the Schrödinger equation gives the motion of the nuclei when subjected to the forces described by the potential energy surface. Because these forces are more complicated than simple electron–electron repulsion, the equations have to be solved numerically rather than analytically—in other words, you run them over and over again while plugging in lots and lots of numbers. Thus it has taken Kuppermann nearly 25 years to progress from hydrogen exchange ($H + H_2 \rightarrow H_2 + H$) to deuterium exchange ($D + H_2 \rightarrow H + HD$), which he and Member of the Technical Staff Yi-Shuen Mark Wu (PhD '92) published in 1993. (Deuterium, or heavy hydrogen, has a neutron as well as a proton in its nucleus.) Why bother? Says Kuppermann, "If you can do something very, very well—even a system as simple as $D + H_2$—it will tell you things that can be applied to much larger systems."

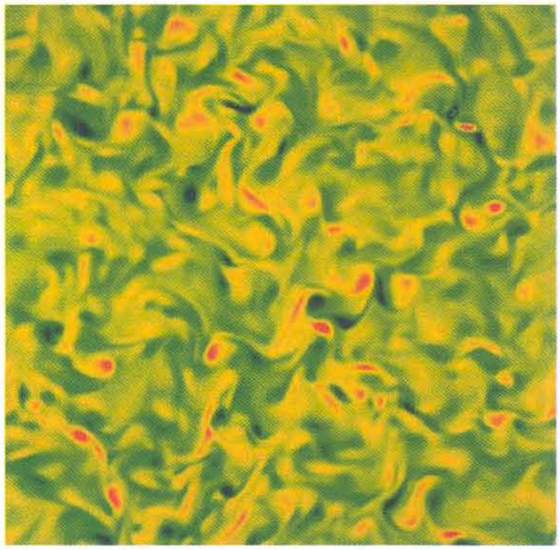
Both the $H + H_2$ and the $D + H_2$ systems, in fact, revealed the existence of a "geometrical phase effect," which has to do with the path that the atoms take to get to their final positions. Like those awkward moments where someone comes toward you, hand outstretched, and then shakes hands with the person behind you, atom A can approach atom B and then veer off at the last moment to embrace atom C instead. Experiments on the $D + H_2$ system by Richard Zare at Stanford have since confirmed that the effect does indeed exist, and can influence the reaction's outcome by a factor of 10.

By 1995, Kuppermann and Wu had moved on to $H + D_2 \rightarrow HD + D$—a chemically trivial but computationally intense leap. This brought a fresh surprise: something called dynamic resonance. In the transition state midway through the reaction, the three atoms are all stuck together: H–D–D. Normally, one bond shrinks while the other one stretches, and the middle D slides from the end D to the H like a bead on a bit of string. Two of the atoms are snuggling and one is estranged—from there it's but a short step to splitsville. But the bonds can also stretch and shrink in unison, like a bodybuilder flexing the pecs. This vibrational mode is unreactive, because the two outlying atoms are equidistant from the center at all times, and thus one is no more likely to leave than the other. The two modes have different energies, so if the incoming hydrogen happens to have nearly the energy of the symmetric mode, the transition state will tend to vibrate in that mode. The effect is to make the threesome "stickier," and the complex lasts for 150 femtoseconds instead of the usual 10 before the transition state slips into the asymmetric mode and falls apart. For reasons too esoteric to go into, this actually increases the probability of the reaction occurring at that energy.

And now, Stephen Walch (PhD '77) of NASA Ames has finished calculating the potential energy surface for the reaction of oxygen and hydrogen in collaboration with Kuppermann's group. Grad student Stephanie Rogers is using the Paragon to explore the reaction on that surface, a project Kuppermann estimates will take six months to a year to finish. Knocking a molecule of H_2 and an atom of O together to give OH and H is an important combustion reaction in, among other things, the propulsion of the Space Shuttle. However, this reaction doesn't occur just like that. Some encounters are ricochets, some are direct hits; the outcome depends on who hit what where, how hard, and at what angle. Like the chaos of a disturbed anthill, most of this activity goes

Right: A slice through a 512 × 512 × 512 three-dimensional "turbulence in a box" simulation by Mei-Jiau Huang (PhD '94). The colors indicate vorticity, with blue being the lowest and red the highest.

Below: Henderson's three-dimensional, spectral-element simulation of the wake behind a cylinder (the black bar at left). The sequence of images shows a slice through the wake along the cylinder's axis. The colors represent velocities, with yellow close to zero. As the color changes from green to blue, the fluid is moving with increasing speed into the plane of the page; similarly, red and white indicate the fluid is coming out of the page. At low speeds (top panel), the wake is very regular, but at a critical speed (second panel), a mode of instability (the scalloped pattern) develops. As the flow speed increases further (third panel), additional modes of instability with shorter wavelengths can be seen, particularly in the two sets of bands closest to the cylinder. At still larger flow speeds (bottom panel) there are so many unstable modes that the bands themselves begin to break up.



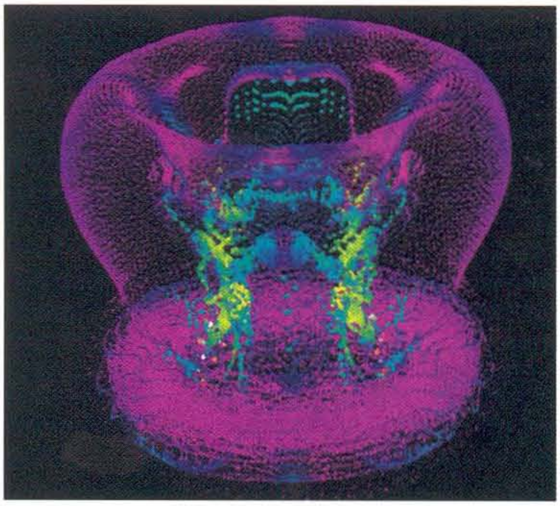
nowhere or is counterproductive. But the calculation has to include every possible collision and breakup mode, because it's the differing rates at which all these processes operate that determine the rate of the reaction. This puts such a load on the computer, says Kuppermann, that "the brute-force approach doesn't work. These are problems you can just barely solve by being clever."

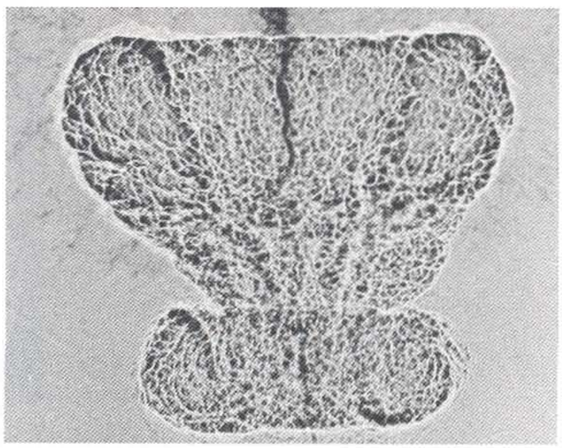
Kuppermann plans to model a number of three-atom reactions to see what the range of behavior is. And, with postdoc Desheng Wang, he's preparing for the arrival of the next generation of concurrent computers by developing methods to do the same with four-atom reactions. This, he hopes, will provide enough information to create a generalized model that will predict the course of almost any reaction, because "most chemistry, including biology—even when you're dealing with proteins—really boils down to three or four atomic centers with one or two bonds breaking and forming at once. The rest is mainly flexible scaffolding."

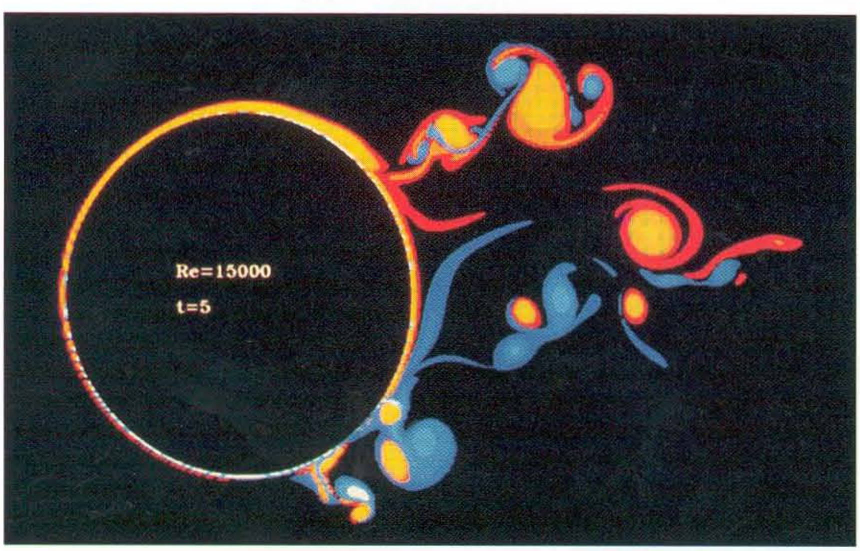
Computational fluid dynamics is another field where the problems can barely be solved by being clever. In the late 1940s, John von Neumann, the father of the modern computer, proposed simulating turbulent flows on a three-dimensional grid. He noted, however, that 20 points on a side would require a whopping 8,000 data points, and that no machine "in sight for several years to come" could do the job. This "turbulence-in-a-box" school of modeling grew with the computers, and today's Delta or the Paragon can handle a 512 × 512 × 512 grid—134 million data points. But the faster an object moves, the smaller the tiniest eddies behind it are. A 512^3 grid suffices for walking-speed turbulence, says Professor of Aeronautics Anthony Leonard (BS '59), but the smallest swirls around a golf ball in flight would fall through the mesh. To calculate the microscopic moil around a jet aircraft in full detail, Leonard says we need a grid that's 2,500 points on a side and a machine he reckons is about 15 years away.

But you can wring more detail from existing machines by using cleverer approaches. One, called the spectral-element method, was developed by Senior Research Fellow Ron Henderson while a graduate student at Princeton. The method approximates the equations of fluid flow with polynomials and can simulate turbulent flows on irregular grids. This makes it possible to handle a much wider range of scales than would a uniform grid with the same number of grid points. The method is being widely used for very detailed simulations of turbulent flow past simple bodies like cylinders, but the flow speeds are still modest—Henderson's simulation of a fully turbulent flow at a speed equivalent to a six-inch putt is the best anyone has yet managed.

Another approach, called the vortex-particle method, is also making fast progress. Here each grid point encodes vorticity—the rate and direction of spin of a fluid particle at that point. Thus the grid only covers the regions where something is happening, which saves the enormous amounts of computation that would be wasted on the smooth-flowing parts of the flow. Even so, there's room for improvement. The original method, developed in the 1970s by Leonard while at NASA's Ames Research Laboratory, calculated every particle's effect on every other particle. The number of computations increased with the square of the number of particles, severely limiting the grid size. But in the late 1980s, then-grad student John Salmon (PhD '91) figured out how to make the calculation grow more slowly by reorganizing it into groups of distant elements that didn't affect each other strongly, and could therefore be approximated, and pairs of nearby elements that still had to be calculated in detail. (Salmon actually studied how sheets and clumps of galaxies coalesce under their own gravitation, a conceptually similar problem because every galaxy attracts every other galaxy.) Leonard and Grégoire Winckelmans (MS '85, PhD







Right: In this three-dimensional vortex-particle simulation by Salmon and Winckelmans (top), the colors represent the strength of the vorticity. The simulation bears a striking resemblence to Haas and Sturtevant's shadowgraph of a Mach 1.25 shock wave striking a helium bubble (bottom). Below: Shiels and Koumoutsakos's two-dimensional vortex-particle simulation of the wake from an oscillating cylinder, which appears in two dimensions as a circle. Red and yellow show increasingly negative (clockwise) vorticity; blue and white show increasingly positive (counter-clockwise) vorticity.

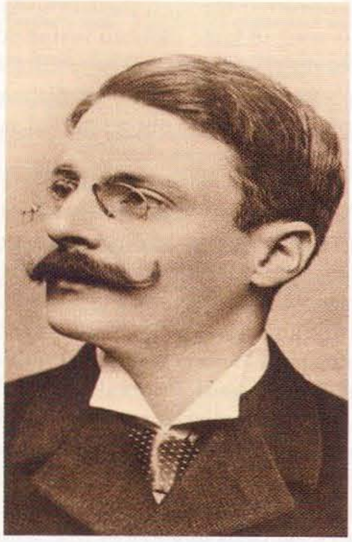
'89) of the Catholic University of Louvain, Belgium, have adapted Salmon's method to three-dimensional turbulence.

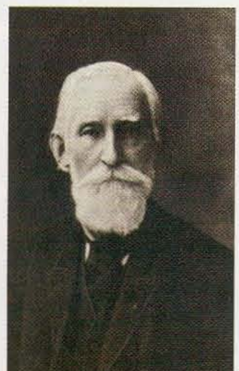
Salmon and Winckelmans have done a three-dimensional computation in which they immersed a sphere in a smooth flow and calculated the infinitely thin initial vorticity layer around the sphere. Then they deleted the sphere to see what the vorticity layer, left suddenly unsupported, would do. This simulation bears a strong mathematical resemblance to an experiment done in the early '80s by then–grad student Jean-François Luc Haas (MS '76, PhD '84) and Bradford Sturtevant (MS '56, PhD '60), Liepmann Professor of Aeronautics, in which a gas bubble was hit by a shock wave transmitted through the surrounding gas—a process used in laser-induced fusion studies to compress the plasma fuel. The results were strikingly similar—the bubble collapsed on itself and imploded, spitting out a braided vortex ring on the side opposite the shock wave.

And last year, armed with a similar, two-dimensional method, grad student Douglas Shiels (BS '93, MS '94) and postdoc Petros Koumoutsakos (MS '88, PhD '93) revisited a phenomenon discovered by then–grad student Phillip Tokumaru (MS '86, PhD '91) and Northrop Professor of Aeronautics and Professor of Applied Physics Paul Dimotakis (BS '68, MS '69, PhD '73): that a cylinder strongly oscillating around its axis can have a greatly reduced drag (*E&S*, Winter 1990). The work, done on the Cray T3D at JPL, has shown that the cylinder's wake is dominated by pairs of vortices that spin in opposite directions. The vortices form spontaneously from instabilities in the boundary layer surrounding the cylinder, and radically alter the wake's flow.

Such discoveries fall naturally out of realistic models. The earlier history of computational science (as opposed to computer science, which is the construction and programming of the machines on which computational science is done) had been a lot like the anonymous broadside posted in the hallway outside more than one Caltech lab. This two-column "Guide to Effective Scientific Communication," which purports to translate phrases commonly found in the literature, lists the English equivalent of "Correct to within an order of magnitude" as "Wrong!" In order to make a model simple enough to actually run, the modelers would have to guess which details could safely be eliminated. "As you simplify the model," says CACR director Paul Messina, "you start throwing out phenomena. Then you wind up not matching the experimental results, because things were left out. Kuppermann's work is an example—who would have thought that a geometrical phase effect would be important?" But now that computers are beginning to reach a level of power where no detail is too small to include, models can be made that replicate the real-world data exactly. And once a model does that reliably, you can begin to take any unexpected results it generates as manifestations of fundamentally new phenomena that are being revealed by, and that are not artifacts of, the model. In the words of Steve Koonin, now wearing his provost's hat, "Supercomputing has been a great enabler for science all across campus. You just can't do science without it." □

The emergence of number theory as a by-product of numerology is analogous to that of another great science, astronomy, which owes its origins to a pseudoscience, astrology.





A Centennial History of the Prime Number Theorem

By Tom M. Apostol

The prime number theorem was proved in 1896 by Charles-Jean de la Vallée Poussin and Jacques Salomon Hadamard, working independently of each other. Both de la Vallée Poussin (top left) and Hadamard (top right) built on the legacy of work by many previous mathematicians, including (in clockwise order from Hadamard) Carl Friedrich Gauss, Pafnuty Lvovich Chebyshev, Georg Friedrich Bernhard Riemann, and Leonhard Euler.

This year mathematicians all over the world are observing the 100th anniversary of the first proof of the prime number theorem, a landmark discovery in the history of mathematics. This famous theorem tells us what proportion of the positive integers are prime numbers. (The positive integers are the counting numbers: 1, 2, 3, 4, 5, and so on; a prime number is a positive integer greater than 1 that is divisible only by itself and by 1.) The prime number theorem is part of a branch of mathematics called number theory, which deals with properties of all the integers—positive, negative, and zero. The first proof was obtained independently in 1896 by two young mathematicians—Frenchman Jacques Salomon Hadamard, age 31, and Belgian Charles-Jean de la Vallée Poussin, age 30. Theirs was a remarkable achievement, the culmination of a century of efforts by an international collection of celebrated mathematicians.

The positive integers were undoubtedly humanity's first mathematical creation. It is hardly possible to imagine human beings without the ability to count, at least within a limited range. Numbers were used for record-keeping and commercial transactions for centuries before anyone thought of speculating about the nature and properties of the numbers themselves. This curiosity developed into a sort of number-mysticism or numerology, and even today numbers such as 3, 7, 11, and 13 are considered omens of good or bad luck. The emergence of number theory as a by-product of numerology is analogous to that of another great science, astronomy, which owes its origins to a pseudoscience, astrology.

The first scientific approach to the study of the integers, that is, the true origin of number theory (still intermixed with a good deal of number mysticism), is generally attributed to the ancient Greeks. Around 600 B.C. Pythagoras and his disciples classified the positive integers in various ways; examples include

Even numbers:
2, 4, 6, 8, 10, 12, 14, 16, 18, 20, ...
Odd numbers:
1, 3, 5, 7, 9, 11, 13, 15, 17, 19, ...
Prime numbers:
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, ...
Composite numbers:
4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, ...

Numbers that aren't prime are composite, except that the number 1 is neither prime nor composite. The Pythagoreans also linked numbers with geometry and with music—the latter by discovering the relationship between the length of a plucked string and its harmonic properties. (For example, a string that is one-half as long as another string under equal tension will sound an octave higher.)

The first systematic study of prime numbers appeared around 300 B.C., when Euclid wrote his *Elements*, a remarkable collection of 13 books that contained much of the mathematics known at that time. Books 7, 8, and 9 deal with properties of the integers and contain the early beginnings of number theory, a body of knowledge that has flourished ever since. It has grown into a vast and beautiful branch of mathematics that for centuries has attracted the attention of both amateur and professional mathematicians. It attracts amateurs because most of its problems are simple to state and easy to understand. It attracts professionals because these same problems are often difficult to solve, and reveal relations of great depth and elegance.

Prime numbers derive their importance from a theorem, called the fundamental theorem of arithmetic, which was first enunciated by the German mathematician Carl Friedrich Gauss. This theorem states that every integer n greater than 1 can

Courtesy of The Huntington Library, San Marino, California

Very little is known of the life of Euclid, who flourished around 300 B.C. and whose 13-volume Elements distills most of the mathematical wisdom of his day. He founded a school at Alexandria, in Egypt, and was a personal tutor to King Ptolemy I. When asked by Ptolemy if there was no shorter way to learn geometry than reading all 13 books, Euclid is said to have replied, "There is no royal road to geometry."

The largest known prime, as of September 3, 1996, is $2^{1,257,787} - 1$; it contains 378,632 digits, which, if printed in the *Los Angeles Times*, would fill 12 pages.

be factored as a product of prime numbers in one and only one way, if one ignores the order of the factors. For example, the number 12 has three different factorizations (1×12, 2×6, and 3×4) in which at least one factor is composite, but only one factorization ($2 \times 2 \times 3$) in which all the factors are primes. The fundamental theorem shows that the prime numbers are the building blocks of the mathematical world, just as the fundamental particles of physics are the building blocks of the physical world.

The fact that every positive integer is a product of prime numbers was known in Euclid's time, but the *uniqueness* of that factorization was first explicitly stated by Gauss in 1801 in his *Disquisitiones Arithmeticae*, one of the earliest books devoted exclusively to number theory. Gauss deduced the fundamental theorem from Proposition 30 in Book 7 of Euclid's *Elements*, which states that if a prime divides a product of two integers, then that prime must also divide at least one of the factors. Gauss, who is hailed as the greatest pure mathematician of all time, made enormous contributions to other branches of mathematics, as well as to astronomy and physics, but he considered the *Disquisitiones* to be his greatest work.

Proposition 20 in Book 9 of the *Elements* states that there are infinitely many primes. Many proofs of this theorem exist, but Euclid's original proof is the most elegant. It is a proof by contradiction that goes as follows. Suppose that there were only a finite number of primes, and let P denote their product. Look at the number $Q = P + 1$. Since Q is greater than 1 it must be divisible by some prime occurring in the product P, because P contains *all* the primes. However, such a prime would also divide their difference $Q - P$, because whenever two numbers (say, 35 and 20) have a common factor, their difference (in this case 15) also has that factor (5, in this example). But in the case of Q and P this is impossible, because $Q - P$ is equal to 1 and no prime divides 1.

A more sophisticated proof of Euclid's theorem was given many centuries later by the Swiss mathematician Leonhard Euler. In 1737, Euler showed that by adding the reciprocals of successive prime numbers you can attain a sum greater than any prescribed number. (This is written symbolically as

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\cdots=\infty$$

where the ∞ represents infinity, and the $\cdots$ indicates that the sum is to be continued indefinitely.) Therefore, there must be infinitely many primes—otherwise the sum would be finite. Mathematicians describe this by saying that the infinite series of reciprocals of the primes diverges.

A question that presents itself at the very threshold of mathematics is this: How are the primes distributed among the positive integers? Detailed examination of a table of primes reveals great irregularities in their distribution.

Some primes are very close together, like 3 and 5; 11 and 13; 17 and 19; or 59 and 61—these are examples of pairs of twin primes, primes that differ by 2. Twin primes keep recurring as far as we can see, as the table below shows.

x	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}	10^{11}
number of twin prime pairs less than x	35	205	1,224	8,169	58,980	440,312	3,424,506	27,412,679	224,376,048

Leonhard Euler (1703–1783) lost the use of his right eye to overwork when only 28. When a friend attempted to commiserate, Euler is said to have remarked, "I shall now have fewer distractions." A cataract robbed him of his other eye at age 51, but his work continued undiminished with the assistance of his sons, an excellent memory, and a remarkable knack for mental computation.

The largest known pair of twin primes is $242{,}206{,}083 \times 2^{38{,}880}$ plus and minus 1. (The largest known prime, as of September 3, 1996, is $2^{1{,}257{,}787}-1$; it contains 378,632 digits, which, if printed in the *Los Angeles Times*, would fill 12 pages.) It would appear that there are infinitely many pairs of twin primes, but to date no one knows whether or not this is true. In 1919, the Norwegian mathematician Viggo Brun tried to use Euler's method to prove that there are infinitely many pairs of twin primes, but instead he found that the sum of the reciprocals of all the twin primes is not divergent but has a finite sum, now called Brun's constant B:

$$B=\left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\left(\frac{1}{17}+\frac{1}{19}\right)+\cdots$$

Its value to five decimal places is 1.90216, which gives you some idea of the scarcity of twin primes, even if there are infinitely many of them.

But there are also large gaps between consecutive primes. For example, there are no primes between 20,831,323 and 20,831,533. In fact, it is easy to prove that arbitrarily large gaps must eventually exist between primes. Choose any integer n greater than 1 and look at the set of $n-1$ consecutive numbers $n!+2, n!+3, n!+4, \ldots, n!+n$. (The exclamation mark, called a factorial, indicates that the n in $n!$ is to be multiplied by all the positive integers less than it—for example, $5! = 5\times4\times3\times2\times1$.) All of the numbers in this set are composite ($n!+2$ is divisible by 2, $n!+3$ by 3, $n!+4$ by 4, etc.), and since n can be as large as you please, this means that there must eventually be arbitrarily long strings of consecutive composite numbers, and hence arbitrarily large gaps between consecutive primes. So we see that consecutive primes can be very close together, or very far apart. This irregular distribution is one of the difficulties inherent in the study of primes. Another difficulty is that no simple formula exists for producing all the primes.

Euclid's theorem on the infinitude of primes can be stated another way. Arrange the primes in increasing order and let p_n denote the nth prime, so that $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \ldots$. We can regard p_n as a function of n. Euclid's theorem states that p_n becomes as big as you want it to be as n increases without bound. Mathematicians describe this by saying that p_n tends to infinity as n tends to infinity; in symbols, $p_n \to \infty$ as $n \to \infty$. How fast does p_n go to infinity? Since not all positive integers are primes, p_n must grow more rapidly than n. But what is the actual growth rate of p_n for large n?

The prime number theorem—the title character of this tale—answers this question. The prime number theorem states that, for very large n, p_n is about the size of $n \log n$, where $\log n$ is the natural logarithm of n (the logarithm of n to the base e, sometimes written as $\log_e n$, or as $\ln n$; $e = 2.71828\ldots$). This is expressed symbolically as follows:

$$p_n \sim n \log n \quad \text{as } n \to \infty.$$

The symbol ~ is read as "is asymptotically equal to," which means that you can make the ratio $\frac{p_n}{n \log n}$ get as close to 1 as you like by pushing n farther and farther out toward infinity.

One can also turn the growth-rate question on its head and ask, how many primes are there that are less than or equal to any given positive value of x? This number depends on x and is denoted by $\pi(x)$. If a table of primes is available, $\pi(x)$ can be determined by simply counting the number of primes up to x. But don't panic if you can't find a table, or if the one you have isn't big enough—a second, logically equivalent version of the prime number theorem states that $\pi(x)$ is asymptotically equal to x divided by the natural logarithm of x. In symbols this is written as follows:

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty.$$

Again, this means that the ratio $\pi(x)/\frac{x}{\log x}$ approaches the limit 1 as x goes to infinity.

People began to speculate about the distribution of primes after extended tables of primes appeared in the 17th and 18th centuries. In 1791, the 14-year old Gauss examined a table (compiled by Johann Heinrich Lambert in 1770) that listed all the prime numbers less than 102,000. Gauss counted the primes in blocks of 100, 1,000, and 10,000 consecutive integers, and made a note in his diary that the function $1/\log n$ was a good approximation of the average density of distribution of primes in the interval from 2 to n. He offered no proof, only the numerical evidence he obtained by looking at the table. In 1797, when Georg Freiherr von Vega published an extended table of primes up to 400,031, Gauss substantiated his hypothesis further, and he kept returning to this work as new tables of primes appeared. Many years later, in 1849, he communicated his observations in a letter to the astronomer Johann Franz Encke, and the results were published posthumously in 1862. (Gauss died in 1855.) Based on tables listing primes up to 3 million, Gauss observed that $\pi(x)$ is closely approximated by the integral of the density function, $\int_2^x dn/\log n$. (This is called the logarithmic integral and is denoted by $\mathrm{Li}(x)$.) The table below is adapted from his letter to Encke. It shows $\pi(x)$ and $\mathrm{Li}(x)$ for x between $^1/_2$ million and 3 million. The agreement between $\pi(x)$ and $\mathrm{Li}(x)$ is striking—the error in each approximation is only about one-tenth of one percent.

Carl Friedrich Gauss (1777–1855) was a child prodigy who, he once said, "could count before he could talk." Gauss reveled in computations for their own sake. When Guiseppe Piazzi of the Palermo Observatory discovered the first asteroid, Ceres, on January 1, 1801, only to lose it again 40 days later as it appeared to approach the sun, Gauss sat himself down and computed its orbit from three of Piazzi's observations. Ceres was rediscovered within a year's time by several astronomers using Gauss's calculations.

x	$\pi(x)$	$\mathrm{Li}(x)$	% error
500,000	41,556	41,604.4	0.12
1,000,000	78,501	78,627.5	0.16
1,500,000	114,112	114,263.1	0.13
2,000,000	148,883	149,054.8	0.11
2,500,000	183,016	183,245.0	0.12
3,000,000	216,745	216,970.6	0.10

The first textbook devoted entirely to number theory was published in 1798 by a Frenchman, Adrien Marie Legendre. In the second edition of this text, published in 1808, Legendre also considered the problem of the distribution of primes. An appendix page from Legendre's second edition displays approximations to $\pi(x)$ for various x up to a million. Legendre asserted that $\pi(x)$ is closely approximated by the quotient

$$\frac{x}{\log x - 1.08366}.$$

On a later page Legendre states that $\pi(x)$ is approximately equal to the quotient

$$\frac{x}{\log x - A(x)}$$

where $A(x)$ is an unspecified function of x that approaches 1.08366 as x goes to infinity. It seems likely that Legendre introduced the number 1.08366 to make his formula approximate $\pi(x)$ more closely.

Neither Gauss nor Legendre revealed how they arrived at the appearance of the natural logarithm in their formulas. Nor did they make any explicit statement about how good they thought these approximations were outside the range of the existing prime number tables. It is generally understood that both intended to imply that the ratio of $\pi(x)$ to each approximating formula tends to the limit 1 as x tends to infinity. An elementary calculus exercise shows that Gauss's logarithmic integral $\mathrm{Li}(x)$ is asymptotically equal to $x/\log x$, so the conjectures of Gauss and Legendre are both equivalent to the statement now known as the prime number theorem:

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty, \text{ which means } \frac{\pi(x)}{(\frac{x}{\log x})} \to 1 \text{ as } x \to \infty.$$

This is one of the most astonishing results in all of mathematics. It describes a simple relation between the primes and the natural logarithm function—which, at first glance, has nothing to do with prime numbers.

It's natural to ask what led Gauss and Legendre to use the natural logarithm in their formulas. They did not leave any written clues; they simply recorded their formulas and the supporting data. Let's see how one might be led to conjecture the prime number theorem by examining a table of primes. Below are some values of $\pi(x)$. This table lists the number of primes less than successive even powers of 10. Gauss had access to tables that only went up to 3,000,000—the last four columns have been added from more modern tables.

x	10^2	10^4	10^6	10^8	10^{10}	10^{12}	10^{14}
$\pi(x)$	25	1,229	78,498	5,761,455	455,052,512	37,607,912,018	3,204,941,750,802

What can we learn by looking at these numbers? Since we want to find how fast $\pi(x)$ grows with x, it's natural to look at the ratio $x/\pi(x)$, which compares the two quantities. The next table shows the corresponding ratios.

x	10^2	10^4	10^6	10^8	10^{10}	10^{12}	10^{14}
$\pi(x)$	25	1,229	78,498	5,761,455	455,052,512	37,607,912,018	3,204,941,750,802
$x/\pi(x)$	4.000	8.137	12.739	17.357	21.975	26.590	31.202

Notice the differences between successive entries in that row of numbers: 4.137, 4.602, 4.618, 4.618, 4.615, 4.612. In each interval where the exponent of 10 increases by 2, we see that the ratio $x/\pi(x)$ increases by an almost constant amount, 4.6, which is 2.3 times the change in the exponent of 10. But if x is expressed as a power of 10, then the exponent of x is the logarithm of x to the base 10. So the table indicates that the change in the ratio $x/\pi(x)$ is approximately equal to 2.3 times the change in $\log_{10} x$. What about this strange factor 2.3? A bright 14-year-old such as Gauss would immediately realize that the factor 2.3 is very nearly the logarithm of 10 to the base e (in fact, $\log_e 10 = 2.3026\ldots$), so

$$2.3 \log_{10} x = (\log_e 10)(\log_{10} x) = \log_e x = \log x.$$

This suggests that we compare the ratio $x/\pi(x)$ with the natural logarithm of x. Our table now looks like this:

x	10^2	10^4	10^6	10^8	10^{10}	10^{12}	10^{14}
$\pi(x)$	25	1,229	78,498	5,761,455	455,052,512	37,607,912,018	3,204,941,750,802
$x/\pi(x)$	4.000	8.137	12.739	17.357	21.975	26.590	31.202
$\log x$	4.605	9.210	13.816	18.421	23.026	27.361	32.236
$\log x/(x/\pi(x))$	1.151	1.132	1.085	1.061	1.048	1.039	1.033

Anyone looking at this last row of numbers would surely be tempted to conjecture that they approach 1 as x approaches infinity. Gauss, Legendre, and many other eminent mathematicians of the early 19th century apparently thought so, but they were unable to prove it. As far as we know, neither Gauss nor Legendre made any significant progress toward a proof.

Peter Gustav Lejeune Dirichlet (1805–1859) was deeply influenced by Gauss, and kept a much-thumbed, well-worn copy of the Disquisitiones Arithmeticae at his side at all times. Dirichlet was said to be one of the first people to actually understand this masterwork, and did much to make it accessible to others. In later years, Dirichlet became a friend of Gauss's as well as a disciple, eventually succeeding him to the professorship at Göttingen.

In the 1808 edition of his book, Legendre made another conjecture—on prime numbers in arithmetic progressions—that plays a tangential role in this story. An arithmetic progression is a sequence of numbers in which the difference between any number and its predecessor is a constant. So if the first term in the progression is h and the common difference is k, the progression consists of all numbers of the form $kn + h$ as n runs through all the nonnegative integers 0, 1, 2, 3, … . For example, if $h = 1$ and $k = 2$, the progression consists of all numbers of the form $2n + 1$; these are the odd numbers: 1, 3, 5, 7, 9, 11, 13, … . This particular progression contains infinitely many primes—in fact, it contains all of them except the prime number 2. The odd numbers, in turn, can be separated into two new progressions—those numbers of the form $4n + 1$,

$$1, 5, 9, 13, 17, 21, \ldots, 4n + 1, \ldots$$

and those of the form $4n + 3$,

$$3, 7, 11, 15, 19, 23, \ldots, 4n + 3, \ldots$$

Again, each of these progressions contains infinitely many primes.

Primes in the progression $4n + 1$ had already been investigated by the leading mathematician of the 17th century, the Frenchman Pierre de Fermat. He discovered the surprising result that every prime of the form $4n + 1$ is the sum of two squares. For example, $5 = 1^2 + 2^2$, $13 = 2^2 + 3^2$, $17 = 1^2 + 4^2$, and $29 = 2^2 + 5^2$. Although he never investigated the distribution of primes, Fermat was the first to discover really deep properties of the integers and is generally acknowledged to be the father of modern number theory.

But returning to the more general progression $kn + h$, you can see that if h and k have a common prime factor p, then each term of the progression is divisible by p and there can be no more than one prime in that progression. Legendre conjectured that there must be infinitely many primes in the progression $kn + h$ if h and k have no common prime factor, but he offered no proof.

In a celebrated paper published in 1837, the German mathematician Peter Gustav Lejeune Dirichlet proved Legendre's conjecture. Inspired by Euler's proof of the infinitude of primes, Dirichlet used an ingenious argument to show that the sum of the reciprocals of all the primes in the progression $kn + h$ diverges, which implies that there are infinitely many primes in the progression. This result is now known as Dirichlet's theorem of the infinitude of primes in arithmetic progressions.

Dirichlet's proof was an incredible accomplishment. It marked the birth of a new branch of mathematics called analytic number theory, in which problems pertaining only to the integers were attacked by going outside the realm of integers. By using concepts that depend on functions of a continuous variable, Dirichlet brought the methods of calculus to bear on problems concerning integers, and changed the way that everyone approached the prime number theorem thereafter. The ideas introduced in Dirichlet's paper laid the groundwork not only for analytic number theory, but also for algebraic number theory, in which the methods of abstract algebra are used to study the properties of the integers.

Dirichlet's proof was an incredible accomplishment. It marked the birth of a new branch of mathematics called analytic number theory, in which problems pertaining only to the integers were attacked by going outside the realm of integers.

But the first real step toward a proof of the prime number theorem itself was made in 1848 by the Russian mathematician, Pafnuty Lvovich Chebyshev. He proved that if the ratio $\pi(x)(\log x)/x$ has a limit as x goes to infinity, then this limit must equal 1. However, Chebyshev was unable to prove that this ratio actually tends to a limit. Then, in 1850, he proved that this ratio lies between 0.89 and 1.11 for all sufficiently large x. So, although he still couldn't make the ratio converge, as it were, he established that the ratio $x/\log x$ does, indeed, represent the true order of magnitude of $\pi(x)$.

Chebyshev also introduced two new functions that are somewhat easier to deal with than $\pi(x)$, and that became the focus of nearly all subsequent work on the prime number theorem. One of these functions, denoted by $\theta(x)$, is defined to be the sum of the logarithms of all the primes not exceeding x. The other function, denoted by $\psi(x)$, is the sum $\psi(x) = \theta(x) + \theta(x^{\frac{1}{2}}) + \theta(x^{\frac{1}{3}}) + \cdots + \theta(x^{\frac{1}{m}})$, where m is the smallest positive integer for which x is less than 2^m. Chebyshev then showed that proving the prime number theorem is equivalent to proving that one of the ratios $\theta(x)/x$ or $\psi(x)/x$ approaches the limit 1 as x goes to infinity. When the prime number theorem was eventually proved in 1896, the argument was based on Chebyshev's functions.

Pafnuty Lvovich Chebyshev (1821–1894) was fascinated by mechanical toys as a boy. His quest to understand machinery led to an interest in geometry and ultimately to the rest of mathematics. He returned to mechanical problems time and again throughout his career, attempting to construct a machine that would draw a straight line when a crank was turned. Although Chebyshev failed to solve this problem (a student of his eventually did), in the attempt he invented the polynomials that bear his name.

A German named Georg Friedrich Bernhard Riemann made the next significant step in 1859, in a famous 8-page paper—the only one he wrote on number theory—that was remarkable for its brevity and for the wealth of its ideas. He attacked the problem with a new method, inspired by a discovery that Euler had made in 1732.

When Euler proved Euclid's theorem on the infinitude of primes by showing that the sum of the reciprocals of all the primes diverges, his argument was based on a formula he discovered that relates the prime numbers and the sum of the sth powers of the reciprocals of all the positive integers

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \cdots .$$

This infinite series is usually written more briefly as follows, using summation notation:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} .$$

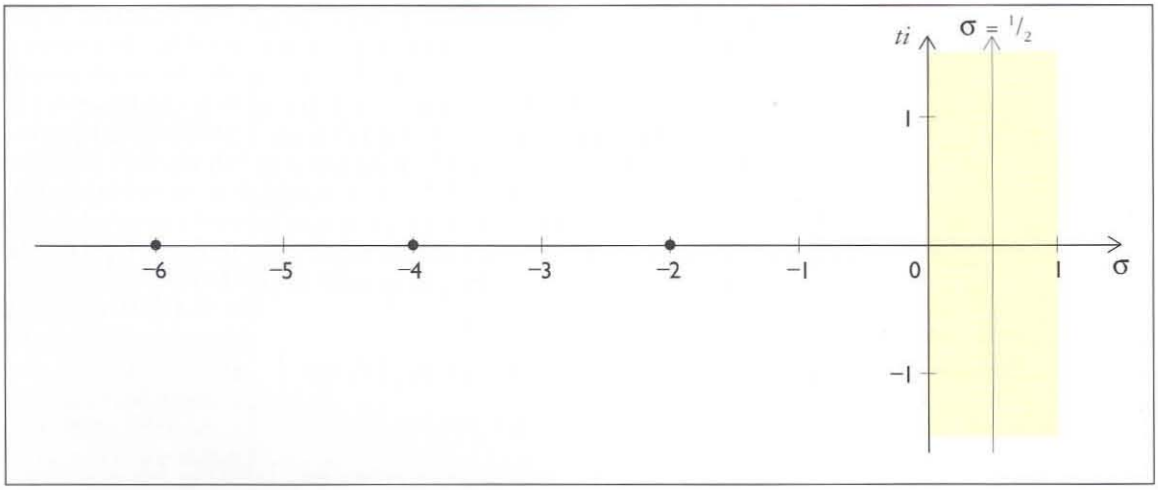
(The embellishments above and below the summation symbol $\sum$ tell us to add up all the terms of the form $1/n^s$ as n goes from 1 to infinity.) Every beginning calculus student learns about this series while studying convergence tests. The series has a finite sum (converges) if the exponent s is greater than 1. For example, when $s = 2$, Euler discovered the striking result that the sum of the series is $\pi^2/6$:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} ,$$

where π is that famous number from geometry, 3.14159…, the ratio of the circumference of any circle to its diameter. He also showed that if the squares are replaced by fourth powers the result is $\pi^4/90$, and if they are replaced by sixth powers the result is $\pi^6/945$. However, if s is less than or equal to 1, the series has no finite sum—it diverges. Euler discovered that for s greater than 1 this series could also be expressed as an infinite product extended over all the primes. This relation is usually written as follows:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \frac{p^s}{p^s - 1} .$$

Right: The complex-number plane maps all numbers of the form $\sigma + ti$. The integers lie on the σ axis; pure imaginary numbers lie on the ti axis. The trivial zeros of the Riemann zeta function are plotted; the non-trivial zeros lie somewhere in the critical strip, which is shown in yellow.



The infinite product symbol means that we are to multiply factors of this type for every prime p. For example, taking $s = 2$, we obtain a remarkable formula for expressing $\pi^2/6$ as an infinite product involving all the prime numbers:

$$\frac{\pi^2}{6} = \frac{2^2}{2^2-1} \times \frac{3^2}{3^2-1} \times \frac{5^2}{5^2-1} \times \frac{7^2}{7^2-1} \times \cdots$$

Georg Friedrich Bernhard Riemann (1826–1866) studied under Dirichlet, and upon his death succeeded him in the professorship that had once been Gauss's. He died of tuberculosis at age 39 while in Italy, on one of several trips he took to escape northern Germany's cold and damp. He borrowed a leaf from Pierre de Fermat when he wrote that the Riemann hypothesis "follow[s] from an expression for the function $\zeta(s)$ which I have not yet simplified enough to publish." Whether Riemann's hypothesis will require 357 years of effort to be settled, as Fermat's last theorem did, remains to be seen.

Euler's infinite product with the general exponent s is the analytic equivalent of the fundamental theorem of arithmetic, which, you recall, said that a positive integer can be divided into prime factors in one and only one way. The series on the left contains powers of all the positive integers, but the product on the right contains only powers of primes. Euler's product identity forms the basis for nearly all subsequent work on the distribution of primes.

Riemann suspected that Euler's product identity might hold the key to the proof of the prime number theorem, because the product on the right involves only primes. Riemann's main contribution was to replace the exponent s, which had heretofore always been a real number greater than 1, with a complex exponent that he also called s. Riemann used the notation $s = \sigma + ti$, where σ and t are real numbers, and i is the square root of -1. (Why Riemann mixed a Greek σ with a Roman t is unclear—he may have intended that it be a τ, but the printer set it as t, and t it has remained. And now, of course, it is enshrined in mathematical tradition.) Riemann then showed that the distribution of prime numbers is connected with properties of the function $\zeta(s)$, defined by the infinite series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} .$$

Because he did so much with the function $\zeta(s)$ it is now called the Riemann zeta function.

Riemann showed that the definition of the zeta function, originally valid only for σ greater than 1, could be extended (using integral calculus) to all complex values of s, and that the prime number theorem is intimately related to the location of the zeros of the zeta function, that is, those points in the complex plane for which $\zeta(s) = 0$. These zeros are of two categories, called trivial and nontrivial. The trivial zeros are the negative even integers, that is, the points $s = -2, -4, -6, \ldots$ along the negative real axis. The exact location of the nontrivial zeros is not known, except that they lie in an infinite strip of width 1 (called the critical strip) in which σ lies between 0 and 1. The critical strip is the region in the complex s plane that lies between the two vertical lines where $\sigma = 0$ and $\sigma = 1$, as shown above.

Riemann laid out an ingenious, highly creative plan for proving the prime number theorem. He showed that the prime number theorem would follow logically if one could prove that there were no zeros of the zeta function on the line where $\sigma = 1$. Unfortunately, despite his best efforts, Riemann could not carry out this crucial step in the plan. (He also conjectured a stronger statement—that all the nontrivial zeros were located on the critical strip's center line, now called the critical line, where $\sigma = 1/2$. This conjecture, called the Riemann hypothesis, is unproved to this day, and is considered to be the most famous unsolved problem in modern mathematics. If true, it has profound implications concerning the error made when $\pi(x)$ is approximated by $x/\log x$.)

Riemann, generally considered to be the intellectual successor of Gauss, came close to proving the

"I have discovered a truly remarkable proof, which this margin is too small to contain." Unfortunately, this truly remarkable proof—if indeed he had one—died with him, as he never wrote it down on anything wider.

Jacques Salomon Hadamard (1865–1963) excelled in Latin and Greek as a child, but was last or nearly last in his arithmetic classes until the seventh grade, when he fell under the influence of a good mathematics teacher. Hadamard was a relative of Alfred Dreyfus (the army officer whose conviction of treason on the flimsiest of evidence began a 12-year controversy, known as the Dreyfus Affair, that rocked France to its foundations) and helped clear his name.

Charles-Jean de la Vallée Poussin (1866–1962) studied religion and engineering successively before turning to mathematics. A lifelong resident of Louvain, Belgium, the third edition of Volume 2 of his Cours d'analyse was lost when the German army overran the city.

prime number theorem, but did not succeed. Not enough was known during Riemann's lifetime about functions of a complex variable to carry out his ideas successfully. After his death, many mathematicians went to work to develop the tools needed to execute his plan. As a consequence of this research, French mathematician Jacques Salomon Hadamard developed in 1893 an important branch of mathematics—the theory of entire functions of finite order—to handle certain classes of previously intractable functions that had bested Riemann. (These functions have since taken on a life of their own in mathematical analysis.) In 1894, Hans Carl Friedrich von Mangoldt used Hadamard's theory to justify and simplify some of the steps in Riemann's method.

By 1896 the necessary analytic tools were in hand. Working independently and almost simultaneously, Hadamard and Belgian Charles-Jean de la Vallée Poussin succeeded in proving the prime number theorem by following Riemann's strategy. In fact, de la Vallée Poussin published three papers on the subject that year—the first contains his proof of the prime number theorem, the second extends his method to obtain a prime number theorem for arithmetic progressions, and the third is on special types of primes.

Hadamard and de la Vallée Poussin each used a different method to prove that the zeta function has no zeros on the line $\sigma = 1$, the step upon which Riemann had foundered nearly 40 years earlier. Of the two proofs, Hadamard's is the simpler. In a two-page note at the end of his third paper, de la Vallée Poussin acknowledged this, and then showed how Hadamard's method could be simplified even further. In just a few lines de la Vallée Poussin showed that the lack of zeros on the line $\sigma = 1$ followed quite easily from an elementary trigonometric identity for the cosine of a double angle:

$$\cos 2\theta = 2\cos^2\theta - 1.$$

He then pointed out that this trigonometric identity can be used to shorten his original proof in the first paper by 24 pages, and that the same identity can be used to simplify the second and third papers as well.

These first proofs were later simplified by many other mathematicians, and new proofs discovered, all using sophisticated methods of calculus and complex analysis. Then, in 1949, Atle Selberg, at the Institute for Advanced Study in Princeton, and Paul Erdös, an itinerant Hungarian mathematician (who died on September 20 of this year, aged 83, while attending a conference in Warsaw), astounded the mathematical world by presenting a proof that makes no use of the Riemann zeta function or complex-function theory. But this so-called elementary proof is very intricate, and is more difficult to understand than the analytic proofs.

The prime number theorem is important, not only because it makes a fundamental, elegant statement about primes and has many applications within and beyond mathematics, but also because much new mathematics was created in the attempts to find a proof. This is typical in number theory. Some problems, very simple to state, are often extremely difficult to solve, and mathematicians working on these problems often create new areas of mathematics of independent interest. Another such example is Fermat's last theorem, which asserts that there are no positive integers x, y, z, and n satisfying the equation

$$x^n + y^n = z^n \text{ if } n \text{ is greater than or equal to 3.}$$

In 1637, Pierre de Fermat jotted that equation in the margin of his copy of Diophantus's *Arithmetica*, along with the note, "I have discovered a truly remarkable proof, which this margin is too small to contain." Unfortunately, this truly remarkable proof—if indeed he had one—died with him, as he never wrote it down on anything wider. The theorem was proved only recently—in 1994!—by Andrew Wiles of Princeton University. The proof of Fermat's last theorem has received more publicity than any other result in mathematics, but Gauss himself considered Fermat's last theorem to be of only minor importance and refused to work on it.

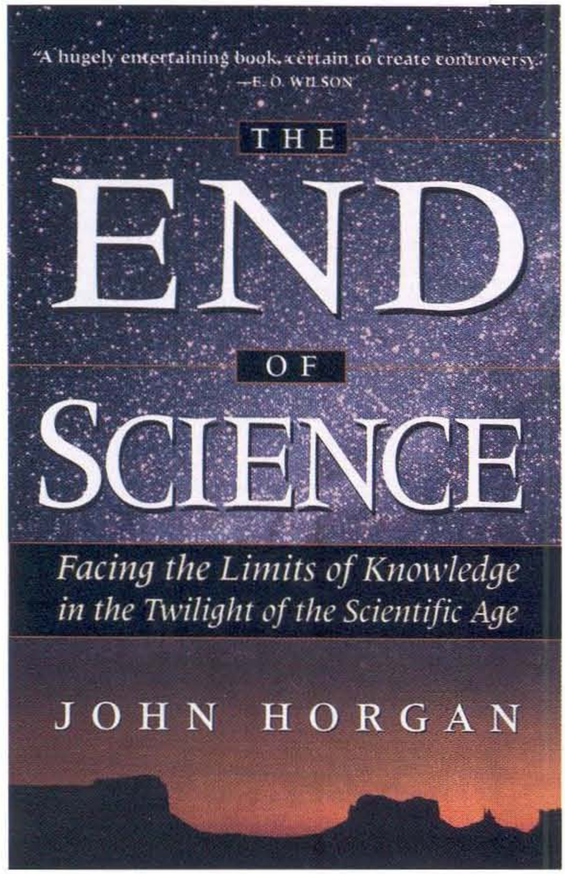
The prime number theorem and Fermat's last theorem are two outstanding examples of problems that have attracted the intellectual curiosity of many individuals but resisted efforts at solution. Repeated failure by eminent mathematicians to settle these problems by known procedures stimulates the invention of new methods, approaches, and ideas that, in time, become part of the mainstream of mathematics, and even change the way mathematicians think about their subject. This is certainly true of the prime number theorem. Early attempts to prove it stimulated the development of the theory of functions of a complex variable—a branch of mathematics that is the lifeblood of mathematical analysis. And efforts to prove Fermat's last theorem led to the development of algebraic number theory—one of the most active areas of modern mathematical research, with ramifications far beyond the Fermat equation. One unexpected application of algebraic number theory is in designing security systems for computers.

There are hundreds of unsolved problems in number theory alone. New problems arise more rapidly than the old ones are solved, and many of the old ones have remained unsolved for centuries. Our knowledge of numbers is advanced, not only by what we already know about them, but also by realizing that there is much that we do *not* know about them. Here are a few of the great unsolved problems from the realm of prime numbers:

- Is there an even number greater than 2 that cannot be written as the sum of two primes? (Goldbach's problem.)
- Is there an even number greater than 2 that cannot be written as the difference of two primes?
- Are there infinitely many twin primes?
- Are there infinitely many primes of the form $2^p - 1$, where p is prime?
- Are there infinitely many primes of the form $2^{2^n} + 1$?
- Are there infinitely many primes of the form $x^2 + 1$, where x is an integer?
- Is there always a prime between n^2 and $(n + 1)^2$ for every positive integer n?
- Is there always a prime between n^2 and $n^2 + n$ for every integer n greater than 1?

Solve any of the above, and your name, too, shall live forever in the mathematical hall of fame! □

Professor of Mathematics, Emeritus, Tom M. Apostol earned his BS in chemical engineering from the University of Washington in 1944, and his MS in mathematics in 1946. He moved south to UC Berkeley for his PhD, which he got in 1948. The southward trend continued when he arrived at Caltech as an assistant professor in 1950, after a side trip to MIT. He became an associate professor in 1956, a full professor in 1962, and emeritus in 1992. His two-volume calculus textbook, written nearly 40 years ago and known to generations of Caltech undergrads as "Tommy 1" and "Tommy 2," is still used to teach freshman math. Apostol has kept up with the times, going electronic in the 1980s as part of the team that created The Mechanical Universe... and Beyond, *a 52-episode college-level physics telecourse. Apostol is currently creator, director, and producer of* Project MATHEMATICS!, *a series of computer-animated videotapes explaining math concepts.*



Books

THE END OF SCIENCE: Facing the Limits of Knowled

BY JOHN HORGAN, HELIX BO

We seem to be coming to the end of a lot of things lately. First Francis Fukuyama proclaimed *The End of History*; then David Lindley announced *The End of Physics*. Now John Horgan goes far beyond Lindley to include all of science. What's going on? Is this just *fin-de-siècle* posturing, inspired by the approaching millenium? Or is it time for all us scientists to start thinking about our next careers?

The main body of this book is distilled from interviews with about 45 prominent scientists. These are organized into chapters, each heralding the end of one field or another: philosophy, physics, cosmology, evolutionary biology, social science, neuroscience, "chaoplexity," "limitology," and machine science. (As a chemist, I'd take encouragement from being omitted, but no: Horgan has chemistry already reaching its end 60 years ago, "when the chemist Linus Pauling showed how all chemical interactions could be understood in terms of quantum mechanics.") Horgan, a science writer for *Scientific American*, is an experienced and able interviewer, and he gets most of his subjects to reveal some of their innermost feelings about where science is and where it is going. But woven through the entire fabric is his own conviction that the glory days of science are coming to an end.

According to Horgan, science is ending in (at least) three different senses. First, all the big problems have been solved, or soon will be: there just aren't many more truly fundamental discoveries—like quantum mechanics, relativity, evolution, the Big Bang, DNA—left to be made. Only less interesting activities will remain, such as exploring the detailed implications of the basic theories, working on applied problems, and so on. Not that those are *unworthy*, exactly, but they don't fire the imagination and attract the kind of intellectual superstars who have made science what it has been over the last couple of centuries. (Horgan anticipates the obvious objection—that's what they thought 100 years ago—and responds that they didn't *all* think that, and even if it was wrong then, it doesn't mean it is wrong now.)

Second, science is approaching its intrinsic limits, in that it is posing questions that it will never be able to answer. Those who keep pushing these limits will be practicing "science in a speculative, postempirical mode that [Horgan] call[s] ironic science. Ironic science resembles literary criticism in that it offers points of view, opinions, which are, at best, interesting, which provoke further comment. But it does not converge upon the truth. It cannot achieve empirically verifiable surprises that force scientists to make substantial revisions in their basic description of reality."

Finally, science is running up against the law of diminishing returns. Experiments are becoming harder and more expensive, at the very moment that society is becoming less willing and/or less able to provide the resources needed.

in the Twilight of the Scientific Age

, 1996, 308 PAGES

Note that these are logically distinct from one another. The last two are true predictions: that we will not find ways to test empirically our latest theories (superstrings, for example); and that the currently unfavorable trend for support of science is irreversible. Both appear (to put it mildly) open to question, but who knows? The first "end," on the other hand, is a value judgment: that discovery of fundamental laws is an accomplishment that clearly stands apart from the "secondary" scientific activity of deducing the detailed consequences of those laws and applying them. Many of the scientists interviewed here do support that position; but many do not, and even when Horgan allows his interviewees to present an opposing point of view, he usually manages to do so disparagingly.

Thus physicist Leo Kadanoff is quoted approvingly—"Studying the consequences of fundamental laws is 'in a way less interesting' and 'less deep' . . . than showing that the world is lawful"—whereas Stephen Jay Gould's contrary suggestion that "[fundamental] laws do not have much explanatory power; they leave many questions unanswered" is demeaned as "ironic science in its negative capability mode." (Either Horgan completely missed Gould's point, or he feels free to redefine "ironic science" at any time to suit his rhetorical needs.) Likewise, "In denying the implication of his own ideas [that science might be ending], Chomsky may have been exhibiting just another odd spasm of self-defiance." Early on, Horgan applies the term "patronizing" to Thomas Kuhn's description of normal science as puzzle-solving, when not 40 pages later he quotes Kuhn as explicitly denying any such intent.

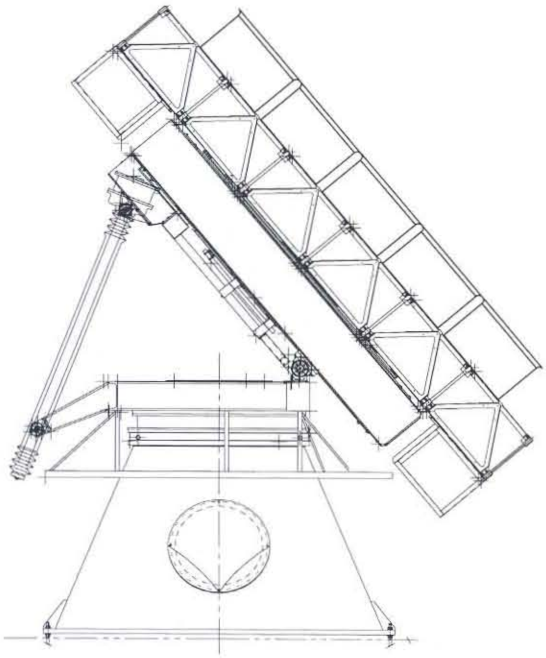
Clearly, Horgan is much more impressed by basic theories and sweeping generalizations than by details. (Inattention to detail in his *own* work shows up here and there, such as a reference to Yo-Yo Ma—born in Paris to Chinese parents—as "the great Japanese cellist.") But it's hard to see on what basis he awards his gold medals. For example, he decides "Quantum mechanics . . . was an enormous surprise. . . . The later finding that protons and neutrons are made of smaller particles called quarks was a much lesser surprise, because it merely extended quantum theory to a deeper domain. . . ." That ranking seems more than a little arbitrary to me.

Speaking of Kuhn, I noted that about half a dozen of Horgan's subjects have died since he interviewed them. That led me to look up the ages of the interviewees: the *average* is just under 65. Might not the fact that so many of them can see that *their* role in science will soon end have something to do with the prevailing mood Horgan depicts? This may be a manifestation of a form of prejudice called "era-ism": that there is something unique and special about the times in which we live and work. Horgan needs to get out and talk to a few more youngsters, who just might be able to sell him on a less depressing outlook for the field. (One hopes his pessimistic beliefs won't become so contagious that there won't be young scientists to talk to!)

The blurb on the jacket has E. O. Wilson calling this "A hugely entertaining book, certain to create controversy." Despite the mostly negative tone of this review, I expect many will agree with the first half of the sentence: Horgan *is* a skilled writer, and provides interesting (if too often unflattering) sketches of a significant number of important contemporary scientists. As for his own opinions, it all depends on whether you find outrageousness entertaining. Unfortunately, a .500 batting average is considerably better than anything Horgan approaches.

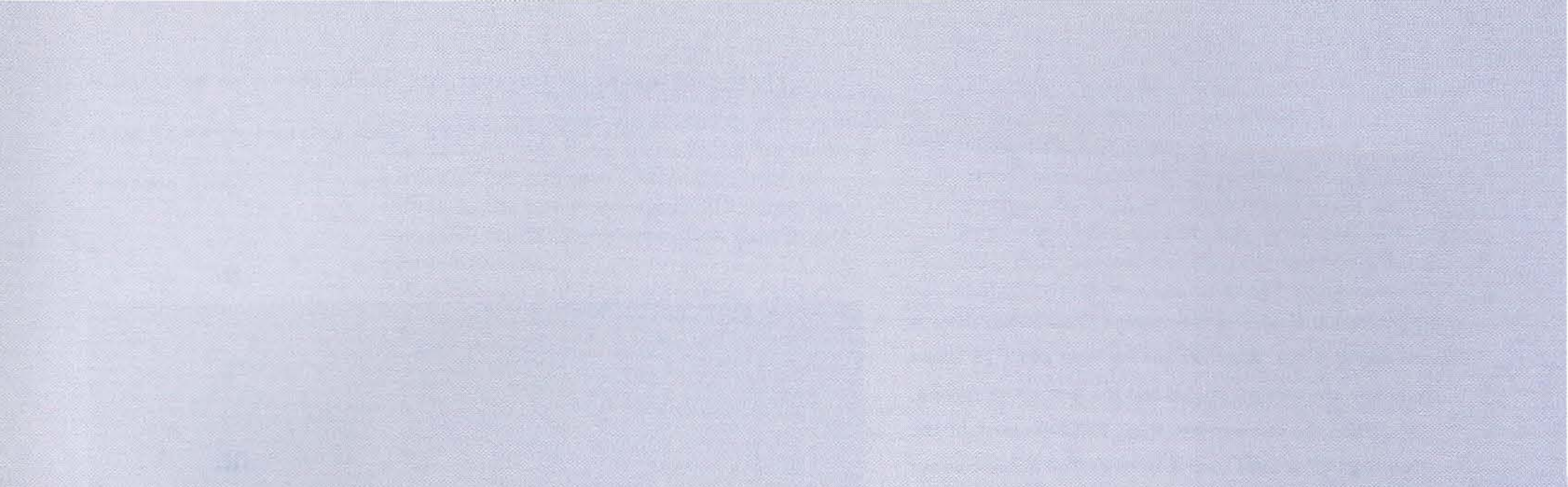
Jay A. Labinger

Chemist Jay Labinger is a member of the professional staff, administrator of the Beckman Institute, and is well under 65.



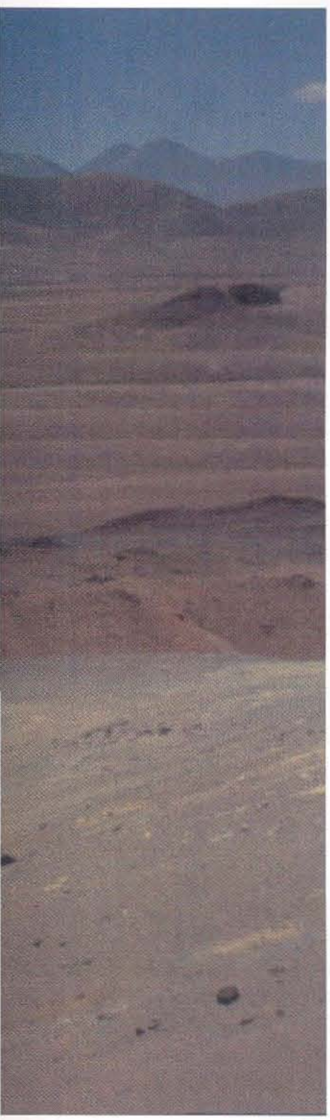
The Cosmic Background Imager (CBI) being designed and built at Caltech will make real images of the universe as it was 15 billion years ago and reveal structures that gave rise to galaxies and clusters of galaxies. The CBI's radio interferometric array consists of 13 one-meter radio telescopes mounted on a 6.5-meter platform. The diagram at left shows a side elevation view of the CBI. In addition to the usual two axes of rotation that point it at the field of interest, the array will be able to rotate about its optical axis in order to eliminate spurious instrumental signals.

Below: One of the prime sites under consideration for the CBI is at 16,000 feet (to minimize the effects of atmosphere) in the Andes Mountains in northern Chile. (Photo by Steve Padin)



Observing the Embryonic Universe

by Anthony C. S. Readhead



What did the universe look like 15 billion years ago? A new instrument we are building here at Caltech should be able to show us. Called the Cosmic Background Imager, or CBI, it will provide real images of what the universe actually looked like 300,000 years after the Big Bang. This is equivalent, when compared to a human life span of 70 years, to imaging a human embryo just a few hours after conception. The images will record the "microwave background radiation," which has been traveling through space for the last 15 billion years and presents an accurate picture of the universe during this embryonic period.

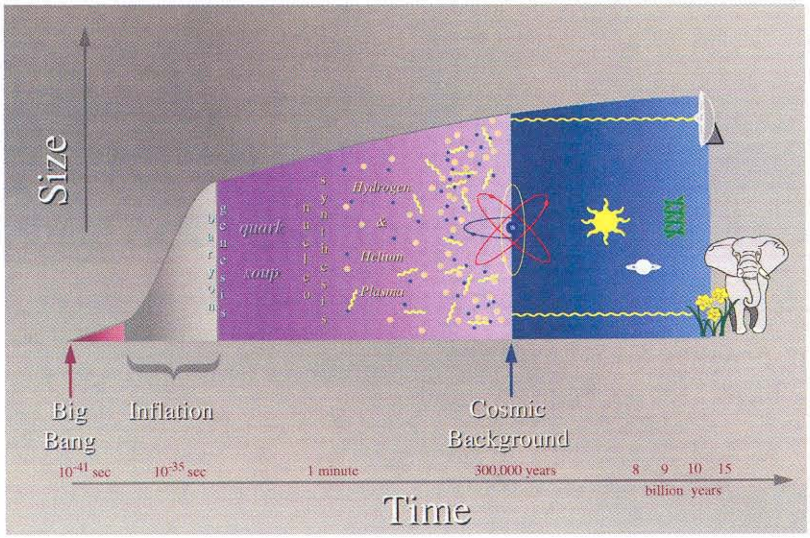
The CBI is designed to reveal minute variations—of down to one part in a million—in the background temperature of the universe as seen in different directions. These variations in temperature are believed to hold the key to understanding one of the central mysteries of modern cosmology—the formation of galaxies and hence of all the structures in the universe, including stars and planets. The temperature variations that we will see in the microwave background radiation are produced by tiny fluctuations in the density, temperature, and velocity of the universe's matter at this early period, and these tiny fluctuations are the seeds that eventually produced galaxies. A region that was slightly denser than its surroundings would slowly accrete mass by gravitational attraction of the surrounding material, and eventually become a galaxy or a cluster of galaxies.

At present we are woefully ignorant about how galaxies form, but there is no shortage of interesting and often bizarre theoretical speculation.

If we could image the microwave background radiation with sufficient precision, we would be able to measure those tiny differences in temperature and density. The physics of how those seed fluctuations grew with time in those early epochs is simple, so the only real problem is the actual measurement itself. With the measurement as a starting point it would be possible to develop a real theory of galaxy formation based on solid observations. At present we are woefully ignorant—as we will see shortly—about how galaxies form, but there is no shortage of interesting and often bizarre theoretical speculation.

In the illustration on the following page, I have taken great liberties with the time axis in order to focus on the most important cosmological events in

These tiny quantum fluctuations would then also be responsible for everything, from galaxies down to stars and planets.



The theoretical history of the universe in a nutshell (time is clearly not to scale). After the hot Big Bang about 15 billion years ago, the universe cooled and the first atoms formed. When the universe was about 300,000 years old, the "decoupling" of light from matter allowed the microwave background radiation to travel freely through the universe, where today it provides us a picture of what went on in that early epoch.

the history of the universe. I will assume, for the sake of clarity, that the universe is 15 billion years old (although estimates of the age vary between about 10 and 15 billion years). The epoch that we will image with the CBI—just 300,000 years after the Big Bang—is marked with an atom. Before this epoch the universe was so hot that the material in it was ionized; that is, the electrons were not bound to the atoms and thus there were lots of free electrons. These free electrons absorbed and scattered photons readily, so that a photon couldn't travel very far before interacting with an electron. The universe was therefore "optically thick"—any light ray could travel only a short distance before being absorbed. It's rather like being in a dense fog, where one can only see a very short distance. It's impossible, therefore, to make direct images of the universe when it was less than 300,000 years old.

However, as the universe expanded it cooled (just as the air escaping from a tire cools as it expands through a valve), and about 300,000 years after the Big Bang the temperature of the universe dropped below 4000 K. The electrons then combined with hydrogen and helium nuclei to form the first atoms. Once atoms had formed, there were no longer free electrons to absorb and scatter light, and photons could travel freely through the universe. We call this epoch—when the universe was 300,000 years old— the epoch of "decoupling." (Before this time light could not travel far without interacting with matter, and light and matter were therefore strongly coupled. However, after this time light could travel across the universe without interacting with matter, so light and matter were no longer coupled together.) Since the epoch of decoupling, the universe has been optically thin. Observations of the microwave background radiation therefore provide a direct picture of the universe at the epoch of decoupling, since these photons were present then, were stamped by the early universe's imprint, and have been unaffected by matter ever since, until they reached our telescopes.

Most of astronomy and astrophysics concentrates on events after the epoch of decoupling, since all of the structures that we are familiar with were formed later. I shall be concentrating primarily on these later epochs in this article. However, some important events from before the decoupling epoch have a bearing on this tale, so I will discuss them briefly now.

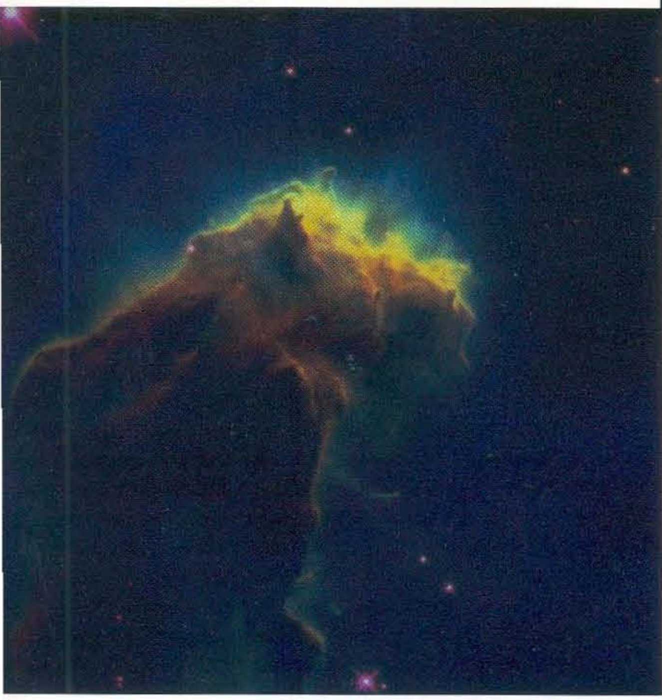
We now know that the universe began with the Big Bang, about 15 billion years ago. Our understanding of the universe's earliest moments is extremely primitive, because when the universe was about 10^{-43} seconds old, gravitational and quantum effects were comparable. At that point, the universe was so dense that gravity, which is usually much weaker than the strong nuclear force, was on an equal footing with it. In order to understand the physics of this period, we need to have a combined theory of gravity and quantum mechanics. (This has eluded theoretical physicists for the last 80 years, although significant progress is now being made.) Many cosmologists think that the universe went through a period of very rapid (exponential) expansion shortly after the Big Bang, when the universe was some 10^{-32} seconds old. This is referred to as the epoch of "inflation."

I will not discuss this epoch in any detail, except to mention that there are a number of puzzles in the standard Big Bang picture that are nicely explained by inflation. After the epoch of inflation, the universe settled down into the standard, much slower expansion that is still going on today.

Continuing with the illustration, the next important milestone occurred when the universe was between one second and a few minutes old. During this epoch the universe cooled to a temperature of a few billion degrees, allowing the nuclear reactions that built up the light elements to take place. Deuterium, helium-3, helium-4, and lithium were created out of the primordial neutrons, protons, and electrons. One of the major strengths of the Big Bang cosmology, and one of the reasons that we accept it, is that it gets the ratio of these light elements right—the abundances calculated from the Big Bang theory correspond to those we see today.

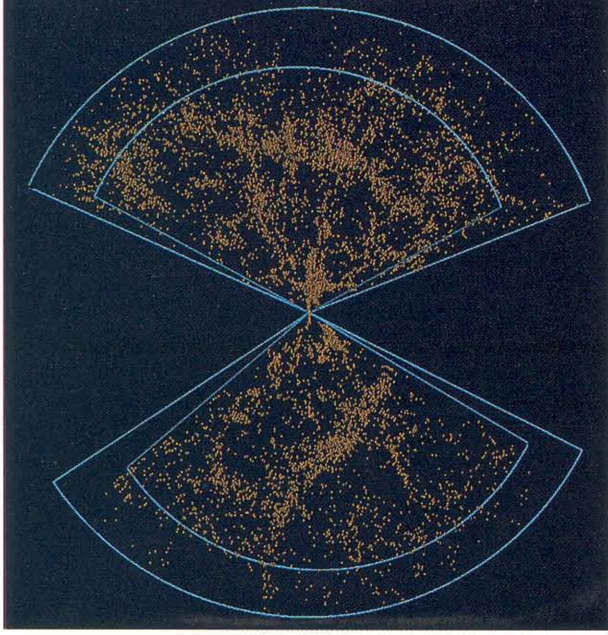
The next important epoch is the one we have already considered in some detail, namely the epoch of decoupling. This occurred, as we have seen, when the universe was about 300,000 years old. This is when the microwave background radiation that we observe today was produced. Imprinted on this radiation are tiny variations in temperature between one line of sight and another, caused by the small fluctuations in the primordial matter's density, temperature, and velocity. These were the seeds that eventually gave rise, through the effects of gravity, to galaxies. These fluctuations must have been caused by something that happened even earlier, and if we can study the microwave background radiation in enough detail, we may even be able to figure out what that something was. The most

Both of these images were captured by the Hubble Space Telescope (using the second Wide Field/Planetary Camera built at JPL). Below is a star-forming region in our own galaxy and, at right, a cluster of galaxies named Abell 2218. All of these structures originated in tiny fluctuations in primordial matter.



popular theory is that quantum fluctuations during the inflationary epoch were responsible. This would be very remarkable if true, since it would mean that small-scale quantum fluctuations in the very early universe gave rise to the largest structures that we see in the universe today. These tiny quantum fluctuations would then also be responsible for everything, from galaxies down to stars and planets.

I now want to review briefly the structures that any successful cosmological theory must explain. These range from planets and stars to galaxies, clusters of galaxies, superclusters, and voids. All of these structures appeared after the decoupling epoch. It should be clear that the formation of galaxies is an essential step in the formation of both stars and planets, since it is in galaxies that the matter density gets high enough to produce these smaller objects under the influence of gravity. A beautiful image, at left, of the star formation process was recently obtained by the Hubble Space Telescope. On larger scales we see clusters of galaxies like the one pictured above. This is an image of the galaxy cluster named Abell 2218, after the late George Abell [BS '51, MS '52, PhD '57], who worked on this cluster as a graduate student at Caltech. The deepest high-resolu-



Until the beginning of this century, only one cosmological fact was known—that the sky is dark at night.

tion optical image of the sky to date is the Hubble Space Telescope image shown at right. Here we see galaxies that span a very large range of distances—from 2 billion light-years away to about 14 billion light-years away. One of the most astonishing discoveries in modern cosmology has been that of gigantic sheets of galaxies, separated by enormous "voids" in which there are fewer galaxies than average. The most famous of these is the "Great Wall," illustrated above, discovered by Margaret Geller and John Huchra [PhD '77]. It lies at a distance of about 500 million light-years from us, and is about 500 million light-years long. Looking at images such as this, one can't help but wonder how such structures were produced by the small fluctuations that existed in the embryonic universe.

As a result of a remarkable discovery in 1963 by Maarten Schmidt (now the Francis L. Moseley Professor of Astronomy, Emeritus), we know that some of the familiar objects in the universe formed in the first billion years. Radio astronomers had discovered objects in the sky that seemed to be unresolved (that is, compact) radio sources, and they had managed to measure accurate positions of these objects. Optical astronomers found starlike objects at these positions, and hence they were called "quasars," short for "quasi-stellar objects." Both Schmidt and Jesse Greenstein (now the Lee A. DuBridge Professor of Astrophysics, Emeritus) were trying to identify the lines seen in quasar spectra when Schmidt had the brilliant insight that the lines might have been shifted far toward the red, toward longer wavelengths, and so were not appearing anywhere near where they were expected. (These lines, which are emitted by the various chemical elements in the source object, lie at characteristic wavelengths that can be used to identify those elements. That's how we know what distant stars are made of.) The redshift, caused by the Doppler effect, indicated that the quasars were moving rapidly away from us. We shall see shortly that measuring the speed of a celestial object

Shown above is the Hubble Space Telescope Deep Field image of a small region near the north celestial pole. This is the deepest high-resolution image made to date. It shows galaxies ranging from 2 billion light-years to about 14 billion light-years away.

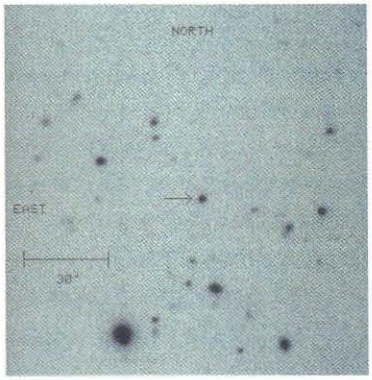
Above, left: The "Great Wall" is a sheet of clustered galaxies with voids in between. It's about 500 million light-years away and 500 million light-years across. (Courtesy of Margaret Geller and John Huchra of the Smithsonian Astrophysical Observatory.)

enables us to estimate its distance, and Schmidt concluded that the object—3C 273—was about two billion light-years away. But 3C 273 is one of the brightest radio sources in the sky, implying that similar, fainter radio sources must be even more distant—which turns out to be true.

Schmidt has continued in this line of research, and he and his collaborators hold the record for the most distant object now known: a quasar called 1247+3406, which is about 14 billion light-years away (right). This object formed just one billion years after the Big Bang. When we look at the spectrum of this distant quasar and compare it with the spectra of intervening galaxies, we find that they all show evidence of the familiar heavy chemical elements that we see in the nearby universe. You need stars to create all the elements heavier than lithium, so this tells us that when the universe was one billion years old, it had already gone through quite a bit of star formation and had already been forging the heavy elements. How the tiny fluctuations from the decoupling epoch could have given rise to stars and quasars in only one billion years is a major problem of astrophysics. We will discuss this in some detail later.

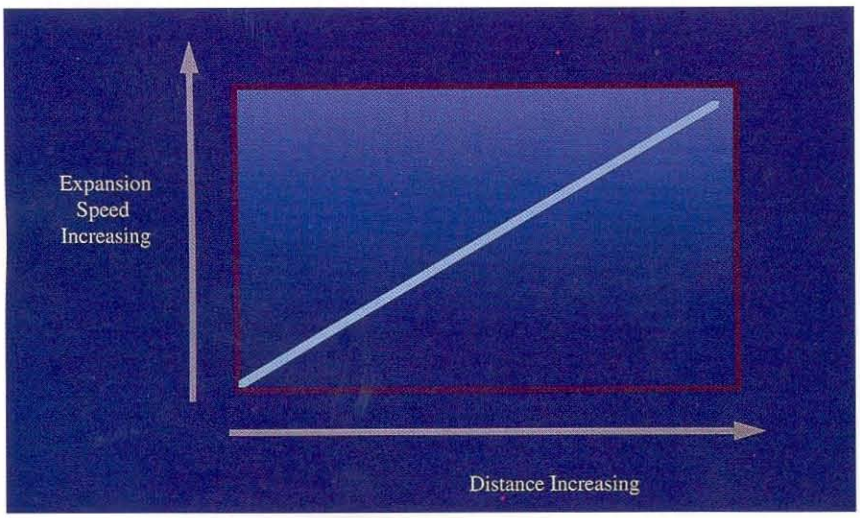
For now, I'd like to step back and talk a bit about the science of cosmology, which deals with the origin and structure of the universe. This is only now becoming a real science, because we're beginning to get enough observational data to make testable predictions.

Until the beginning of this century, only one cosmological fact was known—that the sky is dark at night. In one sense this may sound trivial, but in another sense it's actually very profound. The peculiar aspect of this phenomenon, known as Olbers' Paradox, was first noted by Johannes Kepler in 1610, and later by Edmund Halley and Jean-Philippe-Loys de Cheseaux. Why is the sky dark at night? If the universe were infinitely old, infinitely large, static, and homogeneous, then any line of sight would eventually end up on the surface of a star. The whole sky should be as bright as the sun, which it manifestly is not. Thus one of the above assumptions is wrong. Actually, Olbers' Paradox can be restated in a more fundamental way, as has been pointed out by the cosmologist George Ellis. The real cosmological question is not why is the sky dark at night, but rather, why does the sun shine by day? Because stars shine by thermonuclear burning and because there's a limited amount of fuel available, if the universe were infinitely old, then all the thermonuclear material would have been used up infinitely long ago, unless there were some spontaneous method of generating energy. Therefore, all the stars that had ever existed would now just be dead hunks of rock. The answer to the paradox, as we now know from the Big Bang, is that the universe is not infinitely old (and neither is it static); there is still free energy around to be turned into heavier elements by thermonuclear burning.



The arrow points to the most distant object known—a quasar 14 billion light-years away.

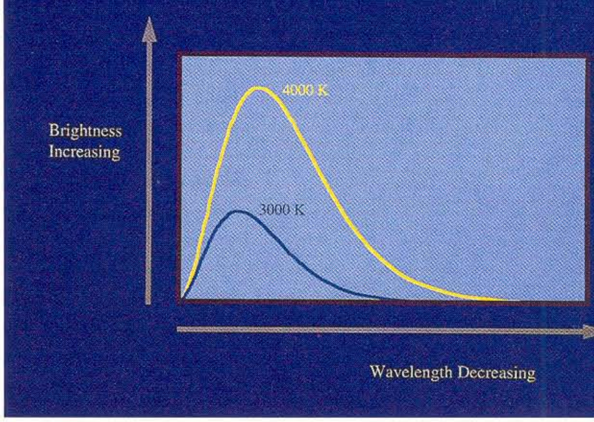
Edwin Hubble discovered that there's a linear relationship between the speed and distance of galaxies (graph at bottom), evidence that the universe is expanding.



The next important step in cosmology came in 1916, when Albert Einstein produced the general theory of relativity. This wonderful theory is the basis of all modern cosmology. But there was a flaw in it—or so Einstein thought. His theory predicted that the universe was not static, but had to be either expanding or contracting. He didn't believe this and was very worried about it, so he added a fudge factor called the cosmological constant. If the cosmological constant is positive, then it endows space with a property that acts like antigravity. With this fudge factor added, Einstein's theory does produce a static model of the universe.

However, in 1929, a few miles away from Caltech up on top of Mount Wilson, the astronomer Edwin Hubble made an astonishing discovery. Building on work that had been done earlier by Vesto Melvin Slipher and others, in which they had found that galaxies outside our small local group are all moving away from us, Hubble found that the farther away a galaxy is, the faster it's moving away from us. He found that there's actually a linear relationship between speed and distance, as shown at left. The scientific community immediately accepted this as evidence that the universe was expanding, at which point

In 1964 AT&T scientists Wilson and Penzias discovered the microwave background radiation, a relic of the radiation produced in the early universe. As the universe expands, it cools, and the peak of this radiation moves to longer wavelengths (right). It can now be detected most easily at radio wavelengths, and this is what the two scientists observed as a uniform temperature of 2.7K (Kelvin) in all directions.

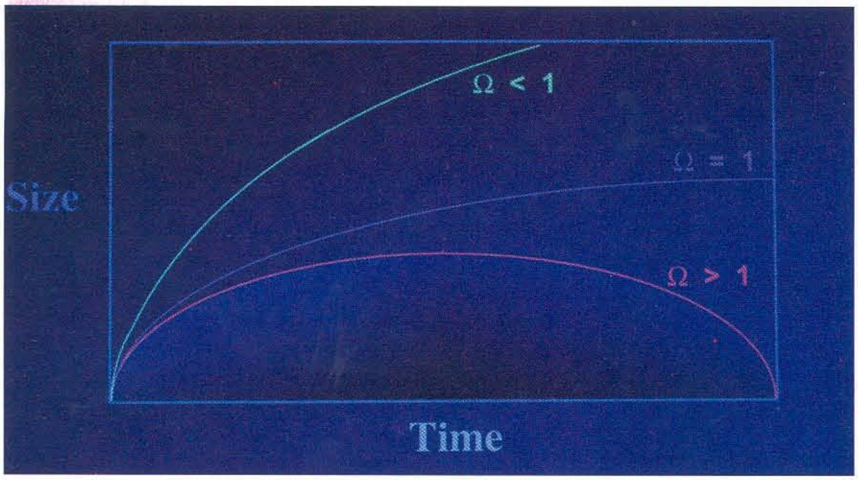


Einstein wished that he had never invented his fudge factor. Had he not, he would instead have made a remarkable prediction. It's there in his theory, but he didn't believe it. The universe is expanding, and Einstein didn't need the cosmological constant. The cosmological constant never really died, however, and has gone in and out of favor over the years. Right now it's back in favor as a very important ingredient in the inflationary theories of cosmology.

The proportionality constant in Hubble's relationship, that is, the slope of the line, is called the Hubble constant. The determination of this constant is important in cosmology, since it tells us the size and age of the universe. Much effort has been put into measurements of the Hubble constant over the last six decades, but it is still known only to within about 50 percent (which is why the age of the universe is known only to lie between 10 billion and 15 billion years). This is one of the major unsolved problems of modern astrophysics.

Another important cosmological parameter is the "density parameter," which describes the amount of matter in the universe. The expansion of the universe is slowing down, due to the effects of gravity. There is a critical density for the universe such that if the actual density is below this value, then expansion will continue forever; if the actual density equals the critical density, then the expansion will just stop in an infinite amount of time; and if the actual density exceeds the critical density then the expansion will one day be reversed, the universe will contract, and it will end in a "Big Crunch"—the opposite of the Big Bang. The density parameter (Ω) is the ratio of the actual density to the critical density, so that the three possibilities we've just discussed correspond to $\Omega < 1$, $\Omega = 1$ and $\Omega > 1$.

The amount of matter in the universe is described by the density parameter—Ω. If the critical density is less than the actual density ($\Omega < 1$), the universe will expand forever. If $\Omega = 1$, the expansion will just stop, and if $\Omega > 1$, the universe will contract in the "Big Crunch."



The three constants, or parameters, that we have discussed—the cosmolgical constant, the Hubble constant, and the density parameter—are the most important in cosmology. If present theories are to be believed, then measurements of the small temperature fluctuations seen in different directions should enable us to determine these three numbers with considerable precision.

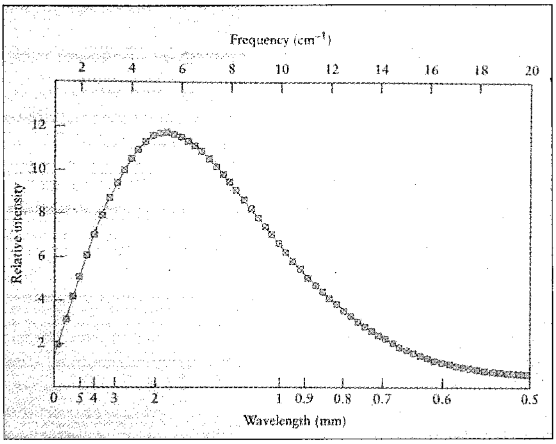
Now, as we have already seen, because the universe is expanding it must be cooling, which means that it must have been much hotter when it was young. In other words this is a "hot Big Bang" universe. This means that there must have been a lot of radiation produced in the early universe. This fact was pointed out by George Gamow in 1948. He also pointed out that the relics of this radiation would still be around today, but its wavelength would have been stretched (by the expansion of the universe) from the optical part of the spectrum all the way down into the radio part. Thus Gamow predicted the existence of the microwave background radiation based solely on a simple Big Bang model of the universe.

In 1949, two of Gamow's students, Ralph Alpher and Robert Herman, calculated that the temperature of the microwave background radiation today would be about 5K (Kelvin). Remarkably, this prediction was ignored by observational astronomers for the next 15 years!

The microwave background radiation was finally discovered accidentally in 1964 by Arno Penzias and Robert Wilson (PhD '62), while they were measuring noise inherent in radio receivers. They had made an extremely sensitive radio receiver and calibrated it very well, but wherever they looked in the sky, they found that it was a little hotter than they expected. They thought something was wrong with their receiver, but it turned out that they had discovered the microwave background radiation, which is seen at a temperature of 2.7K in all directions. This is the relic, or fossil, radiation left over from the Big Bang, and this, more than anything else, has convinced astronomers and physicists that we are indeed living in a hot Big Bang universe. In 1978, Penzias and Wilson received the Nobel Prize for their discovery.

The COBE (Cosmic Background Explorer) satellite, which was launched in 1989 to study this cosmic background radiation, has been very successful. One of its great accomplishments has been to measure the spectrum of this radiation. In the illustration at right, the curve is not a fit to the observed data points; it's actually the curve predicted by thermodynamics for an object that is a perfect emitter and absorber of radiation (blackbody curve). This well-known curve, first worked out by Max Planck in 1900, laid the foundation for quantum mechanics. It is one of the cornerstones of physics and is not in any doubt. It so happens that this is also the spectrum expected for the microwave background radiation. We see that the COBE observations fit the theoretical curve with exquisite precision.

The COBE observations of the spectrum of the microwave background radiation conform precisely to the theoretical prediction of the spectrum from a perfect absorber and emitter of radiation. This is called blackbody radiation.



With the discovery of the microwave background radiation cosmology came of age. It has taken 30 years, however, to develop the tools that will enable us to reap the benefits of that discovery.

Since the discovery of the microwave background radiation in 1964, there have been many observations of this radiation to try to detect variations in temperature along different lines of sight. One of the remarkable facts of astrophysics to emerge over the last 32 years has been the extraordinary smoothness of this fossil radiation. We now know that the temperature of the radiation varies by less than one part in 30,000 on all angular scales, once one has made the correction for the Doppler effect caused by the earth's motion relative to it. The amazing smoothness of the microwave background radiation tells us that the universe underwent a "simple" Big Bang or, to put it more explicitly, the universe is "simple" in the sense that it is both homogeneous and isotropic. In other words, if we look at the content of a small volume of the universe (that is, small compared to the universe but much larger than our local system of galaxies) and compare it to any other such volume, the average properties will be the same (that is, the universe is homogeneous); and if we look in any direction and consider a small area (one that is a small fraction of the celestial sphere) the average properties will be the same (that is, the universe is isotropic).

Now it might have been the case that the universe underwent a "compound" Big Bang, in which case the average properties of the universe from cell to cell would have been very different—the universe would not have been homogeneous and isotropic. In this case we would not have been able to draw conclusions about the large-scale structure of the universe. Cosmology would have been difficult, or most likely impossible, because what you would see would depend critically on just where you happened to be sitting.

Even a simple Big Bang contains many mysteries, but it seems that we are now poised to make critical observations that should elucidate the basic nature of the universe. With the discovery of the microwave background radiation, cosmology came of age. It has taken 30 years, however, to develop the tools that will enable us to reap the benefits of that discovery.

Before going on to discuss the CBI, I want to discuss a major mystery of astrophysics and cosmology that is fundamental to the interpretation of observations of the microwave background radiation. This is the mystery of dark matter, which has been growing more and more mysterious for the last three decades.

M31, the great nebula in Andromeda, must have far more mass than we would deduce from the starlight in this picture, because the stars at the outside of the spiral are orbiting at the same speed as those close to the center. This invisible mass is known as "dark matter."

If we look at our neighboring sister galaxy, the great nebula in Andromeda, M31 (above), we see a system that is pretty much a twin of the Milky Way. M31 is 2 million light-years away; the light that is reaching us today left the galaxy when our australopithecine forebears were stalking the plains of Africa. The galaxy seems to be about 100,000 light-years in diameter and to be fairly isolated. This is an illusion. It turns out that there is much more matter in this galaxy than is revealed through its starlight, which is all that we see in this photograph.

How can we trace mass if not through the light of stars? One simple method is to look for the gravitational effects of the mass. You can figure out the amount of mass in a galaxy by looking at the rotation speed of material around the galactic center. If the mass were distributed like the visible stars are—highly concentrated toward the center and dropping off rapidly toward the outer edges—then we would have a centrally condensed system, like the solar system. In such a system the rotation speed of the stars about the center of the galaxy would fall off with distance from the center, just as is observed for the planets in our solar system. However, in M31 we find that the rotation speed is constant with distance from the center. Even the most distant stars from the center show no evidence of a drop in orbital speed. There can be only one explanation of this—namely, that there is far more material in M31 than is revealed by starlight. Thus this galaxy must have a very substantial component of "dark matter," which is distributed differently than the stars.

M31's behavior turns out to be common. This means that galaxies must be much more massive and much larger than they appear. If you look at a cluster of galaxies, you see that the galaxies appear to be separated by vast regions of empty space that are typically about 10 galaxy diameters across—

If we could see all the dark matter associated with galaxies, we would find that the galaxies are almost touching each other and are composed primarily of dark matter. This poses one of the great mysteries of modern astrophysics—what is this dark matter made of?

about the same as the apparent distance between our galaxy and M31. This is also an illusion. If we could see all the dark matter associated with galaxies, we would find that the galaxies are almost touching each other and are composed primarily of dark matter. This poses one of the great mysteries of modern astrophysics—what is this dark matter made of?

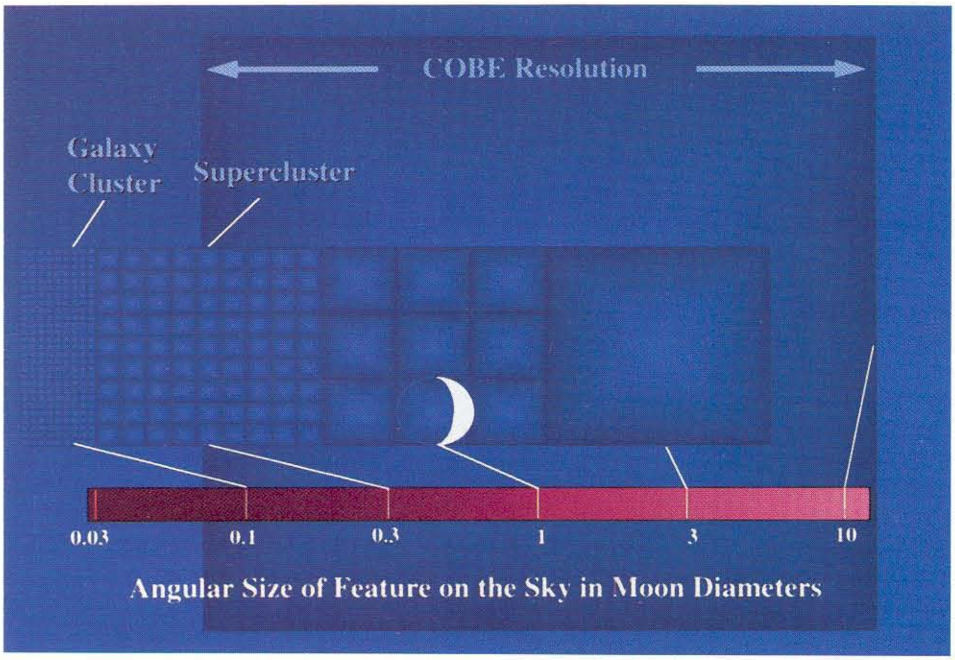
There's a lot of very complicated physics going on in nearby galaxies, which makes it very difficult to address this question. It's much easier to study by looking at the microwave background radiation, which is left over from a time when the fluctuations in the density and temperature of matter were very small. The equations governing this regime are relatively simple, and we think we understand them pretty well.

We have already made a number of observations of the microwave background radiation with the 40- and five-meter radio telescopes at Caltech's Owens Valley Radio Observatory, 250 miles north of Los Angeles. What we have found, and what has been found by a number of other astronomers, is that within about one part in 30,000 there are no variations in the temperature of the microwave background at all. This is perplexing, because it's hard to see how quasars could form within a billion years after the Big Bang, if you start out with a very smooth universe. It takes longer than that for the material to accumulate into clumps by gravitational attraction. In fact, it is now widely accepted that these results show that it is simply not possible to begin with fluctuations of only the strength inferred from the microwave background radiation observations and produce galaxies on the required time scale.

If we wish to produce galaxies on the required time scale, and quasars one billion years after the Big Bang, we must have larger seeds than the density fluctuations that we infer from the microwave background radiation. It's not easy to get around this difficulty as long as we assume that the seed fluctuations are fluctuations of ordinary matter composed of protons and neutrons. Protons and neutrons are called baryons, and wherever there are protons and neutrons there are also electrons—one for every proton. Light photons (and radio photons) interact strongly with electrons. So if the seed fluctuations consisted of baryons, then we would see fluctuations in the microwave background radiation since matter and radiation are strongly coupled. But we don't. This rules out baryons as the major constituent of the seeds. As a result, many cosmologists now believe that the seed fluctuations consist of nonbaryonic dark matter, and that it is the same nonbaryonic dark matter that constitutes the majority of the matter in galaxies.

Therefore, a significant fraction of the universe—perhaps as much as 99 percent of its matter—may consist of nonbaryonic matter. There are other lines of evidence beyond the rotational speeds of galaxies that support this conclusion. For one thing, the relative abundances of the light elements depend critically on the primordial photon-to-baryon ratio, and the observed abundances of these elements tell us that the amount of baryons in the universe is about two percent of the critical density. Yet other observations, based on the calculated gravitation between groups of galaxies, imply that the universe's total density (baryonic and nonbaryonic matter combined) is at least 20 percent of the critical density. Hence baryonic matter can account for, at most, 10 percent of the universe's total matter content. The remainder could possibly be neutrinos, but this seems unlikely, for reasons we don't have space to discuss here. It could also be material that's completely different from any of the material that we're familiar with. But it's still a bit of a reach to try to get out of a theoretical difficulty like this by saying, "Well, 90 to 99 percent of the matter in the universe is in a bizarre form that we just haven't been able to detect yet."

I now want to discuss the CBI, and tell you how it could throw light on all of the above questions. We are currently finalizing the design of this instrument, and testing prototypes of its major components, in Robinson Laboratory here on campus. There is a small local team, consisting of myself; Member of the Professional Staff Steve Padin, the chief scientist, who is responsible for all of the detailed instrument design; Tim Pearson, who is responsible for the data reduction and analysis; Martin Shepherd, who is designing the instrument's computer-control and data-acquisition systems; John Cartwright, an astronomy graduate student; and Walt Schaal and John Yamasaki, engineers working part time on the project. Off campus, Marshall Joy of the Marshall



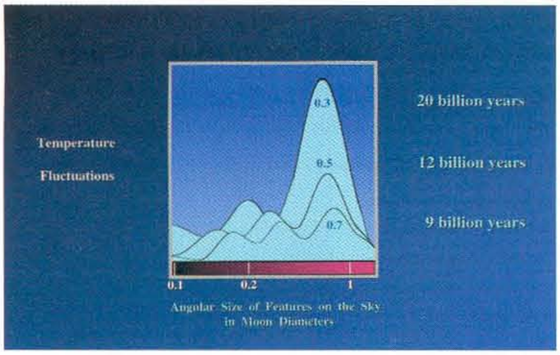
The angular size of features in the sky can be seen here in terms of moon diameters. One moon diameter projected onto the background radiation could resolve structures 300,000 light-years across, the maximum size for regions that are "causally connected." At 0.3 moon diameters, we would be able to see something the size of superclusters, and at 0.1 moon diameter, galaxy clusters. The CBI will be able to measure temperature over all these angular scales—a much higher resolution than that of COBE.

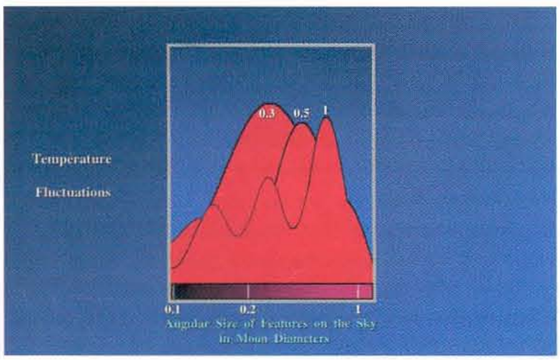
Space Flight Center is responsible for building the telescope dish molds. We're also collaborating with a group, headed by John Carlstrom of the University of Chicago and including people from the Center for Astrophysical Research in Antarctica, that is building a complementary instrument that will look at slightly larger angular scales than the CBI.

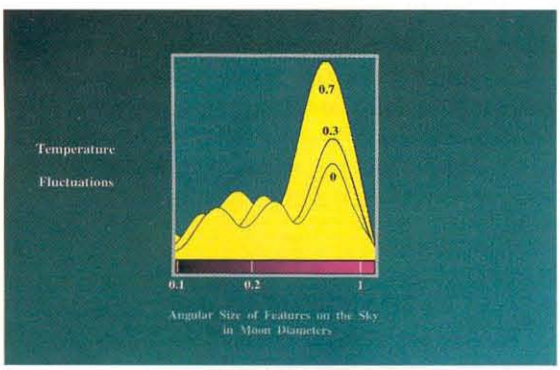
Let me give you some idea of the angular scales of interest. In the plot above and in those that follow, I'll use the size of the moon, which is half a degree, as a fiducial angular size. I don't do this for capricious reasons—it turns out that this angular size is a very important size for the microwave background. If you take the moon's diameter and project it back onto the microwave background, it turns out to be 300,000 light-years across. This is just the size of the regions of the universe that can interact with each other. Larger regions span more space than light could have traveled since the Big Bang, so they cannot interact. We say that they are not causally connected. This means that within one lunar diameter, we can expect to see structures that are collapsing under their own self-gravity. Larger-scale structures will not do this, since they have not had time to feel the effect of their own gravity. (The highest resolution of the COBE observations is much larger than the causally connected parts of the universe.) At one-third of a lunar diameter we're in the regime of superclusters, and at one-tenth in the regime of galaxy clusters. The instrument that we're building at Caltech is designed to observe over this range of angular scales.

Over the last decade there has been a tremendous flurry of theoretical activity to work out different possible cosmological models that predict the expected change in temperature of the microwave background radiation over a range of angular scales. These calculations have been carried out for various assumed values of the three parameters we discussed earlier—the Hubble constant, the density parameter, and the cosmological constant. The three plots at right provide a graphic demonstration of why the angular scale corresponding to the moon's diameter is important—nearly all of the models predict a peak in temperature fluctuations at this or a slightly smaller size. In the top plot we see theoretical predictions for three different values of the Hubble constant if we assume that $\Omega = 1$, that is, the density parameter equals the critical density. (There are no estimates of Ω—which are generally based on observations of galaxy velocities—that are significantly greater than 1.) If we could make good observations of the microwave background radiation, we should be able to discriminate clearly between values of the Hubble constant corresponding to an age of the universe of 20, 12, and 9 billion years. Similarly, if we look at models in which we fix the Hubble constant and vary the density parameter, we see that the main peak in the temperature fluctuations varies (middle). Thus, for example, if the density parameter is 0.3—corresponding to an open universe—we see that the peak occurs at about a half a lunar diameter, that is, at one quarter of a degree. Thus, measuring the position of the main peak should tell us the mean density of the universe, and hence whether the expansion will one day turn into a contraction and head for a Big Crunch. By varying the cosmological constant we produce a third family of curves (bottom). It turns out that there's enough information in these different curves that if you could just measure the temperature fluctuations with enough precision, you should be able to determine the values of these three parameters.

The early predictions of the magnitude of the fluctuations were off by a factor of 100, and any theory that assumes that more than 90 percent of the matter in the universe is in some bizarre form that we know nothing about deserves a certain amount of skepticism.







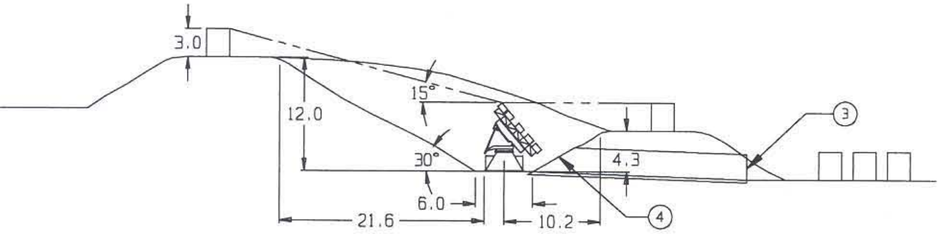
Theoretical calculations for the various cosmological models have predicted the expected temperature fluctuations over a range of angular scales. The top graph of three values of the Hubble constant (related to different ages of the universe) shows a peak at ranges within the CBI's view. In the middle graph, which varies the density parameter (with a fixed Hubble constant), there are several peaks from 0.3 to 1 moon diameter. The bottom set of curves varies the cosmological constant.

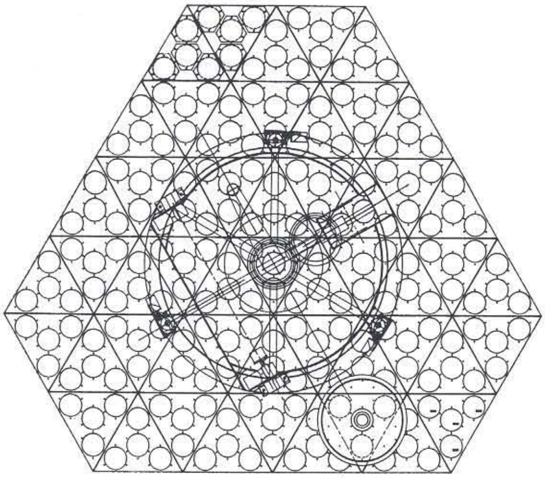
We should, however, be careful about taking these models too seriously; they could all be wrong. The early predictions of the magnitude of the fluctuations were off by a factor of 100, and any theory that assumes that more than 90 percent of the matter in the universe is in some bizarre form that we know nothing about deserves a certain amount of skepticism. On the other hand, the models may be right, so what we want is an instrument that will enable us to make the images needed to measure the actual variations in temperature over this range of angular scales. We may end up finding something that's completely different from any of these models, but whatever we discover is bound to be very interesting.

The CBI will be a radio interferometer consisting of 13 antennas, each with its own receiver, mounted on a 6.5-meter platform. The signals from each pair of antennas (78 in all) will be combined in a correlator and then recorded. This is a standard radio-astronomy technique, which enables us to make images of the sky that reveal, in this case, very small variations in temperature.

The CBI has to measure temperature differences of only 10 millionths of a degree—a very difficult achievement. We have to use extremely sensitive receivers, which have to be cooled to 15K and kept at this temperature for many months. In addition to thermal noise, which we minimize by using cooled receivers, there are a number of sources of systematic error that can easily swamp the signals that we are trying to detect. One of the largest of these error sources is so-called "cross talk" between adjacent antennas, in which some of the noise from one receiver leaks back out of the front of its antenna and is picked up by an adjacent antenna. When the signals from the two antennas are correlated, the noise shows up as a correlated signal—mimicking a signal from the sky. Steve Padin has used a novel antenna design to minimize this cross talk, and his prototype tests show that the cross talk can be further reduced to an acceptable level if we add a third axis of rotation to the telescope. In addition to rotating about two axes to point in the right direction, the telescope will rotate about its optical axis. This will enable us to discriminate between signals coming from the sky and signals generated within the instrument itself—the former will remain fixed in the heavens as the telescope rotates. This extra rotation, plus very careful design of the antennas themselves, has solved the cross talk problem.

Another problem is caused by radio-frequency emissions from our own galaxy. We are observing the microwave background radiation through a rather dirty window—the radiation from dust and from free electrons within our galaxy is comparable to the signals we wish to observe. Fortunately, the spectrum of this foreground radiation is very different from that of the microwave background radiation. Therefore, by observing at a





The plan for the CBI installation and ground screen is shown at top left. The instrument will be placed in a small crater-shaped hollow lined with reflecting material. In this way none of the 13 radio telescopes on the platform will be able to receive radiation directly from the ground, thus reducing the effects of signals from the ground to an acceptable level. At bottom left is a face-on view of the CBI. A single one-meter telescope is shown as a circle near the lower right. It will be possible to center the 13 individual telescopes on any of the smaller circles; the large variety of configurations will enable scientists to tailor the resolution of the instrument to the observations in hand. The large circle in the center is a bearing supporting the array.

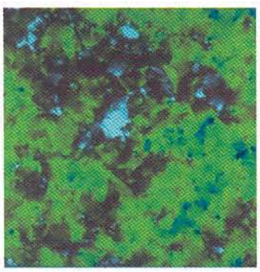
number of different frequencies, we can subtract out the foreground signals to get a clear image of the microwave background radiation itself. The CBI will therefore observe in 10 frequency bands, each one gigahertz wide, between 26 gigahertz and 36 gigahertz.

A third serious difficulty for any instrument observing the microwave background radiation is the atmosphere. Small fluctuations in the levels of water vapor in the atmosphere add noise to the observations, and contaminate the images. Consequently, it can take months to make an image that could be obtained in a few days in the absence of the atmosphere. To minimize the effects of the atmosphere, a number of high, dry sites are being considered for the CBI: the White Mountains, in California; Mauna Kea, in Hawaii; Antarctica; and sites in the high Andes in Chile and Argentina. The photo on page 30 shows the terrain near one of the prime sites we are considering in Chile.

Finally, thermal radiation from the ground underfoot can easily swamp the microwave background radiation signals. In order to eliminate this radiation, the CBI will have a reflecting ground screen, so that the ground will not be visible from any of the antennas at any time.

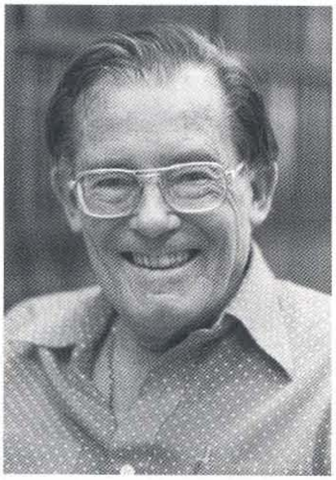
This is a competitive field, and in order to be in the lead we need to have the CBI fully operational and making images within two years. Our work thus far has been made possible by generous gifts from Ronald and Maxine Linde and from Caltech, who provided the required matching funds that enabled us to obtain a $2,000,000 grant from the National Science Foundation. We now have half of the CBI's total cost of $6,500,000 in hand, and are actively exploring funding possibilities for the remainder. We hope we succeed, because a scientific opportunity of this magnitude is extremely rare, and it's hard to imagine a more exciting prospect than seeing this critical stage in the birth of the universe. These images should tell us how all structures formed in the universe. They should also tell us whether the dark matter that we know is out there is composed of normal material, or whether it is an exotic substance that has thus far eluded direct detection. Nature is continually surprising us, and the CBI may well reveal unforeseen aspects of the early universe that could revolutionize our understanding of cosmology, and possibly of basic physics itself. □

One of the more bizarre theories of galaxy formation involves "cosmic strings," which would produce fluctuations in the microwave background radiation with characteristics different from most other theories. All of them may be wrong.



Tony Readhead, who earned his bachelor's degree from the University of Witwatersrand, South Africa, in 1968 and his PhD from the University of Cambridge, England, in 1972, came to Caltech as a research fellow in 1974. He was named professor of radio astronomy in 1981 and professor of astronomy in 1990. From 1981 to 1986 Readhead served as director of the Owens Valley Radio Observatory and from 1990 to 1992 as executive officer for astronomy. This article was adapted from the Watson Lecture he delivered last February.

OBITUARIES



James F. Bonner
1910 – 1996

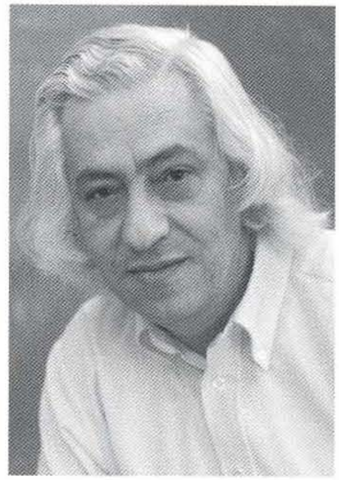
James Bonner, PhD '34, professor of biology, emeritus, at Caltech and a member of the National Academy of Sciences, died September 13, at the age of 86. He had been a Caltech faculty member for 60 years.

Born in Ansley, Nebraska, Bonner earned a BA in chemistry and mathematics from the University of Utah in 1931. After completing his Caltech PhD, he spent a year in Germany and Switzerland as a National Research Council Fellow. He came back to Caltech in 1935 as a research assistant, and received his faculty appointment the next year. He was named professor of biology in 1946.

Bonner's research interests included plant biochemistry and the genetic engineering of agricultural crop plants. He was the inventor of a novel way to increase the yield of rubber trees—which led to the approximate doubling of rubber yield per acre per year in Malaysia—as well as co-inventor of the method used by most Florida citrus growers for the mechanical harvesting of oranges.

His research covered almost every facet of plant biology, from the study of plant growth hormones and the discovery of new plant hormones, to the biochemistry of respiration, photosynthesis, rubber biosyntheses, and chemical ecology. For the last quarter-century of his career, he concentrated on the isolation and study of the genetic material, its chemistry and packaging into chromosomes, and the control of gene expression in plants and animals. Bonner was also active in discussions concerning the future of industrial civilization, particularly with regard to food, population, and the outdoors.

An avid member of several alpine and skiing organizations, Bonner climbed in the Himalayas and served as a ski patrolman for the National Ski Patrol System.



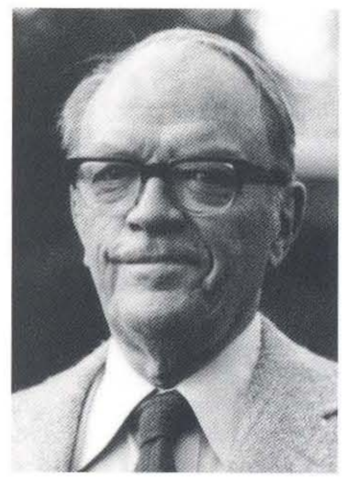
Ricardo Gomez
1930 – 1996

Ricardo Gomez, professor of physics, emeritus, died unexpectedly on October 14 in Pasadena. He was 66.

A native of Bogota, Colombia, Gomez came to Caltech as a research fellow in 1956, after earning his PhD from the Massachusetts Institute of Technology (as well as his bachelor's degree in 1953). He was appointed senior research fellow in 1959, became associate professor in 1971, and was named professor in 1977. He retired in July.

An experimental particle physicist, Gomez in his early years used Caltech's 1 GeV electron synchrotron for photoproduction studies at what was then the high-energy frontier of particle physics. He later helped establish the new style of doing particle physics experiments at remote accelerators, contributing to Caltech-led experiments at Lawrence Berkeley Laboratory, the Stanford Linear Accelerator Center, the Brookhaven National Laboratory, and Fermilab.

His research included studies of the photodisintegration of the deuteron, the photoproduction of various mesons from nucleons, and the interactions of high-energy mesons with nucleons. He also searched for fractionally charged particles, made experimental tests of quantum chromodynamics, and investigated certain meson decay modes of special interest. Gomez was also noted around campus for his commitment to undergraduate teaching.



Hallett D. Smith
1907 – 1996

Hallett Smith, Caltech professor of English, emeritus, died of pneumonia on his 89th birthday, August 15. One of this century's most eminent Elizabethan scholars, Smith served as chairman of the Institute's Division of the Humanities and Social Sciences from 1949 to 1970.

Smith came to Caltech in 1949 as professor of English.

Smith, continued

In 1970 he was made a research associate of the Huntington Library and was named professor emeritus at Caltech in 1975.

Smith's principal area of research was 16th- and 17th-century English literature, and he was the author of many books and articles on Elizabethan literature with a special emphasis on Shakespeare. In 1952 he published *Elizabethan Poetry: A Study in Conventions, Meaning, and Expression*, a work considered one of the most authoritative studies on the subject. His later works included *Shakespeare's Romances: A Study of Some Ways of the Imagination* and *The Tension of the Lyre: Poetry in Shakespeare's Sonnets.*

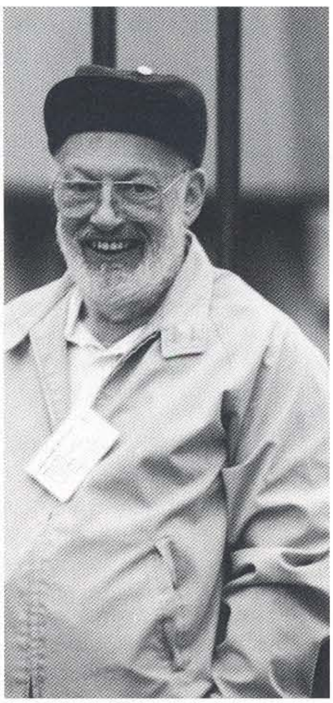
Smith was an editor of *The Norton Anthology of English Literature*, to this day the most widely used and influential anthology of literature in American colleges and universities. He was also the editor of the "Romances" and "Poems" chapters of *The Riverside Shakespeare*.

A graduate of the University of Colorado, he received his PhD from Yale University in 1934. He was a member of the English Department at Williams College from 1931 to 1949, and had been a visiting professor of English at Columbia University and at the Claremont Graduate School. In 1947–48 he was a Guggenheim Fellow, and in 1952 he won the Chapbook Prize of the Poetry Society of America.

Revel Named Dean of Students

Jean-Paul Revel, the Albert B. Ruddock Professor of Biology, has been appointed to a three-year term as dean of students, succeeding Rod Kiewiet, professor of political science, who had served in the post for four years.

Born in France, Revel graduated from the University of Strasbourg in 1949 and earned his PhD at Harvard in biochemistry in 1957. After a number of years on the faculty of Harvard Medical School and Cornell Medical School, he came to Caltech as professor of biology in 1971. He was named the Ruddock Professor in 1978.



Jean-Paul Revel waits to embark for Catalina Island and his first Freshman Camp as Dean of Students.

Revel has long been revered for his warm and enthusiastic commitment to undergraduates. A popular teacher, he is a recipient of ASCIT's teaching award and can attract record numbers of students to 8 a.m. classes. He has served as pre-med adviser and as chair of the Biology Undergraduate Advisory Council, and is currently developing the new core course Biology 1, which he and Scott Fraser, the Anna L. Rosen Professor of Biology, will teach in the spring term.

In his research in cell biology, Revel studies cell-to-cell communication, electron microscopy, and scanned probe microscopy. Many of his fine electron micrographs have graced the pages of *E&S*, and he contributed an article, "Cell Biology of Heart Disease," in the January 1987 issue.

Help Wanted

President Everhart will be stepping down in the fall of 1997, after a decade at Caltech's helm. As part of the search for his replacement, the faculty advisory committee (chaired by Kip Thorne, the Richard P. Feynman Professor of Theoretical Physics) is seeking "outstanding nominees wherever they may be found, and thus would appreciate your calling our attention to highly qualified individuals of either gender and any ethnic background, and from industry, government, and foundations, as well as academia." All members of the Caltech community—faculty, students, staff, administration, JPL, alumni, associates, and friends of Caltech—are encouraged to submit nominations and suggestions about the search process (for example, qualities the new president should have, or issues the incoming president will have to address that should influence the selection process). All input will be kept strictly confidential.

Nominations and supporting material (the committee asks that you also provide a paragraph explaining why your nomination would be an outstanding president; a curriculum vitae, or a suggestion as to where one could be obtained; and a list of people who might provide insightful evaluations of the nominee) should be submitted to:

nominations@pressearch.caltech.edu

Other suggestions should go to:

suggestions@pressearch.caltech.edu

If you wish to encrypt your e-mail via PGP, you can obtain the PGP public key by sending an e-mail (whose contents will be ignored) to PGPKey@pressearch.caltech.edu. You can also send regular mail to:

Presidential Search Committee
Caltech
P. O. Box 60070
Pasadena, CA 91116

The committee can be reached by fax at 818-584-7198.

You can help produce Nobel Laureates like Ed Lewis with financial support of Caltech through our Charitable Trust Program. Receive income for life and a tax deduction, have your money professionally managed, and leave a lasting legacy at the Institute.

GIVING BACK

When Ed Lewis, Caltech's Thomas Hunt Morgan Professor of Biology, Emeritus, won the Nobel Prize in physiology or medicine last year for his work on the genetic control of early embryonic development, there were celebrations, interviews, toasts, and telegrams. Then the "hard reality" set in. What would he do with his check, after splitting the cash award with two other scientists with whom he shared the prize?

You'd think that the first thing someone would do with a windfall would be to splurge and buy something that he always wanted. But Lewis, a modest, energetic man who finds time daily to swim, play the flute, and devote hours to his research, says he already has everything he needs.

The bottom line is that he wanted to do something for Caltech, where he has spent most of his career. And he especially wanted to help students. So he used his prize money to establish a trust with Caltech that will go toward undergraduate scholarships when he dies.

While creating the trust will reduce his taxes on the prize, Lewis says, "That's not the main reason I'm doing this. Caltech has provided the kind of excellent environment that has allowed me to carry out the research that has led to the award of the prize. Also, it has always been rewarding to see how many of our students have gone on to become world leaders in their fields," said Lewis. "In these days of high tuition costs, scholarships are needed more and more."

If you'd like more information about these and other ways to benefit yourself and Caltech, contact us:

Thomas W. Smith, CFP
Susan A. Walker, CFP
The Office of Gift and Estate Planning
California Institute of Technology
Mail Code 105-40
Pasadena, California 91125
(818) 395-2927
fax (818) 683-9891

ENGINEERING & SCIENCE

California Institute of Technology
Pasadena, California 91125