California Institute of Technology

Engineering & Science

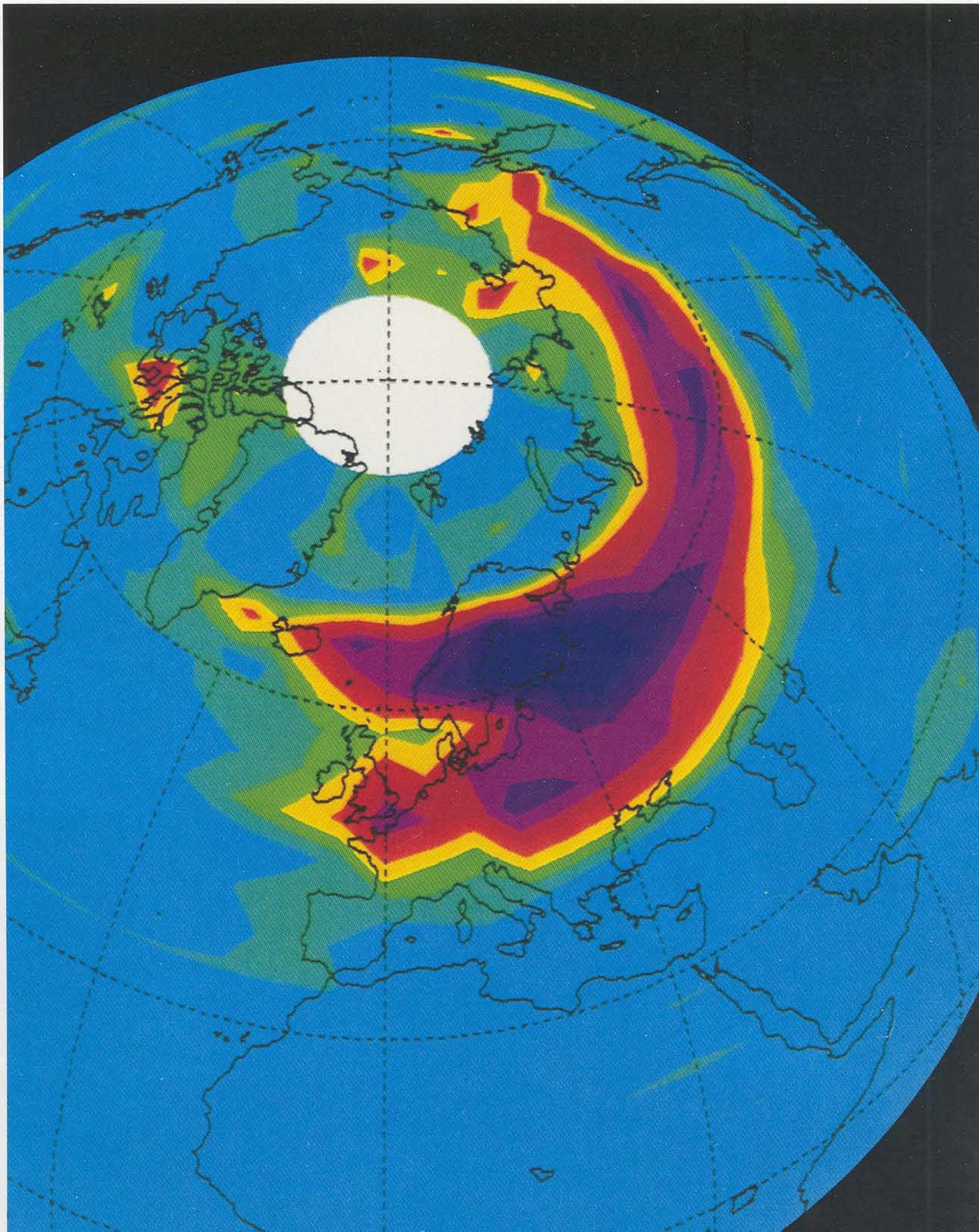
Summer 1993

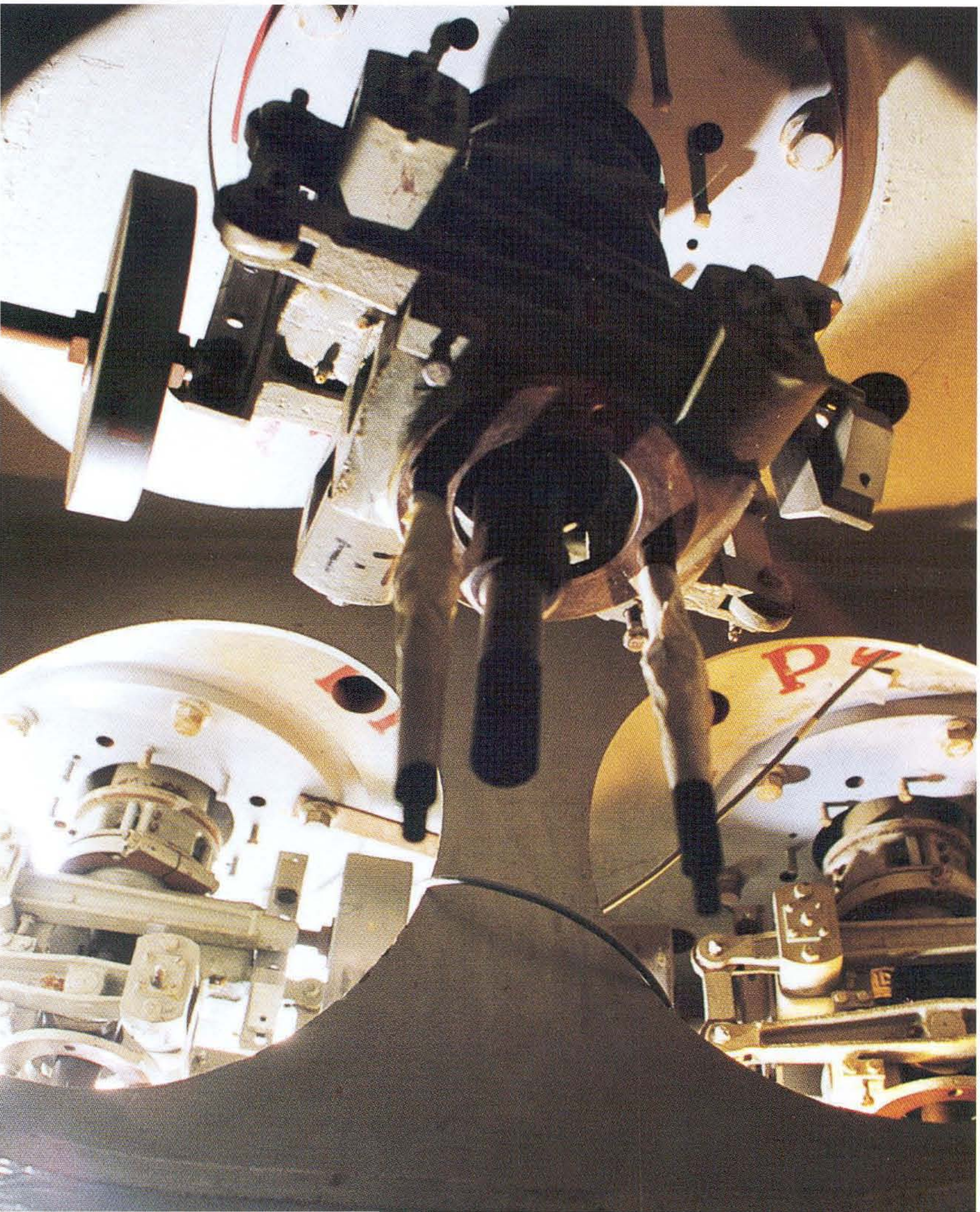
In this issue

Ozone Losses

Squeezed Light

Cancer Genes



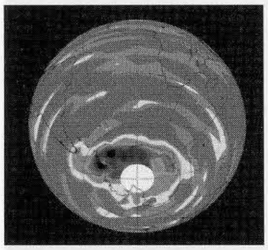


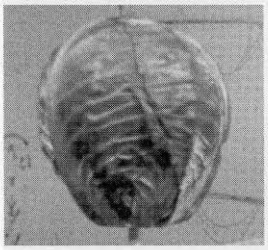
These metal sculptures are actually the protruding ends of three of the 36 back supports that maintain the Hale Telescope's 200-inch mirror in its optimum light-gathering shape. The visible portion of the back supports dangle from the mirror's back side into holes in the mirror cell, a two-foot-thick labyrinth of steel members that holds the mirror. The bulk of the back-support mechanism is up inside the mirror itself, invisible and inaccessible since November 16, 1947, when the mirror was installed in the cell for its trip to Palomar Mountain.

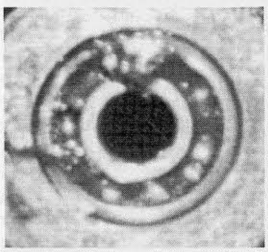
California Institute of Technology

Engineering & Science

Summer 1993
Volume LVI, Number 4







On the cover: Data on the abundance of chlorine monoxide, measured by the Microwave Limb Sounder on January 11, 1992, provided this surprising picture of high levels (from dark purple down to red) over Russia and northern Europe. (The white circle around the pole represents the region out of the instrument's field of vision.) How chlorine monoxide is related to the destruction of ozone is explained in an article beginning on page 2.

2 The Chlorine Threat to Stratospheric Ozone — *by Joe W. Waters*
From JPL's satellite-borne microwave instrument, launched in 1991, has come dramatic additional evidence implicating chlorine chemistry in Antarctica's ozone hole—and the suggestion that similar processes of ozone destruction are occurring in the north.

14 New Light on the Nature of Darkness — *by H. Jeff Kimble*
In the bizarre world of quantum optics, it's possible to make light that is darker than ordinary darkness, and other even stranger things.

26 Roundworm Cells and Cancer Genes — *by Paul W. Sternberg*
By studying a creature the size of this comma, biologists may be closing in on the secrets of the human body's cancer-causing genes.

34 The Bad News Bearings — *by Douglas L. Smith*
How do you get to something hidden behind $14^1/_2$ tons of glass on one side and 20 tons of steel on the other, without separating the two? Such was the Palomar Observatory staff's dilemma when bearings deep within the Hale Telescope's 200-inch mirror began to stick.

Departments

42 Letters

43 Random Walk

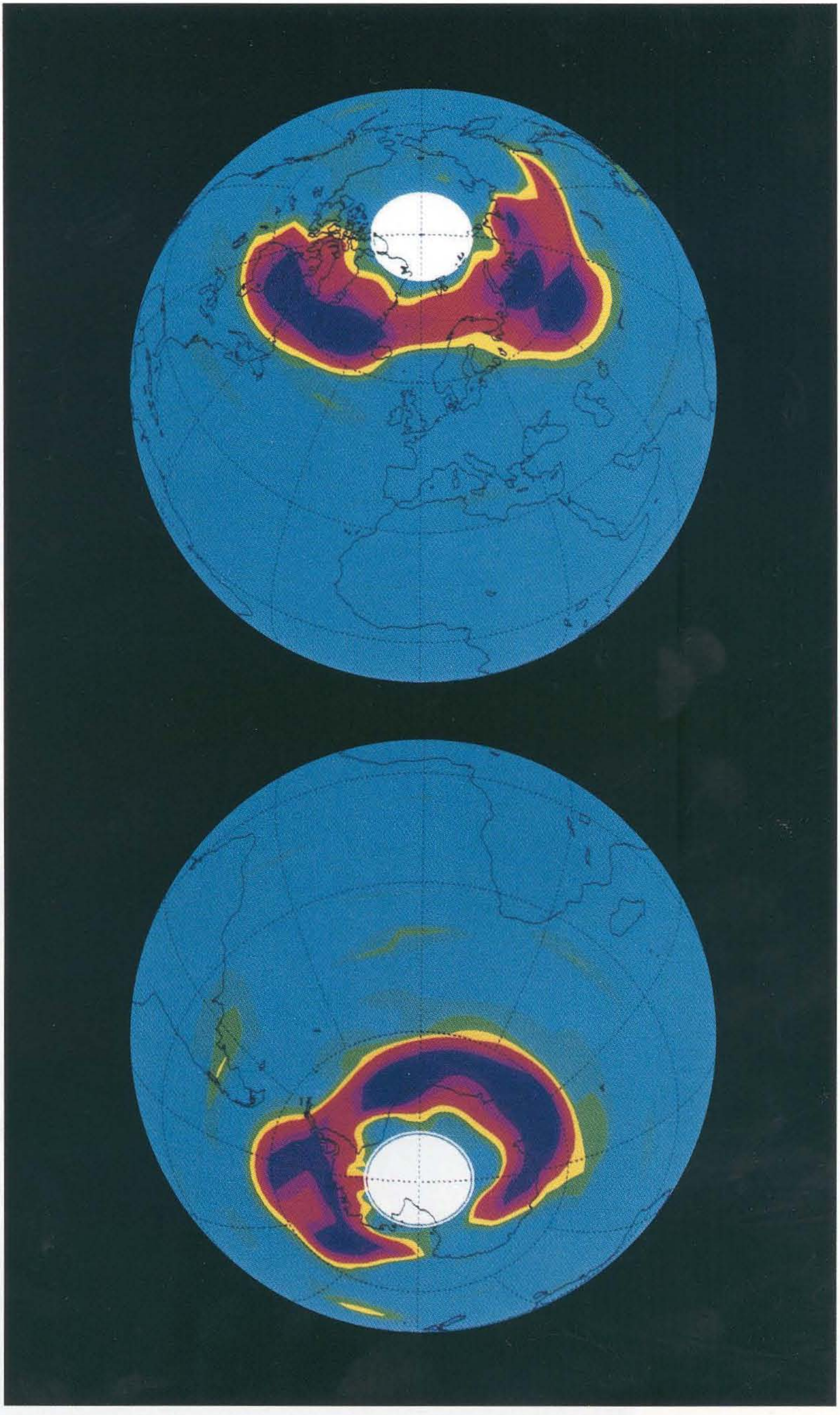
44 Obituaries

Engineering & Science (ISSN 0013-7812) is published quarterly, Fall, Winter, Spring, and Summer, at the California Institute of Technology, 1201 East California Boulevard, Pasadena, California 91125. Annual subscription $8.00 domestic; $20.00 foreign air mail; single copies $2.00. Third class postage paid at Pasadena, California. All rights reserved. Reproduction of material contained herein forbidden without authorization. © 1993 Alumni Association, California Institute of Technology. Published by the California Institute of Technology and the Alumni Association. Telephone: 818-395-3630. Postmaster: Send change of address to Caltech 1-71, Pasadena, CA 91125.

PICTURE CREDITS: Cover, 2–13 – JPL; inside front cover, 19, 21, 23, 25, 32–34, 37–41, 44 – Bob Paz; 14 – Errol Le Cain; 16–18, 22, 23 – Jeff Kimble; 20 – Nikos Georgiades; 24, 25 – Hideo Mabuchi; 26, 29–31 – Paul Sternberg; 28 – Bob Turring; 31 – Linda Huang; 36 – Hal Petrie; 37 – Robert Harrington

William M. Whitney
President of the Alumni Association
Thomas W. Anderson
Vice President for Institute Relations
Robert L. O'Rourke
Assistant Vice President for Public Relations

STAFF: *Editor* — Jane Dietrich
Managing Editor — Douglas Smith
Copy Editors — Michael Farquhar, Danielle Gladding, Julie Hakewill
Production Artist — Barbara Wirick
Business Manager — Debbie Bradbury
Circulation Manager — Susan Lee
Photographer — Robert Paz



The Chlorine Threat to Stratospheric Ozone

Whether we're in for still more unpleasant surprises is a very difficult question to answer.

by Joe W. Waters

Chlorine monoxide (ClO), the form of chlorine mainly implicated in the destruction of stratospheric ozone, as measured on February 14, 1993, in the northern hemisphere, and on August 14, 1992, in the southern hemisphere, exhibits similar patterns of abundance around both poles. Red and darker colors indicate ClO abundances of one part per billion or greater. Data from JPL's Microwave Limb Sounder produced these maps.

Ozone has a split personality. Down near Earth's surface, it's a bad guy—an element of pollution and smog—which those of us who live in the Los Angeles area are all too familiar with. In the stratosphere, however, where most of the ozone resides, it's a good guy—in fact, its presence there is essential for life on Earth. Here I'm going to discuss only the good-guy ozone and how it's threatened by chlorine, currently a very active area of research. Tremendous progress is being made in understanding this phenomenon, but in the history of this research scientists have been unpleasantly surprised by the processes that can deplete ozone. Whether we're in for still more unpleasant surprises is a very difficult question to answer.

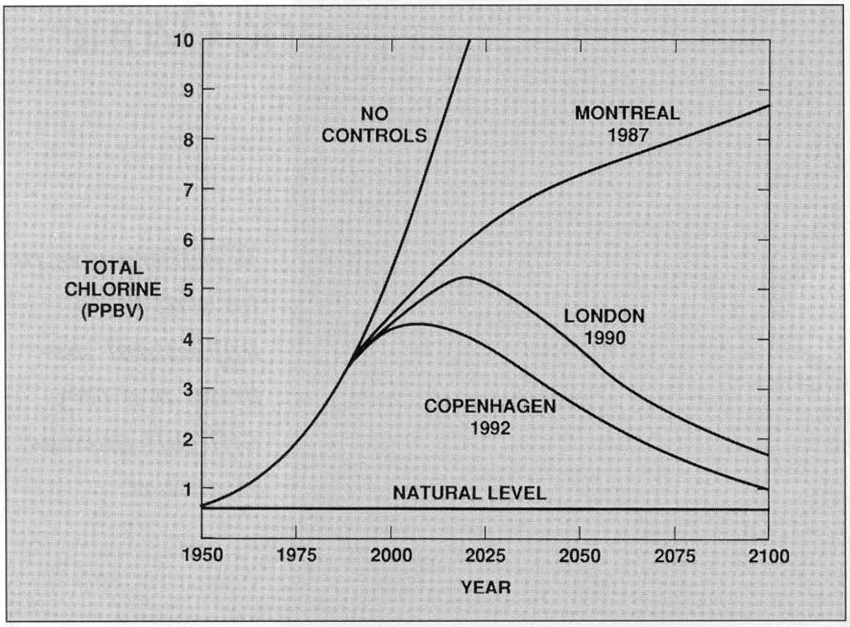
From space Earth's atmosphere appears as an eggshell-thin layer, visible along the horizon. The stratosphere is part of the upper atmosphere, above the region in which most of the clouds form, extending from about 10 to 15 kilometers to about 50 kilometers above Earth's surface. What is commonly called the ozone layer occurs around the region where ozone abundances peak, at about 20 to 25 kilometers above the surface.

Stratospheric ozone and life have a very special relationship. In the primeval atmosphere, before there was life, neither ozone nor oxygen existed in significant amounts in the atmosphere. Scientific theory holds that oxygen was produced by living organisms after life evolved in the sea or other bodies of water; oxygen entered the atmosphere and some of it went on to form ozone. Once the ozone layer was in place, life could then safely climb out of the water onto land. So without life, there would be no ozone in the atmosphere, and reciprocally, that ozone now shields life from solar ultraviolet radiation.

A molecule of ozone consists of three atoms of oxygen bound together, represented as O_3. Its abundance in the stratosphere is relatively slight—only about 10 molecules per million total molecules at maximum. But that small abundance is a very effective absorber of solar ultraviolet radiation. Ultraviolet radiation, a very short wavelength of light, is very damaging to life, a principal manifestation in humans being skin cancer. It has been estimated that a 5-percent reduction in ozone would amount to about 100,000 additional skin-cancer cases a year in the United States. A striking effect of ozone depletion on the ecosystem has been observed in the marginal ice zone in the sea around Antarctica. The productivity of photoplankton there, the base of the food chain in the sea, has been observed to decrease by some 10 percent when the ozone hole is overhead.

Another aspect of ozone is that its absorption of ultraviolet radiation heats the upper stratosphere. Ozone largely determines the temperature structure of the upper atmosphere, which, in turn, determines its circulation. Since the stratosphere is heated from above, it's a relatively stable layer—and therefore relatively comfortable for airplanes. But the troposphere, the layer of the atmosphere nearest Earth's surface, is heated from below (mainly by visible radiation absorbed by the surface), which creates relatively turbulent air and the weather patterns we experience.

The amount of ozone in the stratosphere is



Levels of chlorine in the stratosphere from 1950 projected through the next century, given the controls on CFCs that have been proposed under various international agreements. The total amount of chlorine (whose natural level is 0.6) is represented along the vertical axis in parts per billion. (From *WMO and the Ozone Issue*, World Meteorological Organization, Geneva, 1992.)

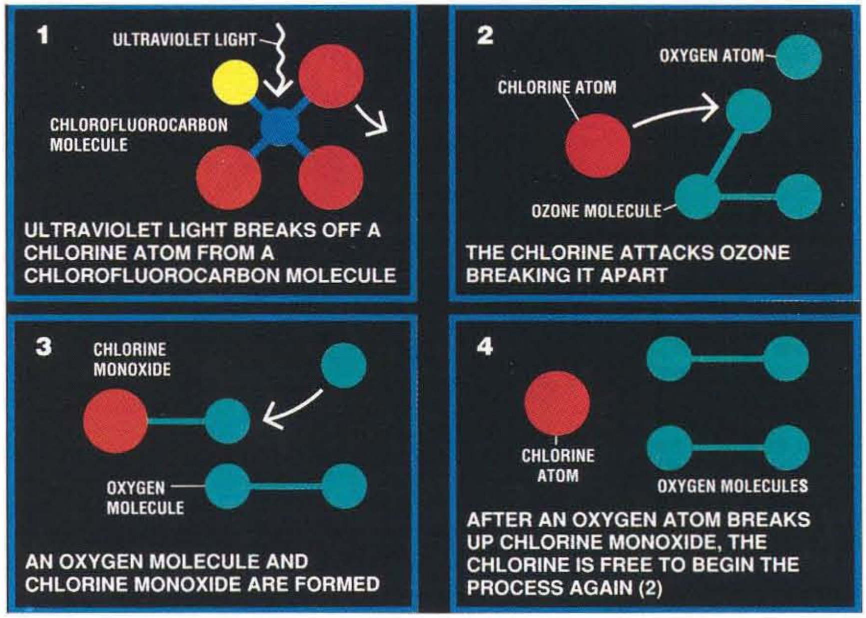
Essentially all of that additional chlorine has come from the production of chlorofluorocarbons (CFCs).

maintained by a dynamic balance. Ozone is being produced and destroyed continually, and the net difference between those effects determines how much there is. So let's consider in more detail how that process works. Ozone is produced from molecular oxygen, O_2, a form that is much more abundant in the atmosphere—about 21 percent. Ozone is produced when an energetic photon from the sun breaks some of the O_2 molecules into two atoms of oxygen. The resulting oxygen atoms react with some of the remaining O_2 molecules to form O_3. A tremendous amount of energy from the sun goes into producing ozone—approximately 2×10^{13} watts, about three times the total power consumed by humankind. It's entirely unfeasible for us to replenish ozone because of the energy required.

Ozone in concentrated form, however, is about as explosive as dynamite. So, although it's extremely difficult to produce ozone, it's very easy to destroy it. This can happen in several ways. For the first 20 years or so of research in this field, which started around 1930, the only known ozone-destroying mechanism was the interaction of O_3 with an oxygen atom to create two oxygen molecules. It turns out that this is a very slow process, and is now thought to account for only about 10 percent of the total destruction of ozone. But since 1950 several catalytic cycles have been discovered that can speed up the destruction of ozone very rapidly. Four chemical families are involved in these speeded-up reactions—hydrogen, nitrogen, chlorine, and bromine. A major source of the latter two is industrial production down on Earth's surface. Chlorine, in particular, is known to be a serious threat to stratospheric ozone.

The graph above shows the total amount of chlorine in the stratosphere, plotted over a 150-year time period with projections forward for various future scenarios. The vertical scale represents the total chlorine in parts per billion. The natural level has been measured at about 0.6 parts per billion, and we know that the source molecule for this natural chlorine is methyl chloride (CH_3Cl), produced, for example, from sea salt. Currently, the total amount of chlorine in the stratosphere is about 3.6 parts per billion—six times the natural amount. Essentially all of that additional chlorine has come from the production of chlorofluorocarbons (CFCs). The curve on the graph that goes up out of sight represents what would happen if there were no controls, no cutbacks in the industrial production of these molecules. The other curves show what would happen under various international agreements or proposals, the first of which is the Montreal Protocol of 1987, which was established after chlorine had been proven to be a threat to the stratosphere, but before the cause of the ozone hole had been discovered. This agreement cut back the production of chlorine from industrial sources, and was strengthened in London in 1990, as a result of proof that the ozone hole was due to chlorine. Just this past year, even tighter restrictions were suggested at an international meeting in Copenhagen.

But even with these cutbacks, there's going to be a lot more than the natural amount of chlorine in the stratosphere for a very long time. The



CFCs (carbon atoms surrounded by chlorine and fluorine atoms) diffuse intact into the upper stratosphere, where ultraviolet radiation can rip off a chlorine atom, setting it free to combine with ozone to form chlorine monoxide (ClO) and O_2. A single oxygen atom can break up ClO, freeing the chlorine atom again to continue its ozone-destroying cycle.

chlorine level at which the ozone hole is generally considered to form is about two parts per billion, and the chlorine in the stratosphere will remain above that level for at least 100 years. Even if we completely ceased production of CFCs today, the reservoir of chlorine in the lower atmosphere, which takes time to diffuse upward to the upper stratosphere, would continue to increase for about 10 years. So, no matter what we do, the problem is going to be around for a long time.

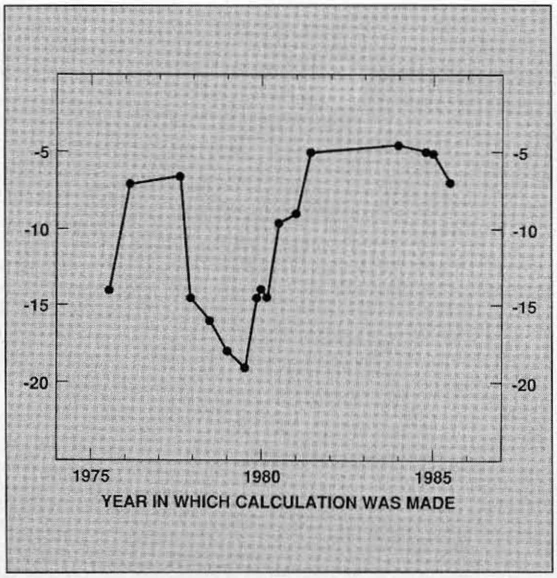
How does chlorine actually destroy ozone? Chlorine enters the stratosphere in the form of CFCs. Economically, these are extremely important molecules, used in refrigeration systems, for example. CFC molecules consist of a carbon atom (colored blue in the illustration above) surrounded by chlorine (red) and fluorine (yellow) atoms. They are extremely stable chemically, and it's this chemical stability that makes them dangerous to ozone, even though that sounds like a paradox. They're initially released in the lower atmosphere, where they don't interact chemically to form other substances that would get removed from the atmosphere naturally. And they're not soluble in water, so they don't get rained out. Eventually they diffuse upward, intact, into the stratosphere, a process that takes some 10 years. If for the first 10 or 20 years of their existence we can assume that these molecules resided in refrigeration systems or other industrial products before being released into the atmosphere, then much of the chlorine that we're looking at now in the stratosphere was actually produced some 20 or 30 years ago.

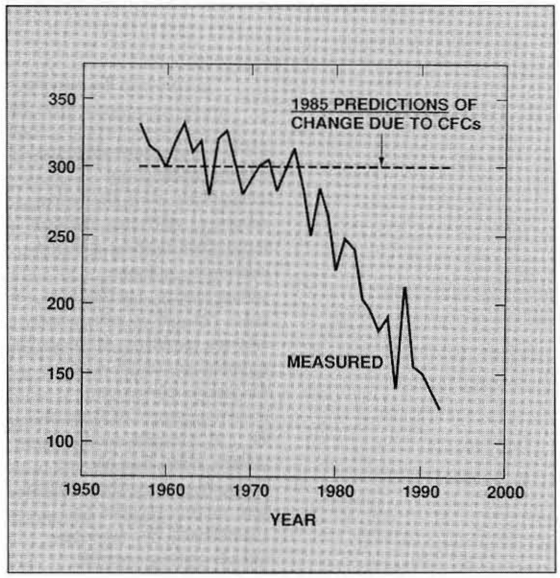
The chlorine level at which the ozone hole is generally considered to form is about two parts per billion, and the chlorine in the stratosphere will remain above that level for at least 100 years.

Because their upward diffusion takes such a long time, the CFCs get mixed around by winds and are horizontally distributed uniformly around the globe. We think they enter the stratosphere mostly through the tropics. Once they get above the ozone layer, they're no longer protected from solar ultraviolet radiation. This radiation breaks loose a chlorine atom, which then attacks the ozone. The chlorine atom rips off one of the oxygen atoms from ozone to form an intermediate molecule—chlorine monoxide (ClO)—and leave an oxygen molecule (O_2). Chlorine monoxide is extremely short-lived. It's very reactive and interacts with the atomic oxygen that is naturally present there. The atomic oxygen rips the O off the ClO to create another oxygen molecule and leaves the chlorine atom free. So this chlorine atom can then go on to attack another ozone molecule and repeat the cycle. That's what makes it so effective: if a chlorine atom destroyed only one ozone molecule, it wouldn't be a problem. But because the cycle repeats over and over, a little chlorine goes a long way.

What breaks this cycle? Eventually the chlorine atom, instead of interacting with ozone, will interact with methane (CH_4) in the atmosphere to form hydrochloric acid (HCl). The HCl diffuses downward, and because it's soluble in rain the chlorine is finally washed out, closing the cycle. This "washing chlorine out of the atmosphere" is a very slow process, taking 100, maybe even 200, years. The ozone destruction cycle can also be broken when ClO interacts with nitrogen dioxide (NO_2) in the atmosphere to make chlorine nitrate ($ClONO_2$).

For about 10 years after the chlorine threat to





Right: After the chlorine cycle of ozone destruction was discovered in 1974, calculations based on laboratory measurements of the chemical reactions were made to predict the percentage change in total ozone due to CFCs emitted at the 1974 rate. This prediction fluctuated over the years before stabilizing at about a 5 to 7 percent total decrease (in 100 years), which caused little alarm in 1985. (From *Atmospheric Ozone 1985*, World Meteorological Organization, Geneva, 1985.)

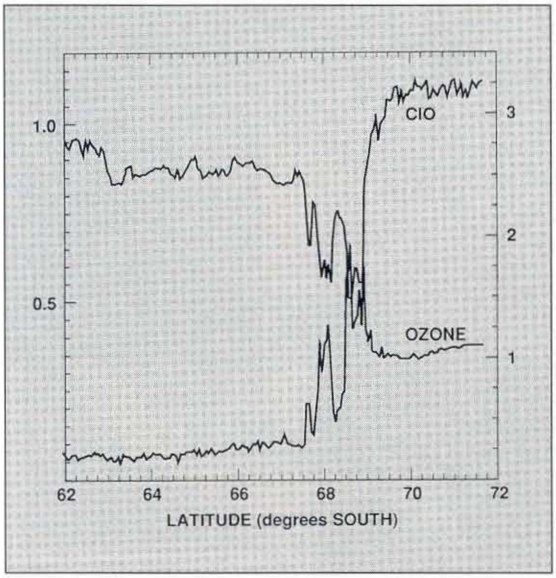
Far right: The relatively rosy prediction of 1985, represented by the dashed line, was shattered when actual stratospheric ozone levels, measured by Joe Farman and a British Antarctic Survey team over Antarctica, started a continuing nosedive in the 1970s. The vertical axis is in Dobson units. (One hundred Dobson units represents a one-millimeter-thick layer of ozone, if that ozone were all concentrated in a layer at Earth's surface.)

ozone was discovered (in 1974 by Mario Molina and F. Sherwood Rowland of UC Irvine), this process was thought to be the whole story. Based on laboratory measurements of the rates of chemical reactions, predictions were made. How much ozone would this process be expected to destroy? How much would this process change the equilibrium of ozone and reduce its natural abundance? The graph above shows the predicted depletion—the new equilibrium that would be set up after the chlorine had done its stuff. In 1974 it was predicted that the ozone depletion would be about 15 percent and that this depletion would be reached after more than 100 years. In the graph, the horizontal axis represents the years the calculations were made, and the curve fluctuates up and down as laboratory measurements of reaction rates improved. After about 1980, it tended to stabilize at about 5 to 7 percent. This gave us a sense of complacency; we thought we knew what was happening. That ozone would diminish 5 to 7 percent was still a serious problem, but because this was supposed to occur over such a long time the problem didn't command any real sense of urgency. It would happen slowly enough for us to have time to think about what to do.

What was really happening over Antarctica, however, was another story, which is shown compared with the prediction (dashed line) in the graph above right. The vertical scale shows the total amount of ozone in the stratosphere above Antarctica in October, as measured by the researchers who discovered the Antarctic ozone hole. These were members of the British Antarctic Survey team, who began making measurements at the start of the International Geophysical Year of 1957. In the mid-seventies ozone over Antarctica really started taking a nosedive. By 1985 it was down by about a factor of two. This is a tremendous effect. At the outset, few really expected that chlorine could be the cause of the Antarctic ozone decrease; three hypotheses were put forward to explain it. One hypothesis ascribed it to the increase in solar activity: the greater the solar activity, the more nitrogen oxides are produced, which can destroy ozone. Another theory had to do with circulation: perhaps the circulation pattern over Antarctica is changing, causing an upwelling of ozone-poor air from below. The third theory had to do with the chlorine. We knew that chlorine was increasing in the stratosphere during this period, but what was being measured was so out of line with the ozone levels that had been predicted that this explanation was difficult to accept.

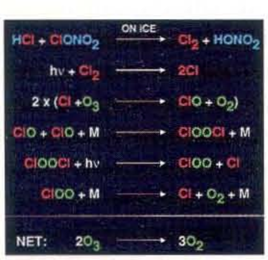
So after this work was published in 1985, scientists led by Susan Solomon of the National Oceanic and Atmospheric Administration (including a JPL group under Barney Farmer and Geoff Toon) hastily organized an Antarctic expedition to test these hypotheses. Measurements indicated that neither the solar activity nor the circulation could explain the dramatic decrease in ozone, but that it was due to chlorine chemistry, which is enhanced in the meteorological conditions over Antarctica. Observations also showed that it's in the lower portion of the stratosphere that the ozone is being lost.

The clincher to implicating chlorine as



Right: The correlation between a rise in chlorine monoxide and a drop in ozone was measured in the lower stratosphere over Antarctica in September 1987 by James Anderson and colleagues. The left vertical axis represents ClO in parts per billion, and the right vertical axis is ozone in parts per million.

Below: The cause of the unexpectedly high amount of ClO turned out to be a chemical reaction taking place on the particles (both water ice and nitric acid trihydrates) of the polar stratospheric clouds that form in cold temperatures. The end result is the conversion of two ozone molecules to three oxygen molecules.



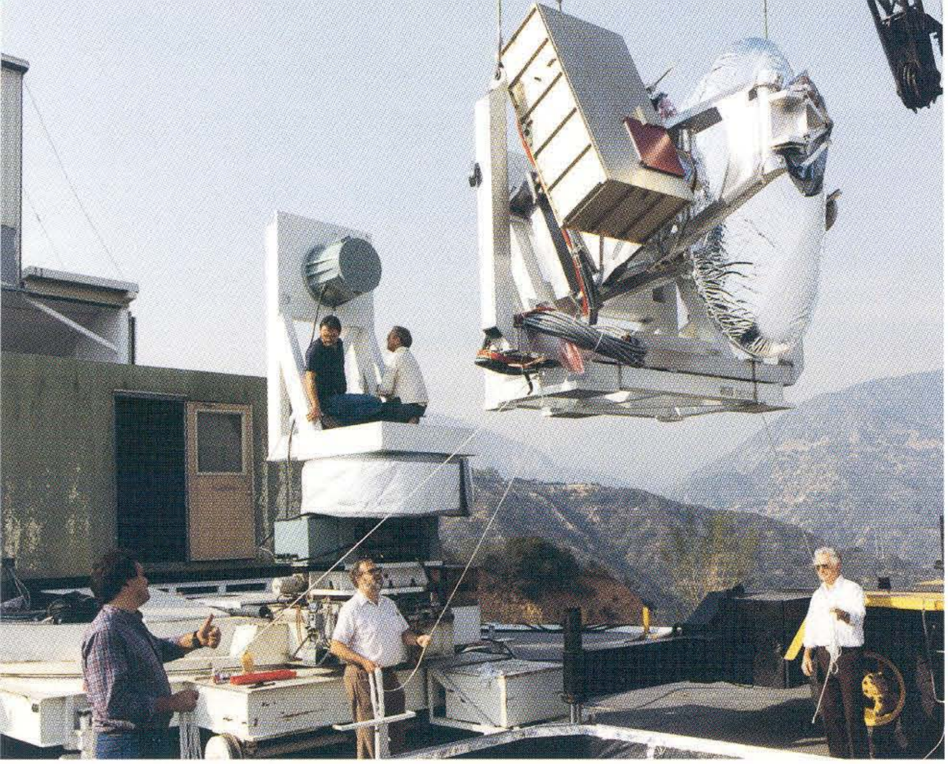
The abundance of the chlorine monoxide molecule is a direct measure of the rate at which chlorine destroys ozone.

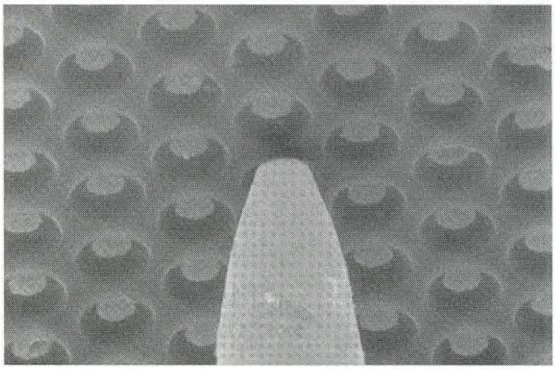
responsible for the ozone hole, certainly the one that has received the widest recognition, is the measurements made by James Anderson and colleagues at Harvard (earlier measurements by ground-based microwave techniques had also provided similar information). They had previously developed a technique for measuring the chlorine monoxide molecule, then put their instrument on a converted U2 spy plane in 1987, and flew from the tip of South America into the ozone hole. At the same time that they hit the ozone hole and the ozone abundances went down, the ClO abundances shot up. The abundance of the chlorine monoxide molecule is a direct measure of the rate at which chlorine destroys ozone. The oxygen atom in ClO has to have been ripped out of ozone, and for ClO to be there in any appreciable abundance, that process must be cycling because ClO is short-lived. In the curve at left you can see some of the impressive anticorrelations between measured ClO and ozone. Although this proved beyond any shadow of doubt that the ozone hole is due to chlorine, there was still a problem. All the theories at the time maintained that it was absolutely impossible to have this amount of ClO in the lower stratosphere. It just "couldn't happen." The theories predicted that the natural abundance would be almost 100 times lower than what was measured.

Something was obviously missing. And it turned out to be that the theoretical models included only the chemistry that occurs in the gas phase of molecules. We now know that additional reactions, which don't occur in the gas phase, take place on particles in the polar stratospheric clouds that form in the cold temperatures over Antarctica. HCl and $ClONO_2$ combine on the surface of these ice clouds to form Cl_2, which is released, and nitric acid ($HONO_2$), which stays on the ice clouds. Even rather weak sunlight breaks up the Cl_2 into chlorine atoms, which can then go on to attack ozone, forming ClO plus O_2. In the upper stratosphere the ClO can get recycled into chlorine atoms through interaction with atomic oxygen, as mentioned earlier, but in the lower stratosphere there isn't enough atomic oxygen to free up the chlorine again. There is, however, another mechanism in which two ClO molecules combine to form ClOOCl. This ClOOCl can then be broken down by sunlight to form ClOO plus Cl, and the ClOO can also be broken up by sunlight or collisions with other molecules to free the remaining chlorine atom. These reactions have the net effect of converting two ozone molecules to three oxygen molecules, and were initially investigated in the laboratory by Mario Molina in 1987 when he was at JPL.

In summary, what's happening over Antarctica is that the cold temperatures in the Antarctic winter cause ice clouds to form, which activates the chlorine—converting it from safe to dangerous forms. The ozone hole then occurs in September because the spring sunlight comes around and breaks up the chlorine, which is necessary to maintain the cycle. (Our recent satellite experiments, to be discussed later, show that this process is actually going on long before September.)

Now, here one might ask a very reasonable question. If the ozone hole is due to the low





Above: Before being shipped east to join the rest of the spacecraft, the Microwave Limb Sounder (MLS) is mounted on the antenna range above JPL to test its reception of signal sources across the valley.

Above, right: The heart of the MLS is this "whisker," which contacts the semiconductor diodes (the dots in the picture, which are about two microns across) to convert the very short wavelengths of the ClO signals to frequencies where they can be amplified by more conventional electronics. The special diodes used in the MLS were developed by R. J. Mattauch and colleagues at the University of Virginia.

temperatures causing activation of chlorine, isn't global warming a wonderful thing? Unfortunately, it doesn't work that way. When we speak of global warming, we mean warming near Earth's surface. This is because of the increase in CO_2, which puts a greenhouse blanket over us. But that greenhouse blanket is effectively below the stratosphere, and this causes the stratosphere to cool. So the net effect of "global warming" is to aggravate the ozone-destroying process.

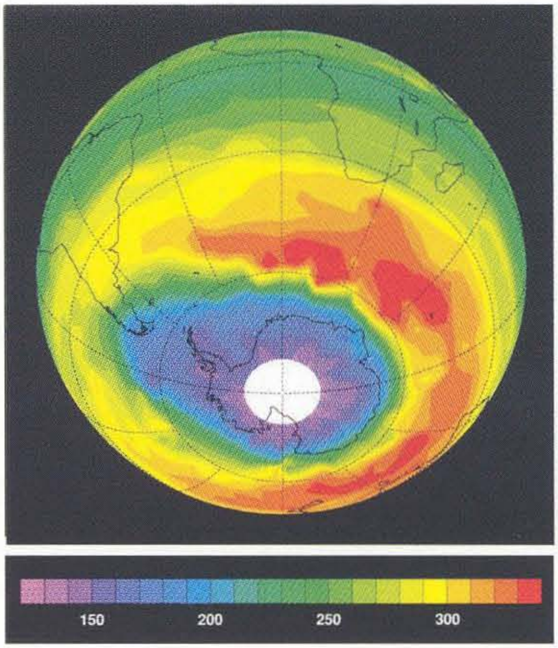
For almost two decades my group at JPL has been developing techniques for making measurements of the stratosphere. We do radio astronomy of the Earth. Through the wonders of quantum mechanics, a large number of molecules in the stratosphere—one of them being this culprit, ClO—naturally broadcast at particular frequencies. We can build sensitive receivers at these wavelengths and listen to the signals given off by the molecules. The intensity of the signal allows us to determine the molecule's abundance. Our work at JPL has not just concentrated on ClO, but is devoted to sorting out the spectrum of all the other molecules in the stratosphere as well. We not only have to know where our target molecules are, but also where there might be interfering lines that could mess up our measurements. A JPL group led by Herb Pickett and Ed Cohen is providing the enormous spectroscopy data needed for this task. We sent instruments up on balloons and aircraft before investing in satellites. Our first balloon launch, in 1980, laid a firm foundation for the satellite experiments to follow.

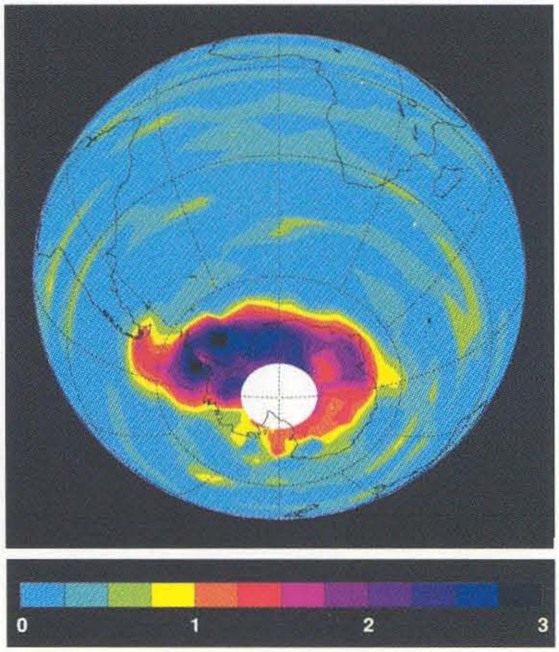
We currently have data coming in from the Upper Atmosphere Research Satellite (UARS), a project established by NASA to perform a comprehensive study of the upper atmosphere, particularly the ozone layer. Our group is responsible for one of 10 instruments on that satellite—the Microwave Limb Sounder (MLS). It picks up microwaves from the edge, or limb, of the atmosphere (seen tangentially through it). The heart of the instrument is the tiny "whisker" semiconductor diode junction (above), which transforms the very short wavelengths (about one and a half millimeters, or 1,000 times shorter than FM radio wavelengths) that the atmosphere is emitting to much longer wavelengths, or much lower frequencies. Then we amplify those signals to a level that can be detected with conventional electronics. Since nothing quite like the MLS experiment had ever flown in space before, it took a tremendous amount of skill and sophisticated design to put it together. MLS follows a long history of JPL microwave experiments in space, starting in 1962 with one going to Venus; Frank Barath and Jim Johnston, who led the overall management of MLS, first "cut their teeth" on the 1962 Venus instrument.

UARS was launched precisely on schedule from the space shuttle *Discovery* on September 12, 1991—an important deadline to make because the ozone hole forms in September. After launch and instrument turn-on, we had a few days to look south at the ozone hole. The UARS orbit is designed to switch observation directions every month; one month we would look mostly south (to 80° south latitude and 34° north) and then switch to looking mostly north the next month. On the next page, top left, is a map, looking

Right: The MLS's first view of the ozone hole on September 21, 1991, less than 10 days after launch, produced this expected picture, color-coded by Dobson units. Purple indicates very low levels of ozone.

Far right: A map of ClO in the stratosphere made on the same date shows an almost exact coincidence of high (red) ClO with the low levels of ozone. Scale is 10^{19} ClO molecules per square meter in a vertical column; one unit corresponds to about one part per billion of ClO in the lower stratosphere.

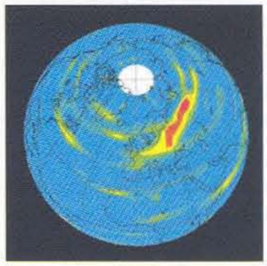


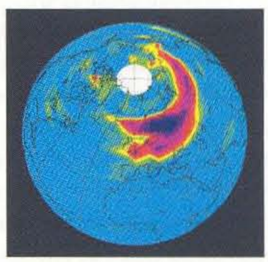


south, made from data taken less than 10 days after launch. This picture of the ozone hole (purple indicates a very low amount of ozone) is what we expected to see. But what our instrument mapped for the first time was ClO in the stratosphere. The ClO map (above, right) was made at the same time as the ozone measurements, with red and darker colors indicating high ClO abundances. You can see that the ClO is concentrated where the ozone is depleted; there's an almost exact coincidence.

When we switched to looking north in early October, we saw a different picture (right, top). Blue indicates ClO abundances below our detectable level. While the northern part of the planet looked pretty clean in regard to ClO in early October, by the middle of December we started seeing ClO signals that were definitely above our noise level. By the middle of January (bottom), our observations gave quite a spectacular picture: about the same abundances of ClO that we had seen in the ozone hole over Antarctica and here occurring over populated areas of the planet. Aircraft measurements had been done before in the north, and we knew that there would be enhanced ClO, but we were very surprised to see so much over such a large area. Now, the first thing we do when we see something like this is to make sure it's real. Examining the raw signals from the instrument made it clear that the abundance of ClO over Moscow on January 11 was comparable to that over Antarctica in the depths of the ozone hole.

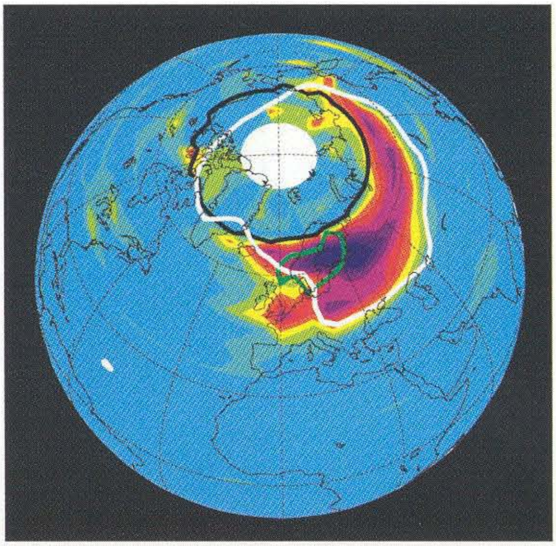
Why was the ClO along one side of the planet in the north on January 11? Why was there

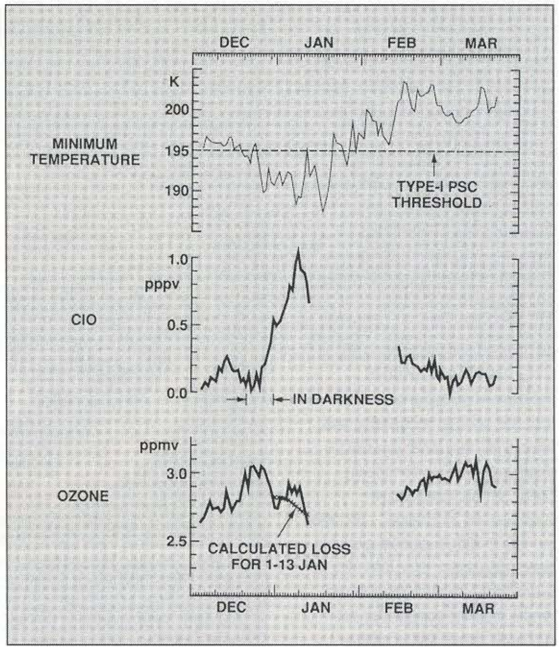




The maps above show lower stratospheric ClO measured on October 2, 1991 (top), December 14 (center), and January 11, 1992 (bottom) in the northern hemisphere.

Aircraft measurements had been done before in the north, and we knew that there would be enhanced ClO, but we were very surprised to see so much over such a large area.





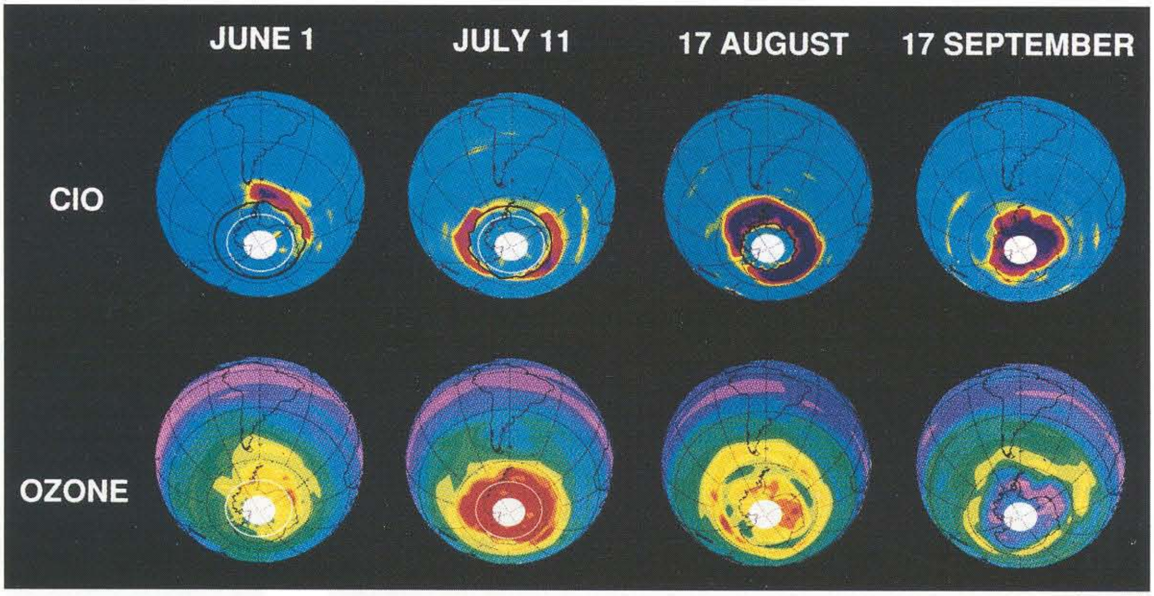
Right: In the same January 11, 1992 map of ClO as on the previous page, the black contour marks the edge of daylight, poleward of which the ClO is expected to be transformed into ClOOCl. The green contour indicates where it was cold enough for polar stratospheric clouds to form, a condition ripe for conversion of chlorine to ClO. This daylight portion coincides with the highest levels of ClO. The white contour indicates the approximate boundary of the Arctic vortex, which serves to contain the ClO.

Far right: A comparison of temperature, ClO, and ozone in the 1991-92 northern winter vortex shows clear correlations between the three. The dashed line is the threshold for the polar stratospheric clouds. As temperatures dip below that line, ClO shoots up, and ozone decreases. The crosshatched line indicates the ozone loss calculated from the ClO; the gap in the two lower curves represents the satellite's switch to the southern view.

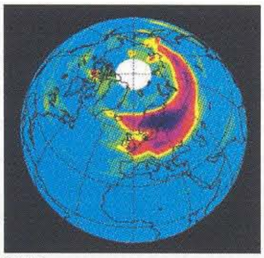
nothing nearer the pole? (Our satellite can't see higher than 80°, but the ClO clearly ends south of that.) We can now explain this. The black contour in the map above marks the edge of daylight; poleward of that contour is winter darkness. Theory predicts that ClO goes into the ClOOCl form in darkness, which explains the ClO decrease toward the north. The green contour marks out the general region in which the temperatures were cold enough for the polar stratospheric clouds to form and trigger the conversion of chlorine to ClO—and that's just where we see the largest abundances of ClO. Another important contributor to this distribution pattern is air motion. The white contour roughly marks the edge of the Arctic vortex, a swirling mass of air about the size of Asia. The air moving around inside this vortex is pretty much contained (just *how* much is currently somewhat controversial), and inside it much of the chlorine in the atmosphere has been converted to the reactive forms, which include ClO.

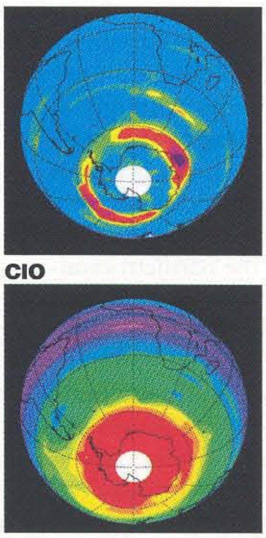
We might expect to see ozone depletion where we see all this chlorine monoxide, but when we compare this picture with a map of the ozone in the same place at the same time, we actually see *more* ozone. It turns out that two processes are occurring in the vortex. One is the cold temperatures that can activate chlorine to ClO. The other is the descent of ozone-rich air from above. Ozone transported up from the tropics into the polar regions is apparently descending into the lower stratosphere, and a race is going on between the rate at which the ozone is being destroyed by the chlorine and the rate at which it's being replenished by descending air. It's interesting to look at the evolution of that situation over the winter. The chart above shows temperatures, chlorine monoxide, and ozone during the four-month period from December 1991 to March 1992. The white, dashed horizontal line in the top column is the threshold below which the polar stratospheric clouds can form, and you can see that it gets cold enough for them to form in the middle of December. When that happens, the ClO shoots up, and ozone, which normally increases at this time of year due to transport from the tropics, starts to decrease. There are fluctuations in the ozone decrease, and it's difficult to just pull out the ozone loss by itself. The gap in the curves occurs because we had to switch to the southward view—a very frustrating period for us, because we really wanted to see what was happening in the north. By the time we looked back north, the temperatures had risen above the cloud threshold. Nitrogen is expected to come out of the evaporating polar stratospheric clouds and quench the chlorine, and that's consistent with the observed decrease in ClO.

It is interesting to compare the levels of ClO and ozone in the south polar region on July 11, 1992, with those of its seasonal equivalent—the same "solar day"—on January 11 in the north polar region (opposite). Just as in the north, there is more ozone in this layer of the lower stratosphere in polar regions. At the same time, chlorine has already been activated to ClO. (These are the first measurements ever made of ClO in the south at this time of year.) We're



In the Antarctic winter of 1992, ClO was already at enhanced levels by the beginning of June and remaining high all winter. More ozone was coming in, however, in the early winter than was being eaten up by the ClO. By mid-August the ozone in the vertical layer shown here had begun to decrease, and the formation of the ozone hole had started.

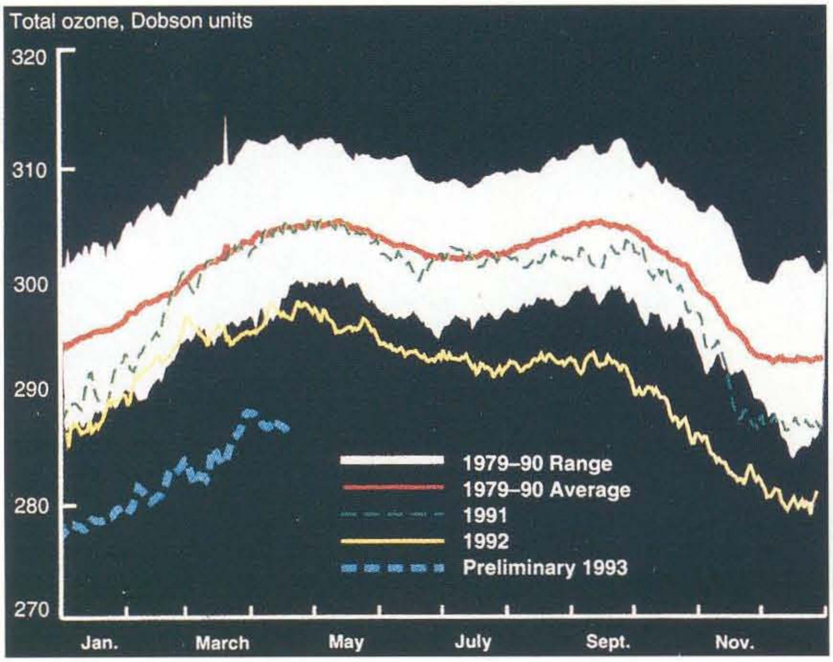




On the same "solar day"—January 11, 1992, in the north and July 11, 1992, in the south—there is more ClO (top) in the north than in the south. More intense planetary waves in the north shift the vortex off the pole more than in the south. This brings air processed by polar stratospheric clouds into sunlight, which is needed to maintain the high abundances of ClO. These planetary waves also cause a warmer winter stratosphere in the Arctic, which is the reason that the ClO doesn't stay around as long in the north and that no Arctic ozone hole has (yet) formed.

seeing the same processes in the south as in the north, although the distribution of ClO in the south is more symmetric about the pole, than in the north, where it's very asymmetric. The northern hemisphere has a much richer distribution of sea and land than the south, causing differential heating of the atmosphere and creating more intense planetary waves. Planetary waves, for example atmospheric wind patterns such as the jet stream, meander back and forth around the planet. The more intense planetary waves in the north shift the vortex off the pole, which is the reason for the asymmetry, and also cause a warmer winter stratosphere. The Arctic stratosphere generally tends to be about 10 degrees warmer than Antarctica, and that's why there's no ozone hole (yet) over the Arctic.

So these planetary waves are our friends; they can warm the stratosphere in the Arctic winter, thereby helping prevent an ozone hole from forming. Once the temperatures do drop below the threshold for polar stratospheric cloud formation, however, the planetary waves work against us by circulating the air out of the polar night into sunshine, which is needed to maintain the ozone destruction cycle. You can see that in the ClO maps: on the same "solar" day, much more of the ClO is present in the north than in the south. But in the north the ClO doesn't stay around as long, because it doesn't stay cold enough for very long in the winter. We can trace what happened through the southern winter in 1992 (above) and again, this picture came as somewhat of a surprise to us. The ClO reaches enhanced levels as early as June 1. The abundance



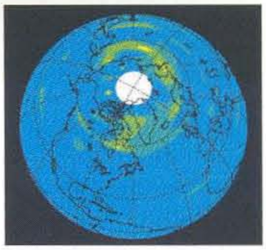
Above: In this chart made by James Gleason and colleagues from a decade and a half of global data from the Total Ozone Mapping Spectrometer, the white area represents the measured ranges of stratospheric ozone between 1979 and 1990, integrated over latitudes from 65°S to 65°N. The curve for 1992 and the segment for 1993 are clearly below the range of values measured in previous years.

of chlorine monoxide is enhanced all winter, but in the early winter ozone is still increasing, because more ozone is coming in than is being destroyed. By mid-August the ozone has started to decrease, indicating that chemical destruction of ozone has become the dominant process, and the ozone hole has started to form.

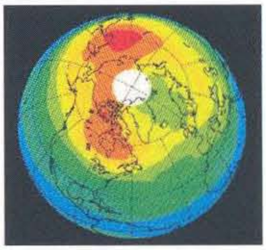
Our observations in the northern hemisphere continued through the winter of 1993. We saw enhanced ClO as early as December 4 over Russia. The vortex, elongated over Canada and Russia, stayed pretty much centered on the pole throughout December and January. On January 8, before we had to switch to looking southward, it was mostly over Canada. But when we looked back north in February, we saw a very different picture from what we had observed in February 1992. In 1993 ClO was enhanced through the end of February into early March. The temperature was a few degrees colder in February 1993 than in the previous year, and this made a tremendous difference because 1993 temperatures were below the point where the polar stratospheric clouds form. We also observed significantly less ozone in 1993 than in 1992, although this was distributed throughout the northern hemisphere. Over the northern hemisphere this past winter it was some 10–20 percent below what it was the previous winter.

The Total Ozone Mapping Spectrometer (TOMS), on another NASA satellite, has been making measurements for 15 years and can put the low 1992–93 ozone in better historical context than our satellite, which has been making measurements for only two years. Above is a

February 1992

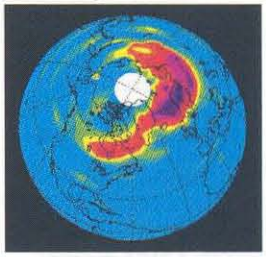


ClO

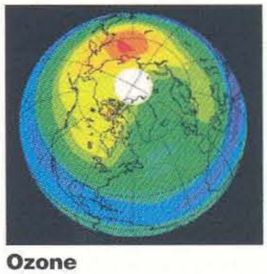


Ozone

February 1993

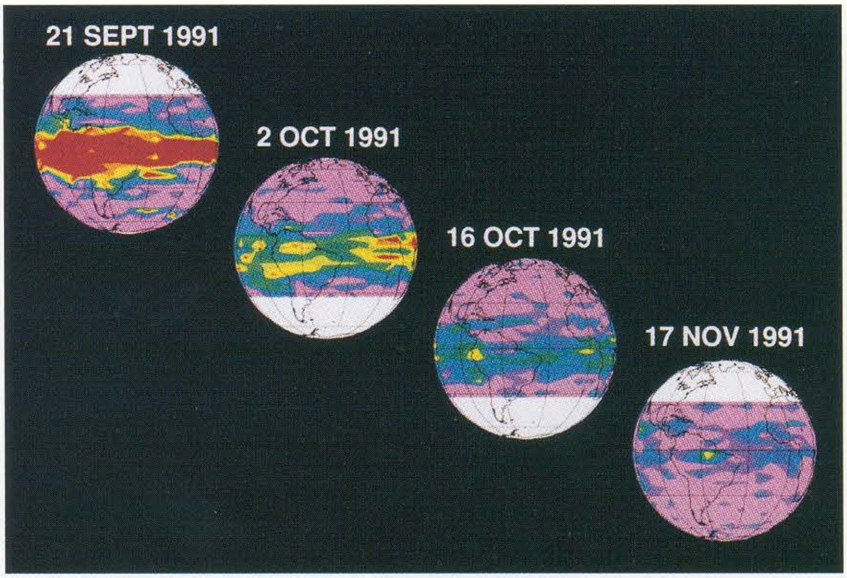


ClO



Ozone

Temperatures just a few degrees colder in the northern hemisphere in 1993 resulted in significantly more ClO than in the previous year. Ozone was observed to be significantly lower in 1993.



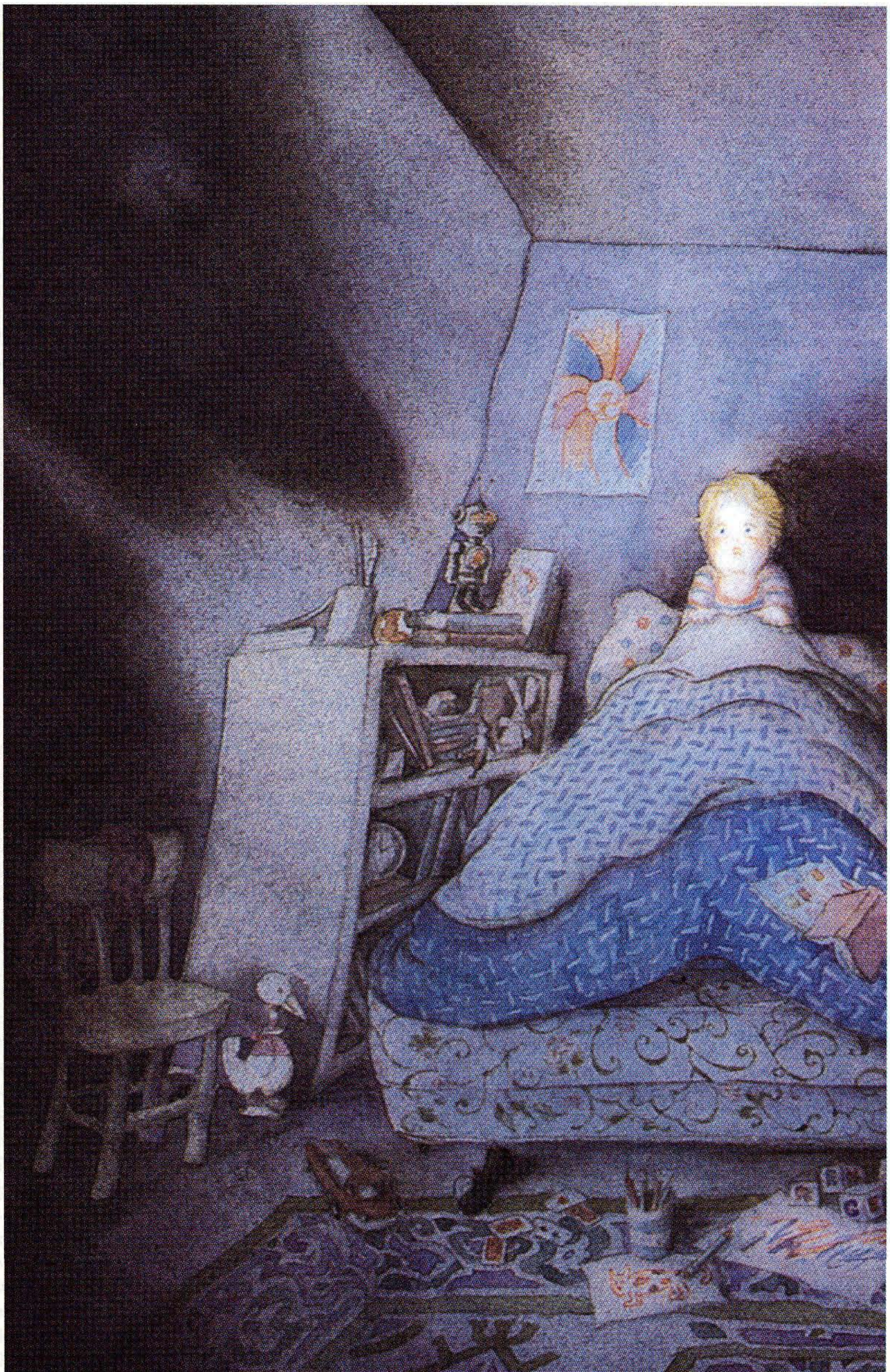
MLS measurements of sulfur dioxide in the atmosphere following the June 15, 1991, eruption of Mount Pinatubo show a significant amount still around in September and then diminishing. Red indicates 10 parts per billion, and purple indicates zero. Sulfur dioxide decay leads to formation of the small sulfate particulates where chemical reactions can take place that shift stratospheric chlorine toward ozone-destroying forms.

plot of the yearly cycle of total global ozone between 65° south and 65° north from TOMS, made by James Gleason (of NASA's Goddard Space Flight Center) and colleagues. The white area indicates the extreme ranges of ozone for the 11-year period between 1979 and 1990. The yellow curve is 1992, and the blue curve segment is 1993. In 1993 ozone is at unprecedented low values when viewed throughout the world. We know that part of the cause is the chlorine activated by polar stratospheric clouds. We are not sure of all the reasons for ozone being low throughout the hemisphere, but Mount Pinatubo, the Philippine volcano that erupted on June 15, 1991, is thought to be a contributor. Measurements showed that while Pinatubo did not inject a significant amount of chlorine into the atmosphere, the volcano did inject a lot of sulfur dioxide (SO_2) into the stratosphere. In the maps above of SO_2 from our instrument, red indicates high sulfur dioxide, about 10 molecules per billion. It's estimated from the TOMS data that Pinatubo originally injected about 20 million tons of sulfur dioxide into the stratosphere. The SO_2 gradually decays as it forms sulfuric acid, which coagulates into tiny particulates called aerosols. Laboratory measurements have shown that heterogeneous chemical reactions can take place on particulates of this sort and can shift the balance of chlorine in the atmosphere toward forms that are detrimental to ozone. So we expect that Mount Pinatubo contributed to the low values of ozone this past winter. Pinatubo would not, however, have been a problem for ozone, had not the stratosphere already been loaded with chlorine. It's the chlorine that's already there, which can be converted on these Pinatubo particulates, that causes the problem.

Discoveries of qualitatively new effects that destroy ozone do not appear to be slowing down.

Now we *hope* we know everything about the processes that deplete ozone, but it's really not certain how much of the whole picture we have at the present time. Discoveries of qualitatively new effects that destroy ozone do not appear to be slowing down. But even though there's still a lot of uncertainty, there is also a much greater sensitivity now to the planet's problems. More effort is being put into understanding the Earth and what we're doing to it. For our own part, an enhanced follow-on experiment is being planned for long-term measurements on NASA's future Earth Observing System (EOS), both to continue studies of ozone depletion and to measure parameters important in climate change. □

Joe Waters is a senior research scientist at JPL, where he holds joint appointments in the Earth and Space Sciences Division and in the Observational Systems Division. He received his PhD in electrical engineering in 1970 from MIT, where he also earned his BS and MS degrees. After working as a research associate at MIT (where he analyzed microwave data from the first Earth-orbiting microwave spectrometer), Waters joined the JPL staff in 1973 to establish a capability in microwave remote sensing there. He led teams that first detected carbon monoxide in Earth's mesosphere and in the atmospheres of Venus and Mars, and he has twice been awarded NASA's Medal for Exceptional Scientific Achievement. This article was adapted from Waters's Watson Lecture on May 19, 1993.



New Light on the Nature of Darkness

Dark is, in fact, an altogether more interesting character than is Light—at least light as most people understand these two characters.

by H. Jeff Kimble

Picture and text from the book *Alfi and the Dark* by Sally Miles and Errol Le Cane. Copyright 1988. Published by Chronicle Books, San Francisco.

In a book called *Alfi and the Dark* that my daughters Katie and Megan have generously lent me, the hero, a young man called Alfi, has a dialogue with the Dark. Although this is a child's story, it would be hard to find a better book in Millikan Library's physics collection to introduce my subject.

"Alfi was lying asleep in his bed
When he suddenly woke with a thought and he said,
'If I switch on the lights I'll be able to see
But where will the Dark go? Where will it be?'"

Alfi's question, in fact, is one of the central themes of my story. What is darkness? Where does it go in the presence of light? A long dialogue ensues between Alfi and the Dark, and each comes to understand the other somewhat better. Alfi learns that the Dark isn't such a happy fellow. Indeed,

"Dark felt so lonely. Dark felt so sad,
As he thought of the fun and the friends that Light had.
Wherever *he* went, people seemed to be scared.
He wanted a friend, just *someone* who cared."

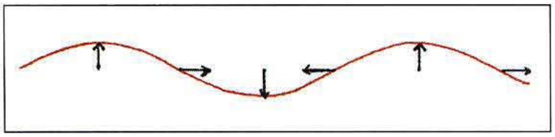
In the end they become friends, and Dark reveals his secret to Alfi.

"Dark was so happy he laughed with delight.
'Now, I'll tell where I go when you switch on the light.
The answer is simple and you'll be amazed—
I NEVER GO ANYWHERE!' Alfi was dazed."

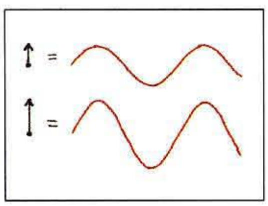
In the spirit of this book, my purpose here is to convey something about the modern view of darkness, and in the process to avoid Alfi's state of bewilderment at left.

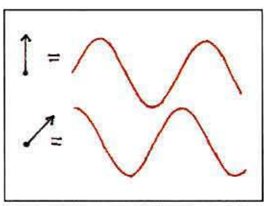
The objectives are really twofold: First, to convince you that Dark is, in fact, an altogether more interesting character than is Light—at least light as most people understand these two characters; and secondly to tell you about the activities of the "Friends of Darkness"—that is, the graduate students and senior scientists in the quantum optics group here at Caltech. The experimental results I'll tell you about are really due to their hard work. In addition, I should note at the outset that the conceptual foundation for much of this research was laid by Caltech's Carlton Caves [PhD '79], visiting associate in physics; Kip Thorne [BS '62], Feynman Professor of Theoretical Physics; and Ron Drever, professor of physics; and by their collaborator (and frequent visitor to Caltech), Professor Vladimir Braginsky of Moscow State University.

Since light is fundamentally a wave phenomenon, we should get straight a few basic concepts about waves. Imagine that you're sitting on a raft in the ocean. As the waves pass, you will bob up and down. Instead of being waves in water, light is an oscillation of the electromagnetic field, so that if you were an electron immersed in the field—that is, if you had a charge—you would bob up and down as the light wave goes by. A raft in the ocean bobs every several seconds. By contrast, an electron bathed in red light oscillates with a frequency of 5×10^{14} cycles per second—roughly a million billion times per second. And while the distance between crests of ocean waves —the wavelength—might be a few meters (a dozen feet or so), red light's wavelength is only 6×10^{-7} meters or 6,000 Ångstroms (about



Above: How a rotating arrow can represent a wave.
Below: Changes in amplitude change the arrow's length.
Bottom: Changes in phase tilt the arrow.





2/100,000ths of an inch, or roughly 100 times smaller than the thickness of this page). So visible light's wavelength is very short and its frequency is very high, which leads to some technical problems that make the experiments I'll be describing somewhat tricky to do.

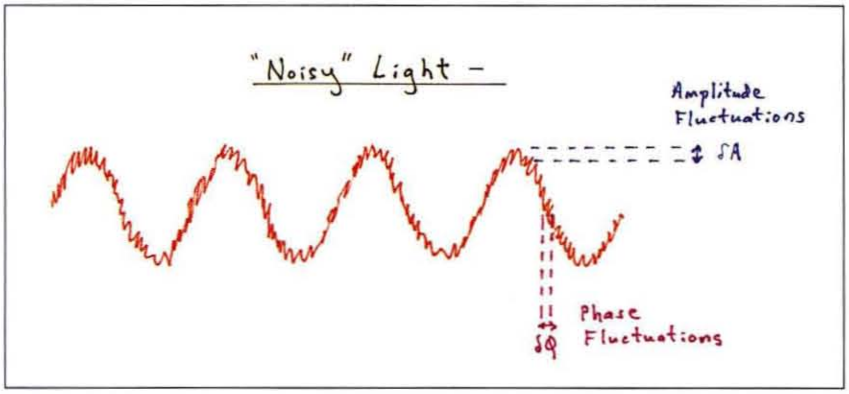
Now, just as two variables, position and velocity, are required to describe the motion of a person, we require two variables to describe light—or, in general, any wave. These two variables are amplitude and phase, and they form the basis for our discussion of the physics of light. The amplitude of a wave is simply the height of the wave's crests, which translates into "brightness" for light. The phase specifies the time (or distance) between zero crossings—the points where the wave's amplitude is zero—and is thus related to the wave's frequency.

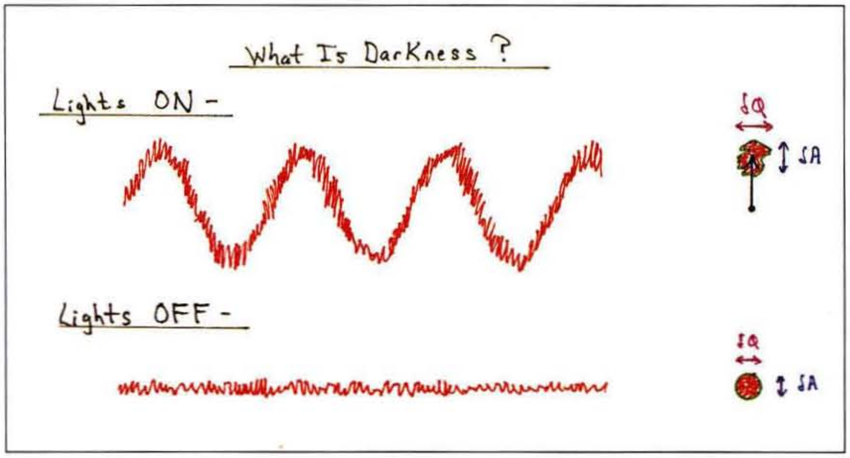
Now, with the objective of creating a more precise and powerful language to talk about light, let me get rid of the wave altogether and replace it by an arrow that rotates like the hand on a clock. The arrow's frequency of rotation represents the wave's fundamental frequency of oscillation. So, when the wave is at its peak, the arrow points to 12 o'clock, as shown above. As the wave's amplitude comes down to zero, the arrow rotates to 3 o'clock. At the wave's trough, the arrow is at 6 o'clock, and when the wave comes back through zero amplitude, the arrow reaches 9 o'clock. Now I don't really want to try rotating the arrow at the frequency of red light, so let's sit still—so that the arrow is stationary in our frame of reference—and assume that the world is instead rotating around us at 5×10^{14} times per second.

Translating our two variables for light into arrow language, we see that the arrow's length gives the light's amplitude, and the arrow's orientation gives the phase. If we increase the light's intensity, the arrow gets longer. If we instead change the light's phase, we tilt the arrow—that is, imagine two arrows spinning at the same rate, and hence fixed in our rotating frame of reference, but with one arrow tipped relative to the other; this is a difference in phase. Hence changes along the length of the arrow are amplitude changes, while deviations of the tip of the arrow perpendicular to its length (with its tail pinned down) are phase deviations.

Anyone who's tried to bodysurf knows that waves have certain irregularities. If you plotted the amplitude of successive ocean waves against their arrival time relative to the preceding wave, you'd find a spread of points clustering around the average wave amplitude and average time between waves. Likewise, light waves—even from a laser—are not perfectly regular, either. There are slight fluctuations in both amplitude and phase for any beam of light. Physicists call these fluctuations "noise" to indicate their random character, and to represent this noise, we "fuzz out" the tip of our arrow, so that its exact length (amplitude) and angle (phase) are now uncertain. I'll call this region of fuzziness a "noise blob." The larger the blob, the noisier the light.

I now want to turn to the fundamental rules and regulations specific to the electromagnetic field—that is, to light. What does physics have to say about the intrinsic amount of noise in a light wave? Nature's rule is simply that the product of the blob's noise in the amplitude dimension times the noise in the phase dimension has a minimum value set by Planck's constant. That is to say, because light is a quantum field, our noise blobs must have a minimum area. This is the Heisenberg uncertainty principle for light—the amplitude and phase of a beam of light cannot be precisely determined simultaneously, even in principle. (Heisenberg formulated the uncertainty principle to explain the quantum behavior of atoms and electrons; it is a direct and unavoidable consequence of the quantum theory.) Note that Planck's constant is a fundamental constant of nature—it sets the scale for the "graininess" of the atomic world. Light is quite remarkable in that, as we will see, this fundamental graininess leads to fluctuations in amplitude and phase that can have import, not only at the atomic level, but also in our macroscopic world. It is worth emphasizing that the fluctuations demanded by quantum mechanics are intrinsic





How noisy light becomes noisy darkness. Top: A light wave has amplitude fluctuations, symbolized by δA, and phase fluctuations, symbolized by δΦ. Bottom: When you turn off the wave or shrink the arrow's length to zero, the noise remains.

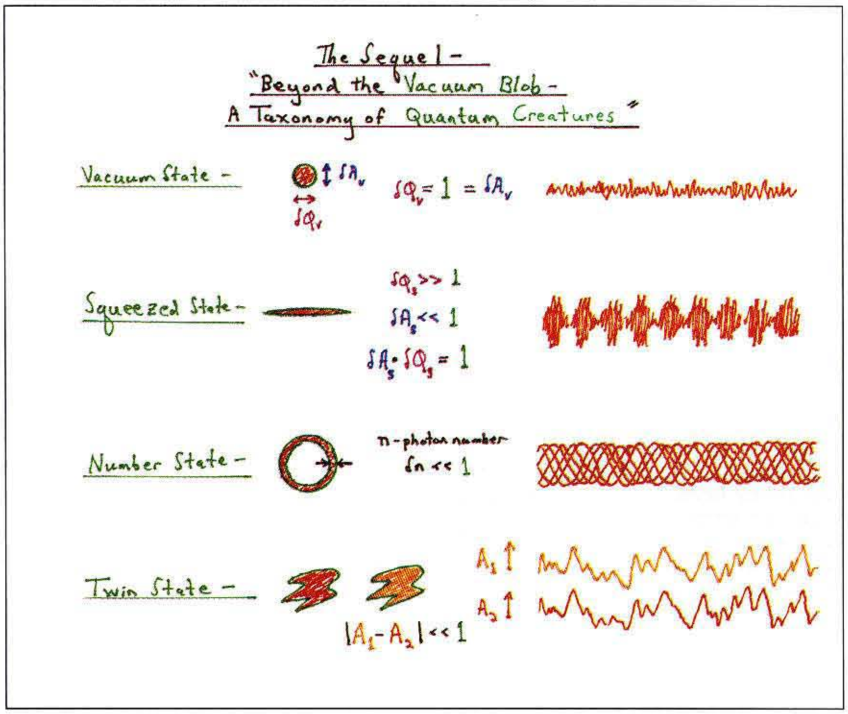
One cannot in principle turn off the quantum noise as well, so in fact there is something to nothing.

and fundamentally unavoidable. Hence the slogan, "Quantum mechanics—It's not just a good idea; it's the law!"

So now that we know something about what light is, we can talk about what darkness is. In terms of our picture of light as an arrow with a quantum noise blob on its end, we simply shrink the arrow's length to zero, leaving only the blob. Thus zero isn't really zero; it is zero plus or minus the noise of the residual quantum blob, as set by Planck's constant and as demanded by the Heisenberg uncertainty principle. An electron still feels a noisy electromagnetic field when the lights go off. Alfi knows this noisy field as his friend, Dark. A physicist knows it as the quantum vacuum state. It's nothing. It's what is left when the arrow—the coherent amplitude of the quantum field—is turned off. But one cannot in principle turn off the quantum noise as well, so in fact there is something to nothing. Note that the vacuum noise blob is symmetric with respect to amplitude and phase fluctuations; any direction is equivalent to any other. Instead of dragging some cumbersome dimensions along, I'll assign the vacuum state's fluctuations a size of "one," in terms of some arbitrary unit. Thus darkness is really a circular quantum blob of radius one.

There is lots of evidence that these vacuum blobs are real. I'll mention two pieces, both of which have to do with the theory of quantum electrodynamics that Richard Feynman, Julian Schwinger, and Shinichiro Tomonaga pioneered in the late 1940s. The first piece of evidence is atomic. Consider the simplest atom—an electron orbiting a proton—to which nature inevitably adds a vacuum blob. (The vacuum fluctuations are everywhere!) Two funny things happen. One, the atom gets measurably bigger—by about one part in 100,000—because the electron is being jiggled by the fluctuations of the vacuum field. This is called the Lamb shift, named after Willis Lamb, Jr., who shared the Nobel Prize in physics in 1955 for the phenomenon's experimental discovery in the hydrogen atom. The other is that the atom spontaneously emits light because its otherwise stable excited state becomes unstable due to the inane and incessant noise of the vacuum. Sodium-vapor lamps glow orange-yellow because the sodium atoms in an excited state decay to the ground state, and that decay is caused by the vacuum jiggling the electron in a way that is perfectly calculable, and well-confirmed by experiment. The second piece of evidence is visible on a larger scale, and can be seen by holding two metal plates very close together. Even though there is nothing between the plates except the vacuum, one finds that the

In science as in Hollywood every successful story has a sequel, and I assure you that the story of the vacuum state has been very successful—in fact, a smash hit—through the past several decades.

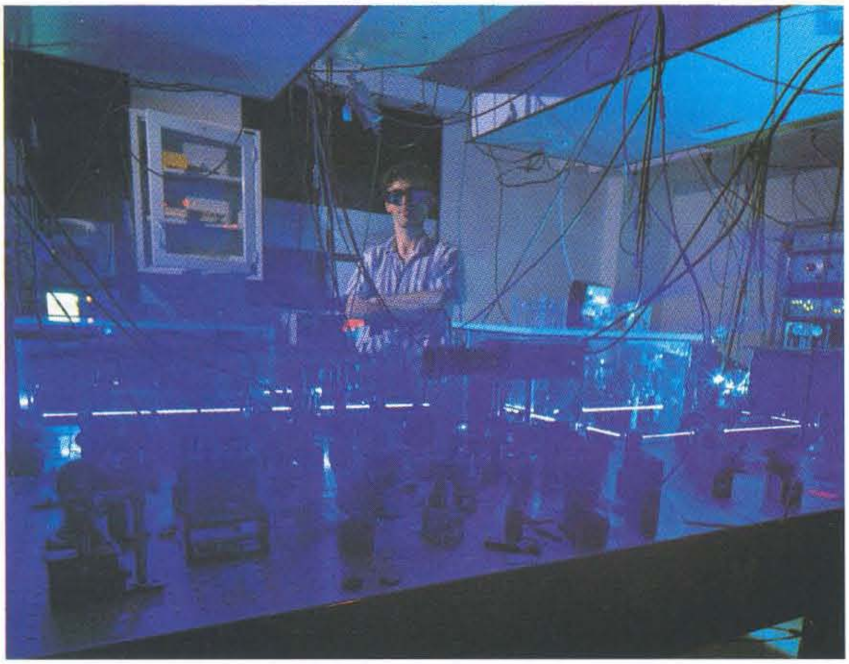


Each column contains (from left) the name of the blob, a portrait of its fuzzball in terms of amplitude and phase, its mathematical description, and the shape of its wave. A stands for amplitude, Φ is phase, n is the number of photons, and δ is the change or uncertainty in the variable.

plates actually attract each other with a tiny force—the Casimir force—resulting from the expulsion of the vacuum energy that's stored between those plates.

With the knowledge that darkness is the vacuum state packaged in quantum blobs, you might be tempted to think the story's over. It's not. In science as in Hollywood every successful story has a sequel, and I assure you that the story of the vacuum state has been very successful—in fact, a smash hit—through the past several decades. The sequel—Beyond the Vacuum Blob—is really what our research in the quantum optics group is about. Recall that our original vacuum state is a rather undistinguished circular blob. However, the only requirement of quantum mechanics is that a noise blob have a constant minimum area. The blob's dimensions along different directions don't have to be constant. For example, if we flatten the vacuum blob along the amplitude axis, the blob bulges out along the phase axis and the area remains constant. In other words, if the amplitude fluctuations get very small, the phase fluctuations must get very large. We call this a squeezed state. There are other ways to preserve the area. We could cut a hole out of the middle of the blob, and then stretch it into a thin ring, or annulus, of equal area, whose thickness in any direction is much less than the vacuum blob's diameter. The number of photons in this state is precisely known, but their phase is undetermined, as indicated by the circular symmetry of the blob. This is called the number state. And we need not restrict ourselves to the topology of only one blob—we can actually talk about two or more blobs at once. The laws of quantum mechanics demand that each blob have some minimum area, but the difference between the fluctuations each area represents—that is, between the "shapes" of the blobs—can be arbitrarily small. This is called the twin state because the blobs are, for our purposes, identical twins. The common ingredient in all these states is that some measurable aspect of the various blobs drops below one while some other aspect increases, keeping the area constant. (Remember, one is the size of the vacuum.) So, from the restricted viewpoint of the shrunken dimension only, we're making electromagnetic fields that have smaller fluctuations—less noise—than even the darkness of the vacuum state.

Moving to a specific example, one can now ask, "How do we squeeze darkness?" We start with a vacuum state—and I should remind you, in this day of conservation, that the vacuum is an unlimited, inexhaustible natural resource. We



The light coming into our detector is four times darker than the darkness that the detector would see if it viewed empty space.

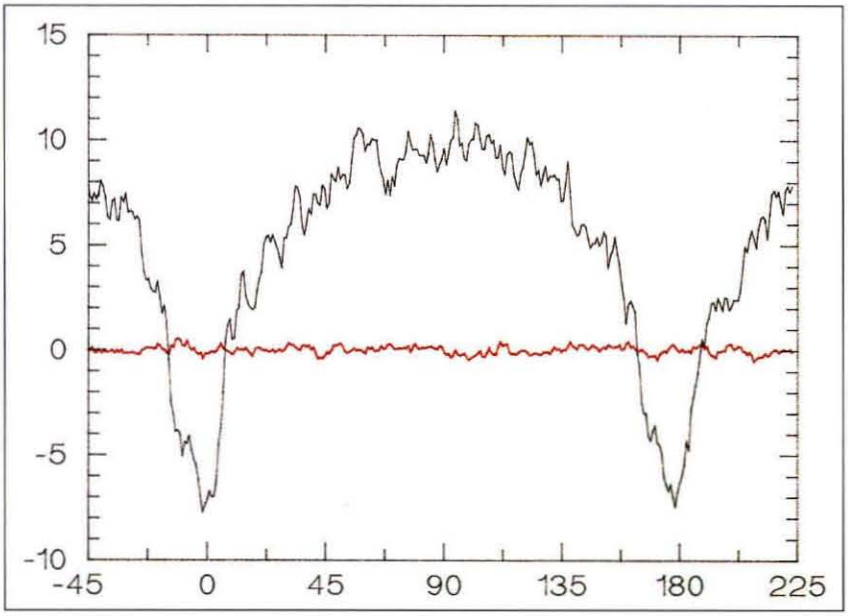
Grad student Nikos Georgiades and the darkness-squeezing factory. The blue lasers in front of him feed into a potassium niobate crystal, where the blue photons fission into squeezed red ones. The squeezed photons emerging from the crystal are actually in the infrared region of the spectrum, and can't be seen. The dark shapes in the foreground are a part of the interferometer that they live in.

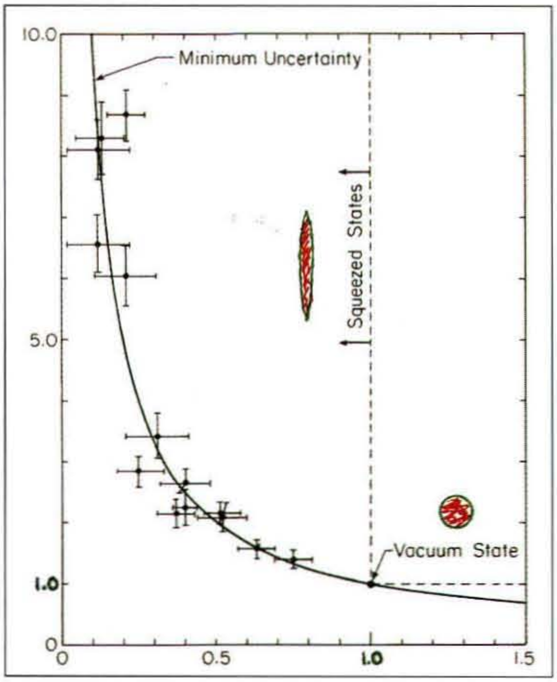
send this vacuum state into our squeezing factory, which is the elaborate arrangement of lenses, prisms, and mirrors shown above. Of course, we have to be very careful in our choice of a squeezing machine. We have to somehow "squeeze" the vacuum without "touching" it—what I call a Platonic squeeze. We can't touch it directly because, after all, it's the vacuum, which is to say it's nothing at all. And once an apparatus has touched or interacted with the vacuum an unacceptable contamination usually results, because at the quantum level, macroscopic beings like graduate students are fairly shaky entities that impart their own uncorrelated fluctuations to the vacuum. There's no easy way to do Platonic squeezing in a satisfying manner, nor is there an easy way to explain it. The process that we use most is called "photon fission," in which a photon of blue light goes into a special "nonlinear" crystal and splits into two photons of red light. The law of conservation of energy must be obeyed, so the sum of the two red frequencies equals the blue frequency. This process doesn't occur to any significant degree in free space, but there are a variety of very interesting crystals, including the potassium niobate crystal that we use, that behave in unusual ways when illuminated. One of the seminal papers describing photon splitting was written by Amnon Yariv, Caltech's Myers Professor of Electrical Engineering and professor of applied physics, some 30 years ago.

Well, how do we squeeze without touching? There's a beam of blue light going into the crystal, but there's also a beam of red darkness, if you will—an initial vacuum state, pure and uncontaminated by the presence of red light—going into that same crystal. Into that red vacuum state, from the distant vantage point of the blue light, we take photons one by one from the blue beam and add them two by two to the red beam. As we do that, the initial vacuum for the red beam—its darkness, if you will—is turned into a squeezed state. And as we turn up the rate at which the photon pairs are added to the initial vacuum, the state is squeezed more and more into an ever thinner ellipse. (The process is mathematically identical to painting our circular vacuum blob on a rubber sheet and then stretching the sheet along one axis.) Surrounding the crystal is the actual apparatus for accomplishing this transformation. The apparatus is very complex—it looks like a kid went wild in a toy store and assembled the ultimate Lego set—because in essence we are trying to process the amplitude and phase fluctuations of a light wave (which is going up and down 5×10^{14} times per second) with a precision that is a small fraction of the size set by the vacuum blob. Therefore the entire apparatus, nonlinear crystal and all, is essentially a large interferometer whose arms are servo-controlled to keep the various waves in near-perfect alignment. In fact, the result of a lot of late-night effort, principally by associate scientist Eugene Polzik, is that we've been able to compress the vacuum state by a factor of four; that is, when measured along the squeezed dimension, the light coming into our detector is four times darker than the darkness that the detector would see if it viewed empty space.

Of course, the rules and regulations for quan-

Reprinted from Wu, Xiao, and Kimble, *Journal of the Optical Society of America B*, Volume 4, pp. 1465-75, October 1987.



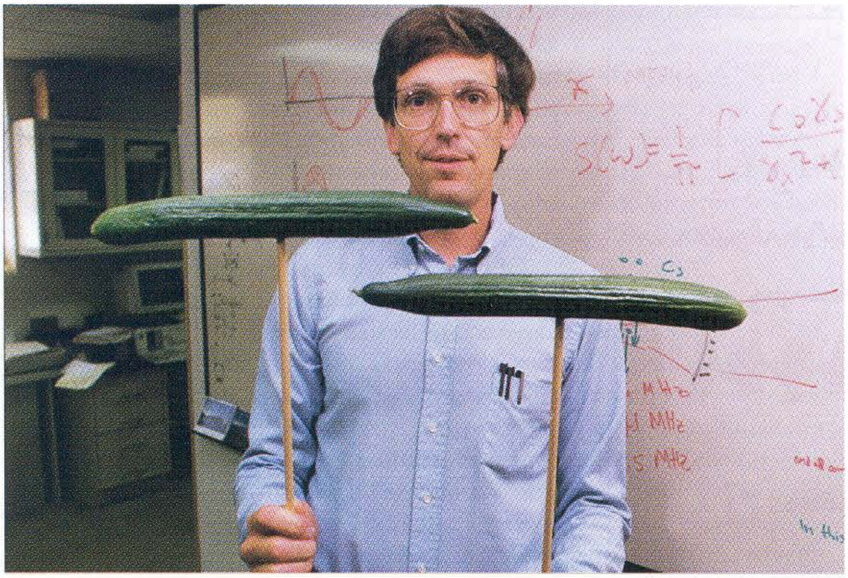


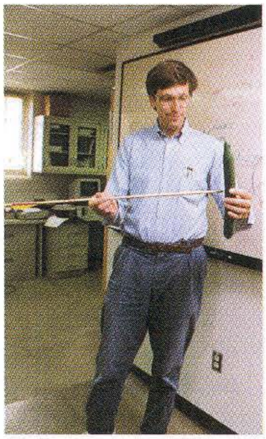
Above: A plot of noise versus the angle of tilt of the arrow for a squeezed fuzzball. The noise in decibels (vertical axis) is plotted logarithmically, so the noise of the unsqueezed vacuum is zero. The phase angle (horizontal axis) is plotted in degrees. The unsqueezed vacuum (red line) is equally noisy at all angles, whereas the squeezed fuzzball (black line) is much quieter than the vacuum at 0° and 180°, and noisier than the vacuum at 90°. Thus, to make a measurement using squeezed light, the detector would be locked at 0° or 180° in this case.
Right: The Heisenberg uncertainty principle for light. Uncertainty in phase (δΦ) is plotted on the vertical axis; uncertainty in amplitude (δA) on the horizontal.

tum fuzzballs require that when we reduce the noise in one dimension, it must bulge out elsewhere. This is shown in the graph above, which plots the amount of noise as a function of angle in the two-dimensional space of amplitude and phase fluctuations. Thus, if we take the valleys in that graph as representing the short (squeezed) axis of the noise blob and the peaks as the long (bulging) axis, then a plot of one versus the other should be a hyperbola. (Remember, our unsqueezed vacuum fuzzball is one unit in radius, and the uncertainty principle sets a lower bound for the area.) And so, independent of how complicated the experiment is or how complicated the theory is, in the end the best that quantum mechanics lets us do is the hyperbola labeled "Minimum Uncertainty" in the graph above right, which agrees with our data reasonably well. Note that the data points have no adjustable parameters; we measure everything in absolute terms. To the right and above that figure's dashed lines, which mark where each dimension of the noise blob equals one, lies the land of classical physics. If all light behaved like that, you wouldn't be reading this. To the left and below these lines is the land of quantum darkness.

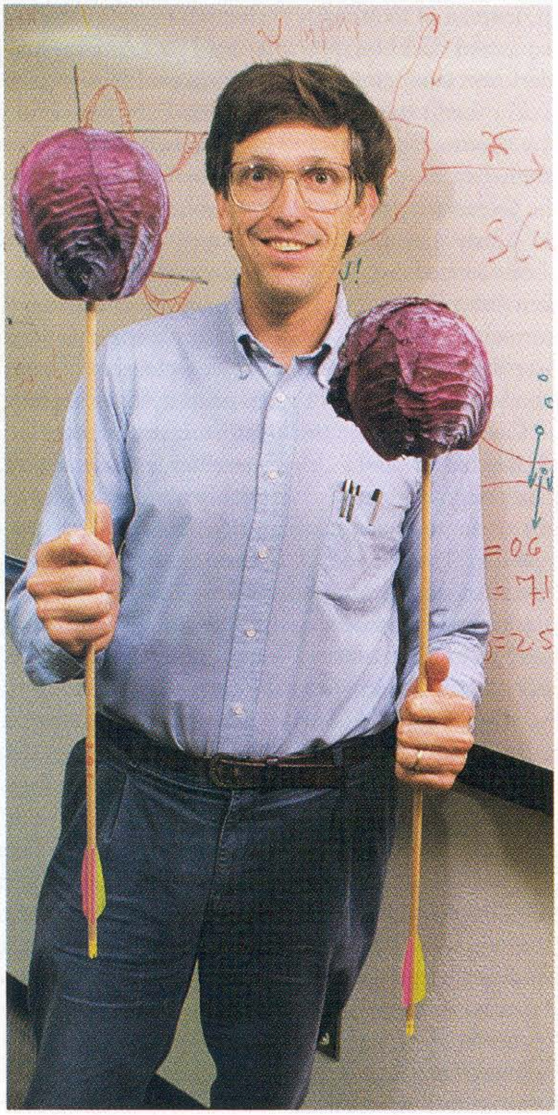
Now we're ready to think about making useful measurements with light. If I want to send a light wave to you, to talk to you on a fiber-optic telephone line, for example, what is the minimum modulation of the light—how much do I have to move the tip of the arrow—in order for you to notice any change? The classical answer is that the modulation can be arbitrarily small, because the position of the arrow's tip that represents the light wave is arbitrarily precisely defined. But this possibility is highly illegal because in quantum systems, the tip's exact location is no longer a defined quantity—it's just somewhere in a fuzzball of uncertainty. Heisenberg's uncertainty principle applies here to state that nature allows no naked arrows. I can represent this rule by impaling a quantum cabbage onto the point of the arrow. The laws of quantum mechanics say there have to be fluctuations —the arrow representing a perfectly smooth wave doesn't exist separately from the cabbage representing the quantum blob. In fact, to a physicist, a naked arrow is a much more heinous crime than is indecent exposure.

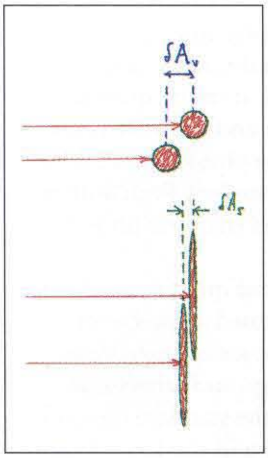
Since I can't remove the cabbage from the arrow, measurements involving a change of length of the arrow have to displace the arrow by an amount larger than the diameter of the cabbage—that is, of the vacuum fluctuations— in order to reliably discern any change at all. This displacement of the arrow by one diameter of the vacuum blob is the standard quantum limit for making measurements of the electromagnetic field. Over the history of the science of measurement, the standard quantum limit has stood as a seemingly impenetrable barrier, both conceptually and practically. And even making a measurement precise enough to approach the standard quantum limit in the first place is not trivial. However, in more recent times—over roughly the past 15 years—it has come to be appreciated that one can, in fact, do better than this limit. To do so, we squeeze our quantum cabbages into quantum cucumbers. Now a





A quantum cabbage (right) has to be moved by roughly its diameter in order to be sure of displacing the tip of the arrow hidden within it. The same applies to a quantum cucumber (above), but it can be moved less, as its diameter is smaller. In terms of quantum fuzzballs (below), this means that a smaller δA is measurable. Are you *sure* Steve Martin got his start this way?

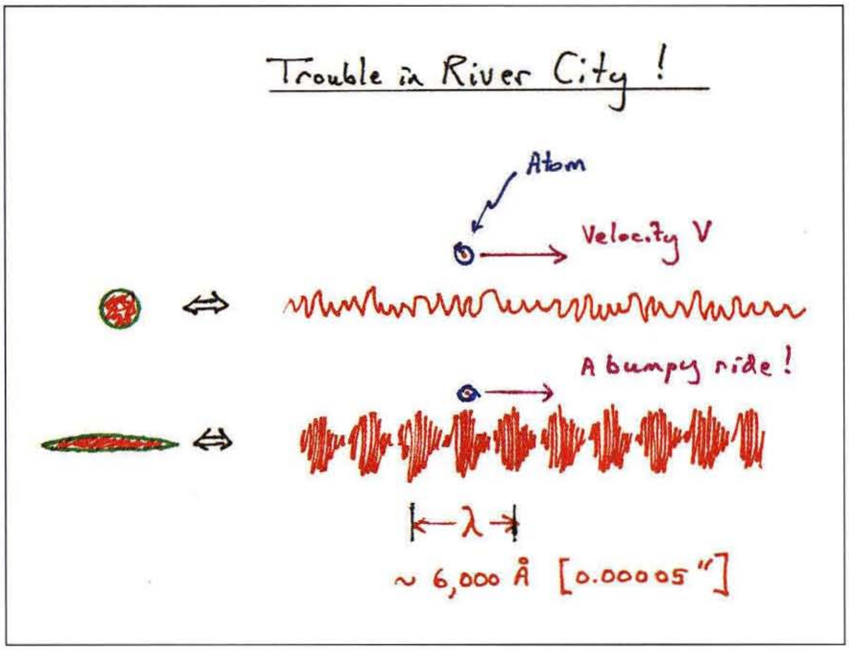




smaller displacement becomes discernible, because a cucumber is narrower than a cabbage—at least along its thin axis!—and we can make better measurements than had previously been thought possible.

It should be emphasized that, unlike cabbages, quantum entities are the same everywhere in the universe. While cabbages come in different sizes, the quantum fluctuations of the vacuum blob don't. Furthermore, these fluctuations are quite small. On a scale where a cabbage denotes a vacuum blob of radius one, the arrow's length, for even a laser of modest power, would be equal to the diameter of the earth.

These otherwise esoteric considerations of the quantum nature of light can be gainfully employed to detect a signal that couldn't otherwise be seen. Imagine that the quantum limit is a sea of fluctuations like the Pacific Ocean, and the signal we're looking for is the Hawaiian Islands. The Hawaiian Islands extend down to the ocean floor, but all we can see is what sticks up above the sea. If we aren't satisfied with this view, we could drain the ocean a bit. If we lower the ocean's level (the noise floor) by a factor of two, the Hawaiian Islands (the spectral peak we want to study) gets bigger relative to the noise by this same factor. That means we can see signals twice as small, or the same signals in half the time, as before. There is a caveat, of course, because this draining—which is really just a redistribution of quantum fluctuations—only happens along one axis. With the freedom to make this noise smaller comes the responsibility to make sure that we push the button that drains the ocean and



The problem of atomic motion: An atom traveling through an ordinary vacuum (top) has smooth sailing, but an atom moving through a squeezed vacuum (bottom) is in for a bumpy ride.

Electrons are reasonably intelligent. If one of them finds out that there's now a quiet dimension to its life where previously there was uniform noise in all directions (the usual vacuum state), it will try to live in the quiet dimension.

not the one that fills it. For if, instead of looking along the quiet dimension's squeezed darkness, we look at the squeezed antidarkness (the long dimension of our "quantum cucumber") the noise level goes up dramatically—by about a factor of ten in our most recent measurements, described below. That means the Pacific Ocean rises tenfold and the Hawaiian Islands vanish altogether.

One such experiment has been carried out here by Polzik and grad student John Carri. They were doing precision atomic spectroscopy—that is, detecting atoms by laser illumination. An atom has resonances—when you tickle it, it gets excited—so Polzik and Carri moved the laser's frequency around until they hit a resonance. The atoms—in this case cesium atoms in a vapor cell—absorbed light at the resonant frequency, betraying their presence. The particular technique we employed is called quantum-limited FM (frequency modulation) spectroscopy. It was pioneered by Drever here; John Hall, of the Joint Institute for Laboratory Astrophysics in Boulder, Colorado, and who was a Fairchild Scholar at Caltech in 1992; and a group at IBM. Our group added squeezed light to this technique, and has gone beyond the standard quantum limit by about a factor of two in spectroscopic sensitivity.

If we can use these "designer" fluctuations to probe atoms in new ways, can we also harness these funny fields to actually *change* the atoms? Remember, an atom coupled to a vacuum gives rise to the standard radiative processes—things like lasers, street lights, and interstellar nebulae. What if we instead couple the atom to a squeezed vacuum? Well, electrons are reasonably intelligent. If one of them finds out that there's now a quiet dimension to its life where previously there was uniform noise in all directions (the usual vacuum state), it will try to live in the quiet dimension. Indeed, there are stacks of theoretical papers indicating that atoms would behave in fundamentally different ways, if only we could couple them to the squeezed vacuum. Such coupling would affect all of traditional spectroscopy, as well as things such as how lasers work.

Of course, there are a few catches, at least one of which is the problem of atomic motion. Ordinary vacuum is structureless, so when an atom moves through it, the atom travels as though it's on a smooth, featureless road in North Dakota, as shown at left. On the other hand, if we use a squeezed vacuum, the atom's in for a bumpy ride—bouncing up and down over the spatially varying noise of the squeezed light. All the while, the atom's electron is trying to find the light's quiet dimension, which unfortunately changes every quarter of a wavelength—about every 1500 Ångstroms. One solution to this motion problem is to cool the atom's motion to almost absolute zero and to confine it to a distance much smaller than one-quarter of a wavelength. Graduate student Zhen Hu is doing such research in my group, trying to nail the atom down by using laser beams to build an atom trap. A trapped atom is also very cold, since temperature, on the atomic level, is really a measure of the atom's energy of motion. The photo opposite shows a cloud of cesium atoms cooled and trapped by laser beams. The cloud is about a millimeter—a twentieth of an inch or so—in diameter, and the atoms within it are cooled to within about 10^{-4} degrees Kelvin of absolute zero (–459° Fahrenheit). At the same time, we are working on ways to make clouds with fewer atoms, until we can eventually just trap a single atom. So we've almost got the atom nailed down, and once we do, we'll bathe it in the quantum quietness of squeezed light and see what happens. (Associate Professor of Astrophysics Ken Libbrecht [BS '80] and graduate student Phil Willems also have a laser cooling and trapping project on campus.)

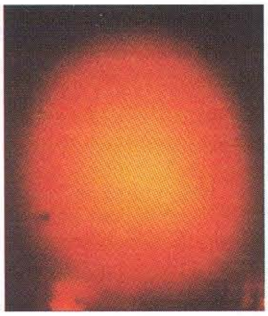
Returning to the theme of quantum measurement, my group has performed a number of measurements over the past six or seven years—spectroscopy, interferometry, and others—at levels of precision beyond the standard quantum limit. But how far beyond will the laws of nature let us go? In terms of our previous analogy, we've lowered the ocean by about a factor of two, but where, actually, is the bottom? As far as I know, there's no totally satisfactory theoretical answer to this question. To find out, we need to optimize our measurement techniques over all possible quantum blobs—all shapes and states, not just the few I've told you about—and over all possible measurement strategies. That's a difficult thing to do. After all, we're using 19th-century techniques—for example, interferometry—and late-20th-century light. Nonetheless, some important theoretical progress has been made in recent years, notably by Carlton Caves and colleagues.

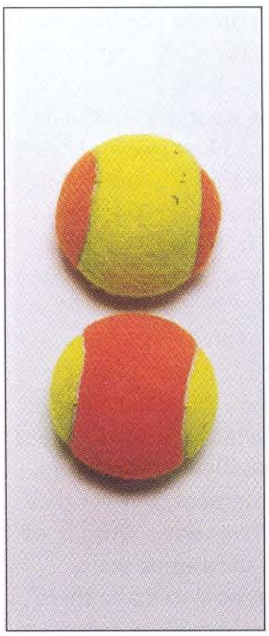
Apart from deep theoretical issues, there is a great deal of practical interest in manipulating the fundamental quantum fluctuations of light for such things as spectroscopy, quantitative analysis, and interferometry. Applications range from things on the scientific frontier, like the LIGO (Laser Interferometer Gravitational-Wave Observatory) program here, to more mundane things like the new aircraft-navigation systems, which use a laser gyroscope working near the standard quantum limit to sense rotation.

At this point, we might stop and ask, what does this all mean? What are these quantum blobs, really? This is, in fact, a very difficult question to answer. To avoid having to answer it myself, I will quote from *Dreams of a Final Theory*, by Stephen Weinberg, a Nobel laureate and one of the eminent scientists of this century. "A year or so ago, while Philip Candelas… and I were waiting for an elevator, our conversation turned to a young theorist who had been quite promising as a graduate student and who had then dropped out of sight. I asked Phil what had interfered with the ex-student's research. Phil shook his head sadly and said, 'He tried to understand quantum mechanics.'" Weinberg goes on to say, "I admit to some discomfort in working all my life in a theoretical framework that no one fully understands." The computational power of quantum mechanics is unquestioned. However, what it all "means" in any satisfactory sense is difficult to explain, even to oneself. Nonetheless, I'll try to illuminate some of the issues and conundrums in the following thought experiment.

Suppose I have a source that emits pairs of colored tennis balls, one to the right and one to the left, and detectors some distance away that catch the balls and register a reading of either red or green. The source always sends out correlated pairs of colored balls heading in opposite directions. Thus if I listed what each detector saw, the left detector would register a sequence of, say, red, red, green, red, green, and so on. And the right detector would register the opposite colors—green, green, red, green, red, and so forth. The question is, what inferences can I draw about the nature of these quantum blobs—here represented as tennis balls—as they propagate from source to detector, using the sequences recorded at the two detectors? For example, if I detect red at detector number one and green at detector number two, can I infer that a red blob actually traveled from the source to detector number one? Or, to paraphrase Einstein, "Do these blobs have any existence independent of one another?" Well, certainly they must. If one blob is just coming by me and the other is way over there in Andromeda, then surely nothing about what happens to this one can affect that one.

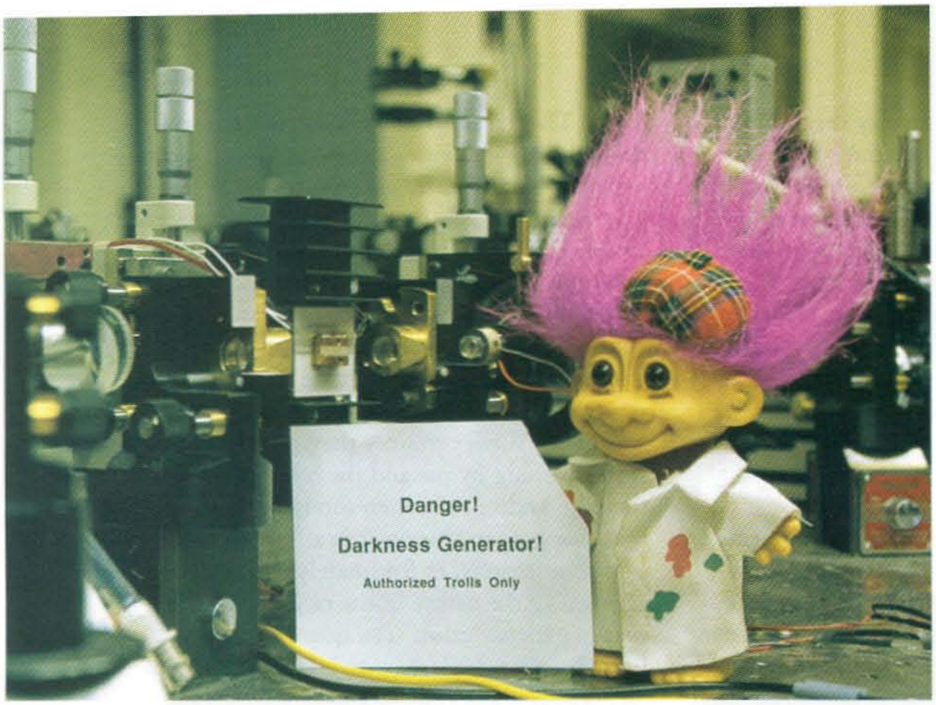
Top: A cloud of trapped cesium atoms.
Bottom: A pair of quantum tennis balls.





Unfortunately, or fortunately, this most sensible view of the nature of the physical world is not, in general, valid. The quantum world is indeed a strange place, with a large domain of exceptions to the rule of objective reality. It turns out that neither blob, for certain kinds of quantum systems, has a "color," where color is used metaphorically to refer to some property of the system in question, as for example its state of polarization. The "color" information is not carried by this blob or that blob, but rather resides in the correlations between the blobs. Either blob has the potential to be red or green—it's neither red nor green as it propagates, but somehow has the potential to be both colors at the same time. Hence physical properties for some microscopic quantum systems don't exist in the sense that I'd like to think that I exist. If you turn around, you can't see me, but I hope that I'm still here with an unchanged, definite set of properties. But for these quantum blobs, for these colored quantum tennis balls, color becomes well-determined—"exists," if you will—only when the blobs are detected. So a red click in my detector here, in some spooky way, means that the other blob must now be green, even if the detection events are light years apart. That's not a very comfortable thing, but that's the way it is. John Bell, who defined the limits of applicability of objective reality, called these correlations the irreducible nonlocal content of quantum mechanics. To paraphrase Bell, the speakable in quantum mechanics is the two detected sequences of reds and greens. The unspeakable, to which we are not allowed an answer in quantum mechanics, is the "real" color of one or the other blob as it propagates.

This is not a particularly comfortable situation, but is it refutable? No. A series of experiments by a number of groups, culminating in the work by Alain Aspect et al. in Paris, says that's the way nature is, like it or not. As for our own efforts in this regard, Zhe-Yu (Jeff) Ou—who has



How darkness squeezing *really* works. (Troll courtesy of Megan Kimble.)

We'd like to find ... systems that are continually evolving and interacting with their environment, but yet that are not describable in objective terms. And if we can learn how to do this on the atomic scale, eventually we'd like to learn how to make them into macroscopic objects big enough to campaign for office.

since left for a faculty position at Purdue University in Indianapolis—and grad student Silvania Pereira have built an apparatus that makes correlated quanta in two spatially separated beams. They've carried out several experiments with this system over the last year and a half, but the one that I'll describe is related to quantum communication.

Imagine that I'm trying to send you a confidential message. Maybe it's about my bank account—a lot of such traffic is financial. Whatever it is, I don't want anybody to listen. Normally, my message would be encrypted in some code, such as the widely used Digital Encryption System (DES), that is nearly impossible to decode illicitly. Although such a code can be made extremely difficult to break in practice, nothing ensures that it cannot be broken in principle by some sufficiently clever person. One would like to protect these messages—not by my ingenuity or yours—but by the laws of quantum mechanics so that they are immune to interception in principle. So, by the process of photon fission, Ou and Pereira made two big, noisy quantum blobs that were arbitrarily large compared to the vacuum blob and that were quantum twins of each other. That is, their fluctuations in amplitude and phase were identical. Then, inside each one of these twins, we wrote a message so small that it was actually smaller than the vacuum level. The twins were then transmitted along different routes. Even if an eavesdropper detected one blob, the message was unrecoverable, because it was smaller than the standard quantum limit. Only when both blobs were detected and properly subtracted did the message emerge. Furthermore, if somebody did try to listen in, this interception would sound a burglar alarm, because detecting one blob destroys quantum correlations, and hence degrades the message to garbage.

Another experiment, which postdoc Olivier Carnal and grad students Robert Thompson and Quentin Turchette are pursuing, is difficult to describe, but the spirit is conveyed by comedian Robin Williams's line, "Reality—what a concept!" The issue is again the nature of reality, but now for a quantum system that's continuously interacting—being "measured," if you will—by its environment. Such "open" quantum systems are both driven by, and decay into, their surroundings, and are the basis for the phenomena that we know on a macroscopic scale. For any given open quantum system, there are many different measurements that we in the external world could choose to make: How many photons are coming out, and how are they distributed in time and space? What do their quantum fuzzballs look like? We could choose to ask a series of such questions by making a series of different measurements on the system. The $64 question —the sum of all questions—is whether there are systems whose "reality" is conditional upon the questions that we ask of them. We'd like to find such open quantum systems for which this is so—systems that are continually evolving and interacting with their environment, but yet that are not describable in objective terms. And if we can learn how to do this on the atomic scale, eventually we'd like to learn how to make them into macroscopic objects big enough to campaign

The Friends of Darkness, seen on the roof of Bridge Laboratory in the unaccustomed light of day. From left: Hu, Carnal, Thompson, Pereira, Georgiades, grad student Hideo Mabuchi, Turchette, and Willems.

for office.

The particular system that Carnal, Thompson, and Turchette are looking at consists of a pair of parallel mirrors facing each other, some 300 millionths of a meter apart, with a stream of cesium atoms passing between them, like a single lane of cars between concrete dividers. The spacing between the mirrors is an exact multiple of a wavelength at which cesium atoms absorb and emit light, and is precisely controlled to within about 10^{-13} meters, or one-thousandth of the diameter of an atom. The cavity formed by the two mirrors serves as a very simple system whereby photons from a laser perpendicular to the path of the cesium atoms can be strongly coupled to them. That is, the interaction between the photon and the cesium atom is much stronger than the dissipative forces that normally cause the photon to lose its coherence. Thus, the excited atoms are in a nonequilibrium steady state, not unlike living beings—they take in energy, they move around and do things, and eventually they dissipate and die. The coupling is so strong that a mere 0.2 photons will evoke a nonlinear response from an atom, and a paltry 0.06 atoms significantly alters the photon's behavior. Hence the escape of a single quantum into the external environment can have a profound effect on the system, even though it contains hundreds of atoms and photons. The strong coupling means that quantum events occur at a faster rate than dissipative events, and the atom-photon system thus has enough time for at least the possibility of leading a life of manifestly quantum dynamics before the grim reaper of dissipation enters. It is to this type of system that we are currently turning our attention in a quest to explore the exquisite interplay of the birth and death of quantum states for driven open quantum systems.

Finally, then, let me come back to where we began—back to Alfi's question. If I turn on the lights, where does the darkness go? I hope I've given some sense of the answer to this seemingly simple question. We now know that darkness is the blob of noise representing fundamental fluctuations in the electromagnetic field. To produce light, we just put that blob on the end of an arrow. What Dark said to Alfi is precisely correct, "I never go anywhere." The dark is still there when we turn on the lights; it's just sitting on the end of an arrow that represents the basic coherent amplitude of the light. I couldn't have told my children about the nature of darkness any better. In fact, I use that book to tell them what I do in the laboratory. We've also seen that there are destinations beyond darkness. For example, I've told you about squeezed vacuum and some of its applications, and about twin states. In general, I've tried to convey a feeling for light that's even darker than the darkness of the vacuum, and about the activities of a group in a "mad pursuit" of the science of darkness. Finally, I would invite everyone to enjoy the darkness, much as Alfi can with his new-found understanding. □

H. Jeff Kimble received his BS from Abilene Christian University in 1971, and his MS and PhD from the University of Rochester in 1973 and 1978, respectively—all in physics. He came to Caltech as a professor of physics in 1989 from the University of Texas, where he was the Richardson Regents Professor of Physics. Kimble's PhD thesis research represented the first observation of a nonclassical state of light, and the research group he established at UT was one of the first to explore the field of squeezed light and related nonclassical phenomena. This article is adapted from Kimble's recent Watson lecture, which combined quantum physics and laser science with elements of Gallagher's vegetable-imperiling stand-up comedy and a tennis clinic. In fact, Kimble's research group had so many quantum and classical tennis balls lying around after the lecture that they recently held the First Annual Quantum Optics Tennis Tournament. (The Forces of Darkness beat the Forces of Light, 7-5, 6-2.) Kimble's daughters Megan and Katie are six and eight, respectively.

Roundworm Cells and Cancer Genes

By studying this... simple, experimentally tractable micro-organism, my lab has been able to help out in the big problem of trying to understand what happens during cancer.

by Paul W. Sternberg

Graduate student Gregg Jongeward watches roundworms through a stereomicroscope, while his inflatable friend appears to be preparing for a doctoral candidacy exam. (Apologies to Edvard Munch.)

My major obsession in life is to understand animal development—how a single cell divides and generates the many specialized cells that form the adult organism. And, as it happens, by studying this process of development in a very simple, experimentally tractable microorganism, my lab has been able to help out in the big problem of trying to understand what happens during cancer. I'll start by introducing the current concepts of what happens in the early stages of cancer, and then I'll tell you about the roundworm we've been studying, and then at the end I'll bring it all together.

A cancer arises from cells that escape their normal growth control and divide continuously. Eventually the cells acquire the ability to invade surrounding tissue—that is, metastasize—or commandeer a blood supply, or both. Imagine a nicely organized tissue—say a layer of cells such as your intestinal wall. The cells are slowly dividing to replenish themselves. Say you get a mutation—a change in a gene in a particular cell that gives it different properties. That mutation, in some instances, might cause that cell to divide faster than its neighbors. Soon the faster-dividing cells are encompassing more and more of the layer. They start to take over, in other words. Then another mutation might cause the cells to grow even faster, and lose their ability to maintain their nice, sheetlike formation. They might start forming a lump. Then there might be a third mutation that divides even faster and has other properties, for example the ability to crawl around and invade nearby tissue. It's by a series of such mutations that most cancers progress. Typically, it's more than three mutations, and they don't happen very fast, which is why some tumors can take 10–20 years to develop. A "genetic predisposition" to cancer often means that the cells have one such mutation to begin with, which shortens the chain of mutations needed for the cells to become cancerous. Certain mutations make the cells pretty sloppy at replicating themselves, increasing the rate of mutation. Normal cells replicate their genetic material very accurately, so a mutation in the machinery that insures this accuracy would quickly lead to more mutations. A recently discovered colo-rectal cancer-predisposition gene might be of this type.

There are two kinds of genes that can mutate to cause cancer. Oncogenes—that is, cancer-causing genes—are one type. This class of genes was discovered about 20 years ago. An oncogene's normal function seems to be to stimulate cell growth and division, so that mutations activating these genes inappropriately would likely lead to cancer. The other kind of genes, discovered over the last 10 years, are called tumor-suppressor genes. These genes tend to inhibit cell growth and division. If such a gene is eliminated from a cell, that cell will grow and divide when it shouldn't.

To understand how these changes can affect a particular gene, we need to review how a gene directs the synthesis of a protein. Proteins are the building blocks of the cell—the structural components that form the cell's architecture, the enzymes that form the cell's machinery, and the messengers that regulate the cell's activities.

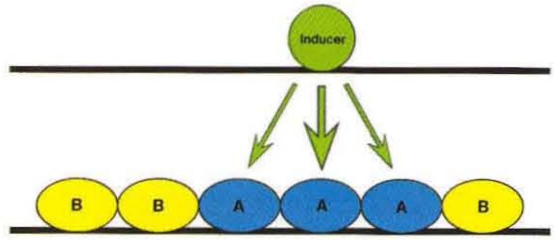
A protein consists of hundreds or even thousands of small building blocks, called amino acids, linked together like beads on a string in a very specific order. The genetic instructions of every organism are encoded in very long molecules known as DNA. Particular segments of that DNA, called genes, are transcribed and copied into another nucleic acid called messenger RNA. Each gene typically contains the instructions for one protein. The messenger RNAs are then translated into proteins by some very specific and exquisite machinery in the cell. The machinery is a complex of perhaps 50 to 80 proteins and several pieces of RNA. The machinery also does proofreading, making sure that each amino acid is put in the right order. The proteins then fold up and form three-dimensional structures determined by their sequence of amino acids, and these structures do the work of the cell.

Some mutations decrease or abolish a gene's activity. For example, the transcription of DNA into RNA could be blocked, or the translation of messenger RNA into protein could be blocked, or the folding of the protein could be abnormal, or the protein could be made but wouldn't work. Or the gene could just be deleted from the genome. Other mutations cause the protein to be more active than normal, or make the gene direct the synthesis of too much protein. All of these things occur in nature. So a mutation could inactivate a tumor-suppressor gene and prevent the synthesis of an inhibitor, which would lead to more cell growth and division, and lead to cancer. Or a mutation could activate or make more of an oncogene, leading to cell growth and division and cancer. Our task is to identify all these genes—and people think that there are at least 100 of them—and figure out what each gene's protein does, and how all these genes and proteins are linked together to form the circuitry that controls what the cell does.

A surprisingly high percentage of cells die during normal development—they either commit suicide or they're murdered.

The normal role of the genes that, when mutated, lead to tumors is to determine a cell's fate during development. A developing cell has to make many choices. It has to decide how many rounds of cell division to undergo—does it not divide at all, or does it generate a million progeny cells? If it divides, what kind of progeny does it produce—skin, nerve, muscle, liver, or what? Does the cell survive, or does it die? A surprisingly high percentage of cells die during normal development—they either commit suicide or they're murdered. And finally, the cell must choose whether to stay where it was formed, or to crawl to another location in the organism, like the neural-precursor cells that Associate Professor of Biology David Anderson studies

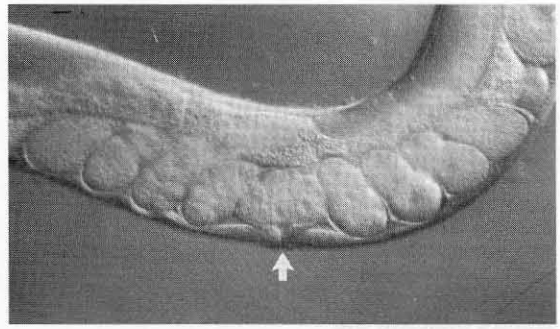
Cellular induction: Whether the individual cells in the bottom row become type A or type B cells depends on whether they are within range of a signal from the green cell above them. The black lines are generic tissue structures.



[*E&S*, Spring 1990]. The problem that I set out to study 10 or 15 years ago is: How are the instructions for the fate of particular cells coded in their DNA?

Now, in most organisms, what a cell does depends on signals from its neighbors. In the simplest possible case, consider a type A cell, colored blue in the drawing above. The fact that this cell is a blue A cell as opposed to a yellow, or B, cell, depends on a signal from a neighboring green cell, which I'll call an inducer cell. We can demonstrate this by surgically removing the green cell, and when we do, the cell that should be an A is instead type B. Or we can get rid of the A cell; then its neighbor, which is normally a B cell, becomes an A. So we conclude that the A cell becomes an A by virtue of the fact that it receives a signal from the inducer cell, and the B cell can't become an A because it doesn't get the signal. The signal is a chemical—usually a protein, in the examples that I've been studying—that is secreted, or released, from the inducer cell and interacts with a protein on the surface of the A-cell-to-be and directs its development.

The organism I spend most of my time working with is *Caenorhabditis elegans*, one of the nematodes, or soil roundworms. Nematodes are as common as the dirt under your feet—there are perhaps a hundred of them per cubic inch of soil—and they literally stick to your shoes as you walk through the grass. But you're not in a constant state of being grossed out by this because they're so small that they're almost invisible to the naked eye. At right is a worm in its normal habitat in the laboratory. It's crawling



Right: A full-grown, one-millimeter-long roundworm takes its constitutional on a petri dish. The wavy lines are tracks left by other worms. The dark blot below and to the worm's left is an egg; below that is a baby worm.
Far right: The vulva (arrow). The line of nine spudlike objects above and flanking the vulva are fertilized eggs. The dimples in the eggs are cell nuclei; thus the egg directly above the vulva has already divided into at least eight cells.

on a petri dish, in a slurry of the bacteria it eats. These small creatures have a number of advantages as lab animals. They're very easy to raise. They're also easy to handle—we can pick them up with very small, sterile platinum wires, and move them from petri dish to petri dish. And they grow very rapidly, going from an egg to an egg-layer in three and a half days. We get two generations a week for genetic studies, so we can do a lot of experiments. One worm on a petri dish will give rise to 300 progeny in, say, four or five days. Of course, there's a slight disadvantage in that you have to look at the worms daily to follow their growth, as opposed to most other organisms, where you can ignore them for a week at a time because things don't happen very fast.

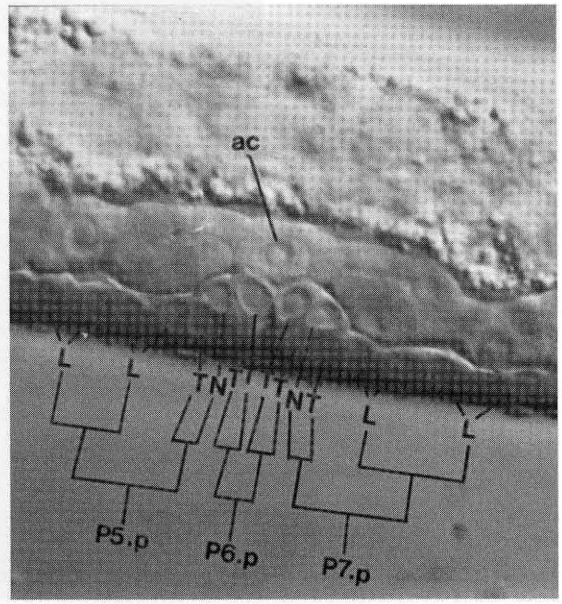
The key to our technique is that the animals are transparent, so that we can actually watch individual cells as they grow in the intact organism, and follow what becomes of them. (This approach was developed in 1976 by John Sulston at the MRC Laboratories of Molecular Biology in Cambridge, England.) We put the worm under a microscope that magnifies it about a thousand times, and as the worm goes about its business crawling all over the petri dish, we twiddle knobs under the microscope stage to move it around and keep the worm in our field of view. This skill takes some practice—it takes most students several weeks to acquire the knack—but it has the added advantage of making us tough opponents in the video arcade. We can also remove a particular cell by focusing a laser microbeam through the microscope's optics onto that cell, boiling it. Furthermore, roundworms only have a small number of cells. Excluding the germline—the eggs and the sperm—the hermaphrodites have 959 cells, and the males have 1031. (Hermaphrodites are females that make sperm as well as eggs.) The number isn't completely precise, because occasionally a worm is plus or minus one cell. So, after years of study, we now know all the cells in the organism as individuals. In many cases, we know what the cell is going to do before it does. We can tell by its position that a cell is going to make skin instead of a vulva, for example, yet we can show by doing the sort of microsurgical experiment I described above that the cell hasn't yet made the choice itself.

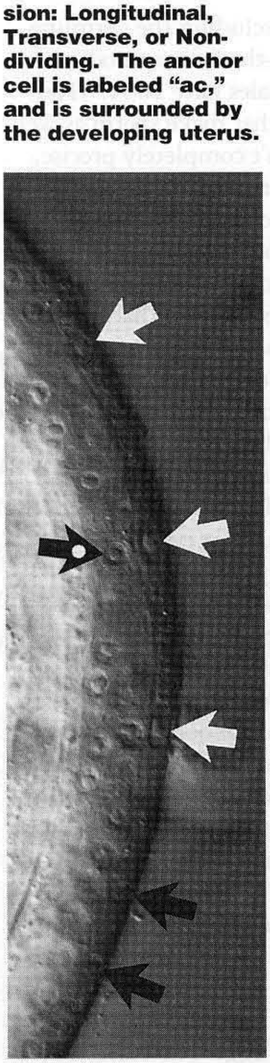
My lab has been studying the process by which the vulva is formed on the belly of the developing worm. The vulva is easy to study, because it develops rapidly—in just a few hours—and it involves only a handful of cells, making it easier to track their individual fates. And since the vulva is not vital to the worm's growth or reproduction, we can easily grow viable mutant strains that have inborn (hereditary) defects in vulval development. The vulva is the organ that gets the eggs out of the animal. Once eggs are produced in the ovary, they get fertilized in the gonad by the worm's own sperm, or by sperm from a male worm. (These eggs are quite small—1,000 would fit on the head of a pin.) The fertilized eggs start dividing. Once an egg has divided into a 20-cell embryo, it is forcibly ejected through the vulva and onto the petri dish to make room for another egg. The vulva is actually a specialized piece of skin, as Sulston discovered. In the embryo's developing gonad he found

Below: Although they don't yet know it themselves, the cells indicated by white arrows are fated to become vulval cells, while those marked with solid black arrows will become skin. The anchor cell (dotted black arrow) is the divinity that shapes their ends.

Right: A few hours later, precursor cells P5.p, P6.p, and P7.p have each given rise to a family of cells, as shown by the black lines. These cells, which look like sunny-side-up eggs, are now moving inward to form the vulva, visible as a dark, arrow-shaped indentation. The letters indicate the cell's mode of division: Longitudinal, Transverse, or Non-dividing. The anchor cell is labeled "ac," and is surrounded by the developing uterus.

Reprinted from Sternberg and Horvitz, *Cell*, Volume 58, August 25, 1989, pp. 679-693. Copyright 1989 Cell Press.





one particular cell, called the anchor cell, that signals three precursor cells in the skin to divide an extra time, start moving into the worm's body, go through a complex series of shape changes, divide again, and generate the cells of the vulva.

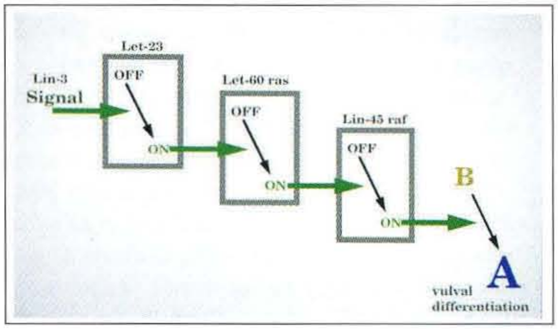
The problem my lab is working on is this: How do these cells know to become specialized and make a vulva instead of remaining nonspecialized and making just skin? In the smooth belly of the adolescent worm at left, the three white-arrowed cells will give rise to the vulva, and the two black-arrowed ones won't—they'll just become skin. But if given the chance, they would make a vulva. The signaling cell, shown here with a dotted black arrow, produces a signal that reaches the three nearby cells but not the more distant cells. If we destroyed those three cells, there would still be a vulva because the outer cells would move in and make one. So these cells really have two choices—they can make a vulva or skin.

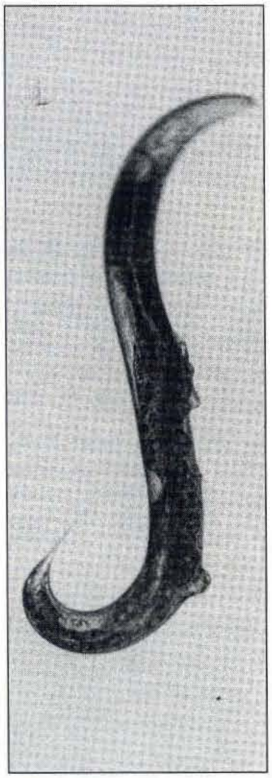
And the beauty of roundworms is, we can see it all happen. If we sit in front of that microscope for eight hours, we can actually watch these three cells divide, move into the worm's body, and connect up to the developing gonad and form the vulva. This technique allows us to do a variety of experiments with unparalleled precision, because every animal is the same, and we know all its cells. We get the same reproducible effect from the same perturbation, a level of precision that you rarely get with more complicated animals.

We've found lots of mutations that affect vulval development, and I'll give you examples of two classes. One class we call vulvaless. All the cells are present—the signaling cell, and the cells that normally respond to the signal—but no vulva is formed. There are two things that could be going wrong here: The cells could be failing to respond to the signal, or it could be that the signal is not being sent. The vulvaless class contains examples of both kinds of malfunction. In the other class of mutations, called multivulva, not only do the three normal cells make the vulva, but the other three more distant cells also try to make vulvas. These mutants are easily recognizable to the trained eye, because they have lumps on their bellies formed by cells in the wrong location that are trying to make vulval structures but can't quite do it. (They also have a normal vulva, so they can still lay eggs.) One of the really interesting properties of these multivulval mutants is that, even if we get rid of the signaling cell, all the cells still make vulvas. The cells act as if they are constantly getting the signal. I'll explain why shortly.

These mutations allow us to identify the genes involved in the signaling process, but our real goal is to understand the order in which they act. Over the last few years, genetic studies have told us that these genes make proteins that act like switches. That is, the proteins can exist in two states—active and inactive, or ON and OFF. On the opposite page is a simplified model of three of these switchlike proteins acting in series. The *lin-3* signal activates the *let-23* protein, which impinges on the next switch (*let-60 ras*) and turns it on, which in turn throws the third switch, *lin-45 raf*, and that switch then makes the cell turn

Below, left: A multivulva mutant worm. The three growths on the worm's right (i.e., belly) side are vulva wannabes. The normal vulva is also visible, midway between the lower two growths.
Right: This simplified signaling pathway consists of three switchlike proteins acting in series to decide a cell's fate.





from type B into type A, which differentiates into the vulva. (These genes' arcane names come from abbreviations describing what their proteins do—*let* stands for lethal, for example, and the number 23 indicates it was the 23rd gene discovered that, when eliminated, causes the worms to die.)

Mutations can affect this process in several ways. For example, if we make a mutation that eliminates the activity of the *lin-45 raf* gene, the third switch is now broken, locked in the OFF position. The *lin-3* signal comes on, and turns the first switch on, which turns the second switch on, which tries to turn the broken switch on, and nothing happens. The cell stays as type B. There's no vulva formed. This worm is one of several strains of vulvaless mutant worms we've made. Other mutations that cause a particular protein to be much too active—locked into the ON state—cause multivulval worms. If, say, the second switch (*let-60 ras*) is always on, it will turn the third switch on, and make the cell become an A, even if there's no *lin-3* signal. Because the switch is stuck in the ON position, it doesn't need anything beforehand to turn it on. In some cases, like the *let-60 ras* gene, we have one mutation that locks it ON and another that locks it OFF, so we can set the switch in whichever position we want.

So the key experiment is, if we have one mutation that locks one switch ON, and another mutation that locks another switch OFF, what happens if we put both mutations together in one animal through a simple genetic cross? Which mutation wins? There are two possibilities: Say the switch that's stuck in the OFF position acts after the switch that's stuck ON. The signal comes in and turns on switch number one. Switch number two is broken in the ON position anyhow and is already trying to turn switch three on, but can't because number three is stuck OFF, and the cell stays a B. Switch number three wins. Alternatively, if the broken OFF switch is earlier in the pathway, say at switch number one, when the signal comes, nothing happens at number one, but since number two is stuck ON, it will turn on number three regardless, so number two wins. Either way, the mutation farthest downstream prevails.

By doing many such experiments, we can come up with the order in which the genes act. (In fact, all Caltech biology majors are required to take a worm-genetics lab where they make such crosses and try to deduce a pathway.) There are considerably more genes involved than just these few, and tracing their interactions is much more complex than what I've just described—for example, some genes are inhibitors that send a signal downstream that tells another gene *not* to turn on; an inhibitor gene stuck on ON acts like an ordinary gene stuck on OFF, but that's the idea. I started working on this pathway about a decade ago, and we've probably only figured out one-fifth of it.

But what does this have to do with cancer? It turns out that all four of these genes have counterparts in humans. Raffi Aroian, Min Han, Andy Golden, Russell Hill, and Jane Mendel in my laboratory have demonstrated this in two ways. First, recall that every protein consists of a particular sequence of amino acids that

.	.
.	.
.	.
N	S
G	G
G	S
Q	Q
Q	Q
Q	V
Q	E
Q	Q
P	P
T	T
G	G
S	S
I	V
L	L
W	W
M	M
A	A
P	P
E	E
V	V
I	I
R	R
M	M
Q	Q
D	D
D	N
N	N
P	P
Y	F
T	S
P	F
Q	Q
S	S
D	D
V	V
Y	Y
S	S
F	Y
G	G
.	.
.	.
.	.

Left: A portion of the amino-acid sequence for the *lin-45* roundworm protein (left) and the human *raf* protein (right). Right: Postdoc Wendy Katz maneuvers a roundworm cell into the line of fire of the laser—the thin vertical black box sitting atop the microscope.

determines the protein's chemical structure. There are 20 different amino acids, each of which we represent by a one-letter code. Thus A stands for alanine, G is glycine, R is arginine, S is serine, T is threonine, and so on. One part of the roundworm gene, *lin-45*, tells the cell to assemble this sequence of amino acids: TGSILWMAPEVIRM. When we look at the sequence of the human *raf* gene, we find almost the same sequence—TGSVLWMAPEVIRM. In fact, all the letters shown in red at left are the same, and this degree of similarity extends over several hundred amino acids. So when we see two proteins this similar in sequence, we have a pretty good bet that the proteins have a similar structure. We can then hypothesize that they have the same function in the cell, which generally turns out to be true.

This is a very important hypothesis, and it is supported by our second proof: We take a gene from one organism and put it in the other and see if it still functions in the same way. This technology, called transgenic technology, is now available for a variety of simple experimental organisms, including worms, fruit flies, mice, and yeast. We inject into the ovary of the worm the human (or whatever) DNA that has the gene we want to compare. Some of that DNA gets incorporated—we don't know the details of how it happens, we're just lucky that it does—into some of the worm eggs. These worms, called transgenic worms, now have acquired that injected gene, and they'll pass it on to their progeny. We don't want to have to take a DNA sample from each worm to find out if it has the gene, so we put a marker on the gene we inject. One nice marker is a mutant worm gene called roller, which causes the animal to have a twisted body. When roller worms try to move forward, they twist to the right and end up rolling. It's very easy to spot them—all their siblings are crawling around very nicely, and the rollers just go in circles.

I'll give you a particularly spectacular example of such an experiment, using a rodent gene that's about 80 percent identical to a roundworm gene that's part of a different pathway we're also studying. If you took five normal, or "wild-type," worms and put them on a microscope slide and waited for five minutes, they'd all be gone—they'd have crawled out of your field of view. But if you put five mutant worms there and waited 40 minutes—eight times as long—all the worms would still be in your field of view. Some of them might not have moved at all. The worms have a neurological defect that makes them have serious trouble moving forward or backward. Junho Lee and Gregg Jongeward in my lab have identified this gene and the protein it makes. It turns out that they're very similar to a gene and protein that other people have discovered in rodents and humans, where it is involved in moving membrane proteins around inside the cell. So we asked ourselves, if we take the rodent gene and inject it into our mutant worm, will the mouse gene function? If it does, the worm should move normally. There are subtleties, of course—we have to make the mouse gene into a hybrid so that it stands a chance of functioning in the worm, yet have most of the protein encoded by the mouse's DNA. We've done this experiment, and the "rescued" worms crawl just fine. So this mouse gene will function in the roundworm, and we can confidently say, to a first approximation, that the two genes are the same.

We know from similar examples that each of the genes involved in vulva differentiation in the roundworm has a counterpart—or several counterparts—in humans. Thus the protein encoded by the *lin-3* gene looks like human EGF, or epidermal growth factor, protein. And just as *lin-3* is a signal between roundworm cells, EGF and related proteins act as signals between human cells. Then, on the responding cell, there's a protein that acts as the receptor—in the worm it's *let-23*, which resembles the EGF receptor protein in humans. This receptor binds to the signaling protein and controls what that cell does in response. Inside the cell, the signal is somehow transduced, or changed in form, by other proteins—switches like *let-60 ras* and *lin-45 raf* in the worm, and their human twins, genes called simply *ras* and *raf*. The transduced signal travels down pathways that many research groups are just beginning to explore, and eventually reaches the cell's nucleus. There the signal controls what genes are turned on to make the cell proliferate, or change shape, or otherwise choose its fate.

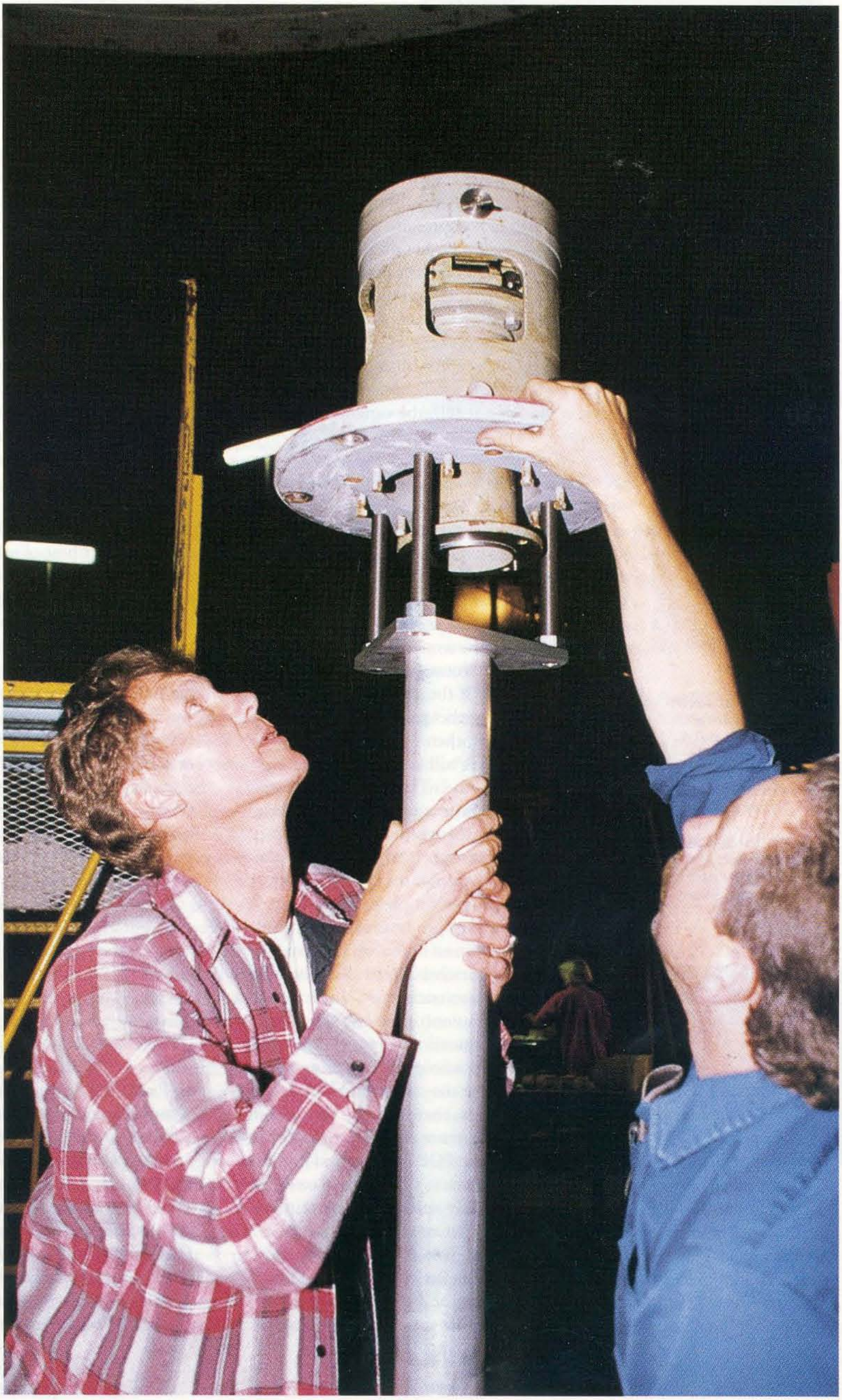
Above: Grad student Junho Lee pulls a worm out of a petri dish. The worm is impaled on the tip of the stainless steel probe in his right hand.
Below: A three-day's supply of fresh, nutrient-laden petri dishes for the Sternberg lab.

Since we can draw a one-to-one correspondence between the worm genes and the human genes, we can say, "If the genes work in this particular order during this particular process in worm development, then we predict that in humans, these genes will act in the same series to control cell growth." The genes' actions may have different effects because they are triggering other switches that the worm doesn't have, but we expect the order of their triggering to be the same. And this prediction turns out to be correct. So we can use the simple genetics of one organism—the worm—to learn about some really important genes in an organism that we care a lot more about—ourselves. And all of the human equivalents are known oncogenes. In fact, *ras* is a particularly infamous oncogene—it's the one most frequently mutated in colo-rectal cancer.

But this isn't the whole story. If it were, we could probably solve the cancer problem in a few years. Unfortunately, there are a lot of genes still to go. For example, there are at least two other proteins between the EGF receptor and *ras*. Just in the last few months, it's been discovered that *ras* interacts physically with the *raf* protein. Then, after *raf*, but before cell growth, there are a lot more genes. We still need to figure out what they are, and the order in which they act, and then we need to know the details of what controls them and how they function. That's the level of understanding we're going to need in order to look at a tumor and say what went wrong. And *that* knowledge will enable people who are good at that sort of thing to design ways of intervening—that is, to come up with therapeutics or new drugs.

There are two ways to eradicate cancer: One is to prevent it from happening in the first place. We can all stop smoking; we can get rid of a lot of environmental carcinogens. We know that most agents that lead to cancer are either mutagens that mutate the DNA or tumor promoters that stimulate cell proliferation. And the more cells divide, the more likely they are to mutate and cause cancer. That's something we can take care of without any fancy science—we just have to use common sense. The other way, to eradicate cancers that have already started, is to come up with the next generation of very specific anticancer drugs. The drugs we have now essentially kill any and all dividing cells. This has nasty side effects, because the cells that line the stomach, and the cells that make hair (not to mention the ones that do a host of other things) also have to divide. You wind up killing them, too, which is why chemotherapy patients suffer nausea and hair loss. But as researchers discover which protein binds to which receptor to send a signal, they can try to come up with drugs that interfere only with those specific interactions. No one's done it yet, but it's promising—last year, a number of biotech start-up companies formed to take advantage of the knowledge we've gained about the signaling pathways in these oncogenes. The point is, the basic understanding of the mechanism will lead to large-scale efforts to come up with drugs based on those mechanisms. □

Paul Sternberg chose biology as a major because "I couldn't get an appointment with the economics advisor." Sternberg earned his BA in biology from Hampshire College in 1978, and his PhD from MIT in 1984. He came to Caltech as an assistant professor in 1987, and was promoted to associate professor in 1992. Sternberg holds a joint appointment with the Howard Hughes Medical Institute in Pasadena, where he was appointed assistant investigator in 1989, becoming associate investigator in 1992. This article is adapted from the Seminar Day talk he gave in May.



The Bad News Bearings

Casting and polishing a replacement mirror today would cost an estimated $15 million, so the Palomar staff were understandably reluctant to risk treating the original like a twist-off bottle cap.

by Douglas L. Smith

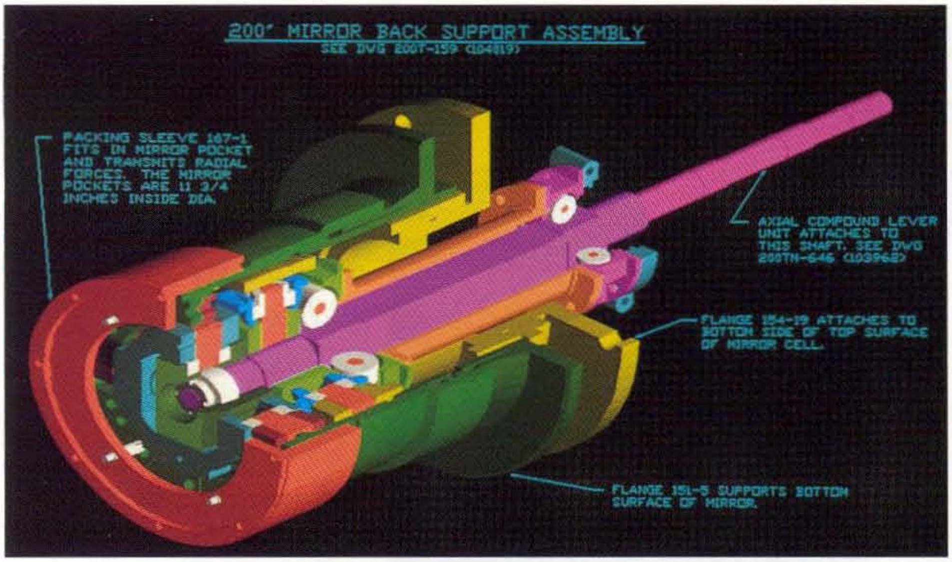
Removing the central gimbal assembly is the tensest part of the process. At 166 pounds, this is the heaviest piece to come out of the mirror and far too heavy to lift by hand. Even though Hal Petrie, Palomar's chief engineer, designed a special hoist for the job, getting the gimbal to ground reminds one of the plight of the fellow holding a bucket of water against the ceiling with a broomstick. Petrie (left) and preventive maintenance mechanic Bruce Baker (right) stabilize the irreplaceable chunk of 1930s engineering bal-anced one false move away from a 14-foot plunge to a concrete floor. Not a sigh of relief, perhaps, but once the assembly is safely down on the forklift, there's certainly a collective letting go of the breath.

Palomar Observatory's 200-inch Hale Telescope—the world's biggest for nearly 30 years, and still one of the most productive—began to show its age recently. Even after 45 years of nonstop use, the telescope remains a premiere scientific instrument, thanks to advanced instruments and aggressive maintenance. But now parts of the sensitive system of supports that maintain the mirror's shape were beginning to stick. And the parts were unreachable—way up inside the mirror itself. In 1947, while the mirror stood balanced on its edge in Caltech's campus optical shop, supported by a giant cradle, the supports had been inserted into pockets cast in the mirror's honeycombed underside. The mirror was then gently tipped into its mounting—a two-foot thick labyrinth of steel members called the mirror cell—on which it lies flat and from which it has never since been removed. Casting and polishing a replacement mirror today would cost an estimated $15 million, so the Palomar staff were understandably reluctant to risk treating the original like a twist-off bottle cap. But everyone, from Observatory Director Gerry Neugebauer on down, agreed that something had to be done.

Says Hal Petrie (BS '68), Palomar Observatory's chief engineer, "For some time, the infrared observers, who use very high magnifications, had been complaining about astigmatic images—images that weren't round, and in some cases were very strongly out of round. What you should see is a round point of light that gets bigger and blurrier as you go out of focus. But as they went out of focus in one direction, they'd see an oval oriented one way; and on the other side of focus, they'd see an oval oriented 90 degrees to the first. And at focus, they basically got the intersection of the two ovals. It looks sort of round, but it's not as sharp as it should be.

"Even in its worst state, it's not a bad shape. Generally the images are better than the seeing. But on nights of good seeing, the mirror's figure can deteriorate the image. And we do get some nights of very good seeing here." The "seeing" is the distortion imparted to the image by turbulence in the atmosphere. With the telescope at zenith, and the support systems working properly, 80 percent of the visible light from a distant source focuses in 0.45 arc seconds, and about 50 percent in $\frac{1}{3}$ arc second. But off zenith, the numbers get about 50 percent worse. Depending on the time of year, the average atmospheric distortion at Palomar is about one arc second, but on good nights it's only half that. In the infrared at two to three microns, it's even better.

Trouble is, like all large telescope mirrors, this one can't support its own weight. It sags out of its optimum light-gathering shape the moment it no longer points straight up. Even though only the top $4\frac{1}{2}$ inches of the 22-inch thick mirror are solid glass—the rest is ribbed into a honeycomb pattern to save weight—it weighs nearly $14\frac{1}{2}$ tons. A 200-inch mirror thick enough to be self-supporting would be impossibly heavy, as is the alternative—a perfectly rigid telescope that could hold a flexible mirror in shape. (The telescope tube weighs 138 tons as it is.) Thus, in a compromise between stiffness and weight, the mirror cell flexes by about a millimeter as the telescope moves from zenith to horizon.



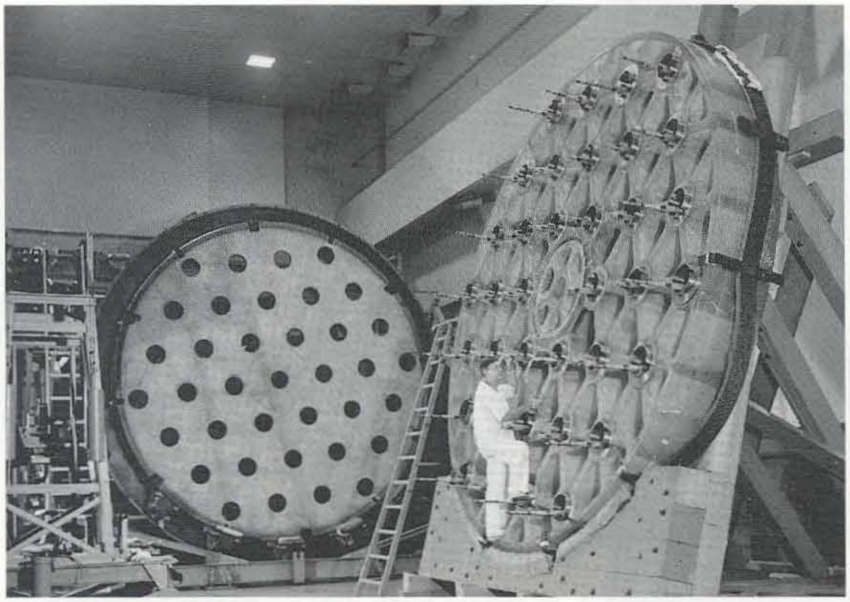
A CAD drawing of the back support, rotated 90° from its vertical position in the mirror. The dark green sleeve fits snugly in the mirror pocket, while the yellow-green flange bolts to the mirror cell. The axial and radial forces needed to keep these two fixed points in their correct relative positions are transmitted through a common shaft (purple). The axial lever unit (not shown) bolts to the lavender casting's bottom and imparts a force to the shaft's exposed end. The force acts via the concentric gray-blue and maroon rings of the upper gimbal assembly on the lime-green sleeve, which is wedded to the dark green sleeve, and hence on the mirror. The radial force is mainly generated by the weight of the axial lever-arm assembly. The white bearings in the salmon pins in the dark blue ring act as a fulcrum for this force, transmitting it through the upper gimbal assembly to the red sleeve, which imparts it to the mirror.

And this is where the sticky supports come in. Between the cell and the mirror are 36 "back supports" that maintain the mirror's figure, or precise optical shape, despite the flexure of the mirror cell. Each support is an elaborate system of weights and levers that exactly balance gravity's force on the mirror no matter where it points. And since both the magnitude and the direction of the compensating force change as the telescope moves all over the sky, the lever systems are masterpieces of subtle mechanical engineering—a pinnacle of 1940s high technology. (See the cutaway drawing above.) Each support counterbalances the force's axial component (which acts along the telescope's length) and radial component (which acts perpendicular to it) with separate sets of weights acting on separate sets of levers with different lever ratios, but then applies both forces to the mirror through the action of a single shaft on two sets of gimbals. (Other telescopes of a similar vintage use two separate sets of simpler lever systems for radial and axial support. And the simpler systems provide radial support only at the mirror's edge, but the Hale's back supports distribute the radial load uniformly over the entire 200-inch mirror, giving it a better figure.) The supports near the center of the mirror balance about 700 pounds' worth of mirror each. Those at the periphery, where the mirror is thicker and the supports are more widely spaced, carry up to 1,100 pounds. Since the lever systems are all floating—"they give you a feeling like a waterbed," says Petrie—three supports at 120° intervals around the mirror's periphery are locked down to "define" the mirror, keeping it parallel to the mirror cell.

When the mirror started to lose its figure—as so many of us do at age 45—the back supports were immediately suspect. After all, there were bearings up in there that hadn't been lubricated in 45 years! But were just a few of them sticking, or all of them? Maps of the mirror's shape, made by a sophisticated wavefront analyzer built especially for the Hale Telescope by Gary Chanan, professor of physics at UC Irvine, were inconclusive. They did show that there was a lot of hysteresis in the mirror—in other words, the changes weren't reproducible. If you mapped the mirror's precise shape when pointed at the zenith, then tilted the telescope as far south as it would go, and then as far north as it would go before bringing it back to the zenith, the figure would change constantly as the telescope moved. But when the telescope was finally brought back upright, the mirror didn't always return to its initial figure.

To ferret out the root of the problem, says Petrie, "during a long series of nights when we were taking lots of engineering data last fall, Keith Matthews [BS '62], a member of the professional staff in physics who has designed many of the infrared instruments for the telescope, actually rode in the Cassegrain cage of the telescope as it was pointed at different places in the sky. [The Cass cage, as it's called for short, bolts onto the back side of the mirror cell, and provides access to instruments mounted at the telescope's Cassegrain focus point.] This was not an easy thing to do, because although the Cass cage has a floor to stand on, the minute you tip the telescope off zenith, the floor tips too. You start sliding around, and the next thing you know, you're standing on instruments or things. But we put a safety belt on him, and gave him a long broomstick. Then we took the telescope from the zenith, where the image was quite good, to a place in the south where we could frequently get a bad image. And when Keith just touched one of the back supports with the broomstick, the back support made a clunking sound and the astigmatism immediately went down by a factor of two. By moving the telescope around and touching every single one of the back supports we determined several things. One was that the mirror's shape could be improved by jiggling the back supports and getting the 'stiction' out of them, but that it never got as good as it was at the zenith. Also, about one-third of the supports went 'clunk' when you touched them, indicating that they had severe stick-slip problems."

How to get to those buried bearings was a real poser. Popping the mirror off the cell was not a



When the mirror stood on its side in the optical shop in the 1940s, getting the back supports out was easy. It's a bit trickier nowadays. From top: Petrie, facilities maintenance mechanic Russ Day, and Baker.

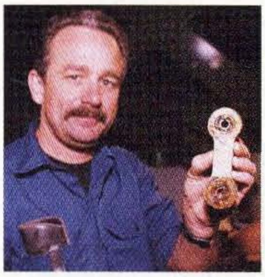
popular option, but nobody knew how much of a back-support assembly was accessible through the holes in the mirror cell below. Petrie explains, "These giants built the telescope, and as they retired and left the scene, not all their knowledge got transferred. Real folklore grew up around this. I go out with my own amateur telescope once a month, and other amateurs come up to me and say, 'Is it true that there are so many thousand parts in there, and nobody alive understands how they work?' In fact, we did know how they worked, although the details of their installation process had been lost. And there *are* a lot of parts, but they're conceptually fairly simple once you get into it.

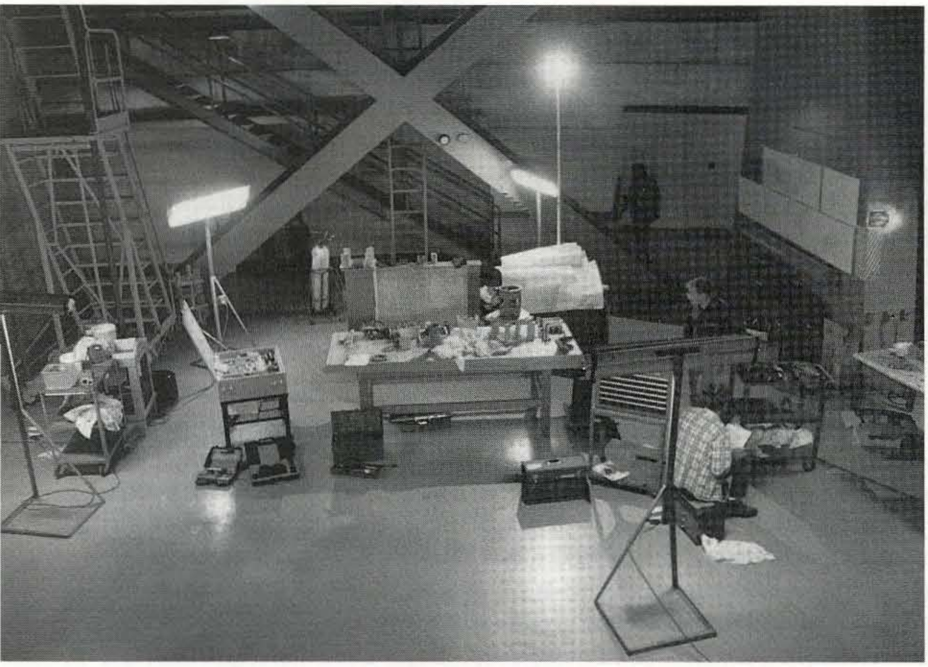
"So Keith and I looked at the assembly drawings for the back supports, and attempted to understand how the mechanism was put together. We deduced correctly that you can't get the packing-sleeve assembly out, because it's bigger than the hole in the cell. But it wasn't clear that we could get all of the bearings out safely. So I started taking the old machine-shop drawings and converting them into 3-D CAD [computer-aided design] models using a program called AutoCAD." It took Petrie a month to get the drawings into the computer. AutoCAD builds objects by adding and subtracting appropriately sized and oriented "primitives"—simple geometric solids such as spheres, cylinders, and cones—from each other. "These parts are mostly castings, and with all their complex, whittled-out shapes, it often takes some imagination to figure out which primitives should be added and subtracted, and in what order." Petrie initially intended to use the model to see if he could get a fiber-optic camera or a tube full of lubricant up into the works. "Once the model was made, we realized that, in fact, you could get all the stuff apart without removing the mirror from the cell as had been done in the optical shop. We were also concerned that if we got it all out, could we get it back in? But the clearances are fine."

After several weeks of testing their procedures on the computer to assure themselves that they weren't about to do anything irretrievable, they very cautiously removed and inspected one of the worst back supports, unit K_5. Recalls Petrie, "When we took the first one apart on Memorial Day weekend, it was pretty clear that we had a lubricant failure rather than a bearing failure. We didn't have corrosion of the bearings, we didn't have pitting, or cracking, or anything like that." Nor were the individual components deteriorating. All the back-support parts that go inside the mirror are made of Invar, a high-nickel stainless steel whose thermal expansion exactly matches that of the Pyrex mirror. As a bonus, Invar doesn't corrode. (Corrosion could have caused parts to stick together, and perhaps break during disassembly.) The grease, however, had oxidized and polymerized into a tough, rubbery solid that had frozen some bearings outright. Others ran very rough as their balls ground up the dried grease. "We cleaned at least six bearings by hand, with toothbrushes, solvents, and hot, soapy water. A lot of that gunk was very hard to get out, but once we did, the bearings were fine."

The bearings had been out of sight all this time, but they hadn't been out of mind. Bruce Rule (BS '32)—Palomar's chief engineer when the mirror was installed, and one of Petrie's "giants"—and colleagues discovered early on that the telescope's performance improved if it was "exercised" periodically by driving it all over the sky for several minutes. This flexed the telescope, working the bearings and freeing them up. If this wasn't sufficient, the levers in one mirror-defining back support would be unlocked and cycled back and forth through their full range of travel a couple of dozen times, rocking the mirror about the axis created by the other two supports. This pushes all the other levers through a much wider arc than they normally move and tends to free up sticky bearings. Says Petrie, "that was done periodically, according to Bruce Rule, to 'break the crust that was forming on the grease.' I don't know if they really knew what was happening, because most of the bearings they were servicing by this process were up inside the mirror, inaccessible. We don't know how often it

Top: As seen from the dome-level catwalk, the telescope monopolizes the eye, dwarfing the huddle of people and paraphernalia on the hydraulic ram below it. The ram, built into the dome floor to provide access to the mirror cell, becomes the scaffold for most of the removal and reinstallation work, although a mobile scissors lift was needed to reach the outer supports. In order to get to the bearings, the telescope had to be locked at zenith, and the Cass cage unbolted. Once lowered on the ram and shoved out of the way into the lower-left corner of the photo, the cage and its multitudinous electrical outlets became a handy power strip for the work lights on the ram. Middle: Baker shows off the culprit grease. Bottom: The work area, which is out of sight behind the gargantuan steel pier that dominates the right side of the top photo.



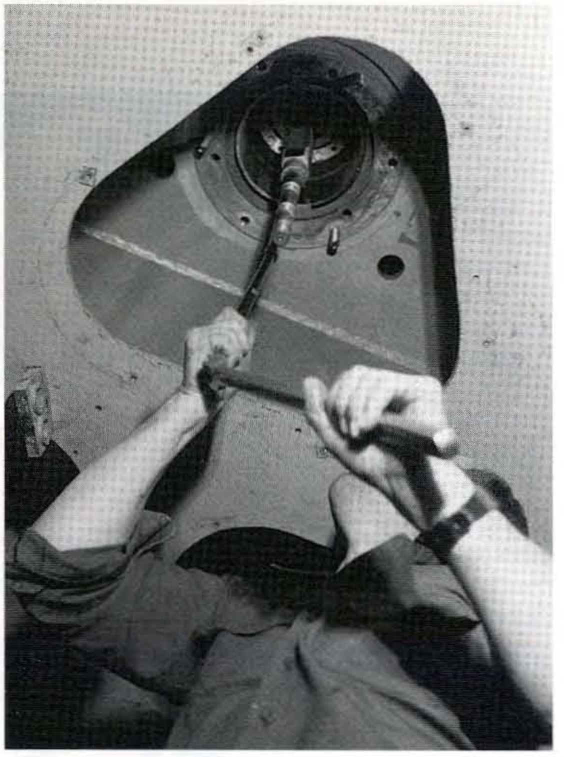


was done, either." It is indisputable, however, that the force required to move a lever after exercising it was smaller.

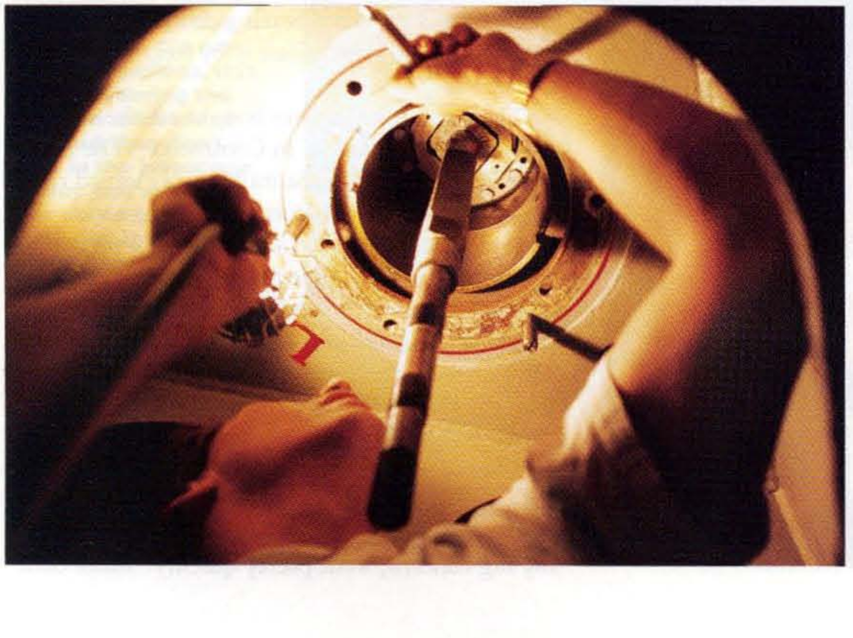
Not all of the grease is 45 years old. The axial assembly, a double-compound-lever system of near-baroque complexity that Rule designed, hangs down into the mirror cell and has accessible grease fittings. "We don't have a good history of how often those were greased, or what they were greased with," Petrie notes. "There are some vague comments in some of the reports about a class of greases, but no brand name or chemistry is given." In 1987, Palomar Superintendent Bob Thicksten flushed a hot silicone-based oil through the grease fittings to clean out whatever could be cleaned. "They saw at least three different kinds of grease come out," Petrie continued. "They saw 'Gargoyle grease,' which is an old, red grease from the Gargoyle Oil Company; and they saw black grease, which is the molybdenum disulfide grease commonly used around here now; and then there was some other, clearer, grease of unknown composition." Once the bearings were flushed until clear fluid came out, silicone grease was injected.

But now they're discovering that the hot oil didn't get all the way through every bearing. "We're still finding mixtures of different kinds of greases," Petrie says. "In some cases, the grease fittings were plugged and no grease got in. So although it helped a lot, and those bearings are in better shape than most of the ones in the mechanism, it wasn't a complete fix. And it only affected the 13 bearings per assembly that are accessible from under the telescope. The other 26 were up inside, getting no attention."

Once computer technology had shown the bearings to be reachable, decisions had to be made: What grease to use? Would a redesigned system be better? The Palomar staff looked into a number of different greases, and considered such exotica as ceramic bearings that never need lubrication. They opted to stick with the existing design, since its mechanical aspects were sound, and put in new bearings lubed with high-performance grease. This grease had to have long-term stability, offer corrosion protection to the bearings, and not absorb water. And finally, the lubricant is periodically exposed to a 10^{-5} torr vacuum every two or three years for five or six hours, when the mirror gets realuminized. Thus the lubricant has to have a low vapor pressure, or it will boil away into the vacuum and contaminate the mirror's reflective aluminum coating. While not quite the void of interplanetary space, this vacuum is equivalent to the rarefied air of Earth's ionosphere some 80 miles up. So Petrie,



Clockwise, from upper left: The axial support unit comes out first, once Baker rotates it in place and removes parts that won't otherwise clear the cell. Baker and engineer Bob Weber lower the 50-pound axial unit out. Once the central gimbal assembly is gone (p. 34), it takes a socket wrench with a four-foot extension to reach the bolts securing the upper gimbal assembly. The orientation of the upper gimbal is marked before it's removed from the cell. The gimbal comes out, its 33-inch shaft dangling within the hollow interior of Petrie's hoist.

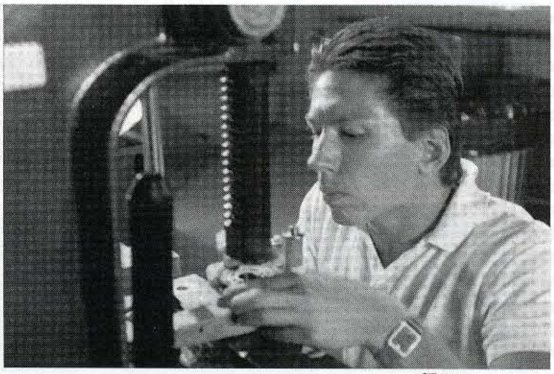


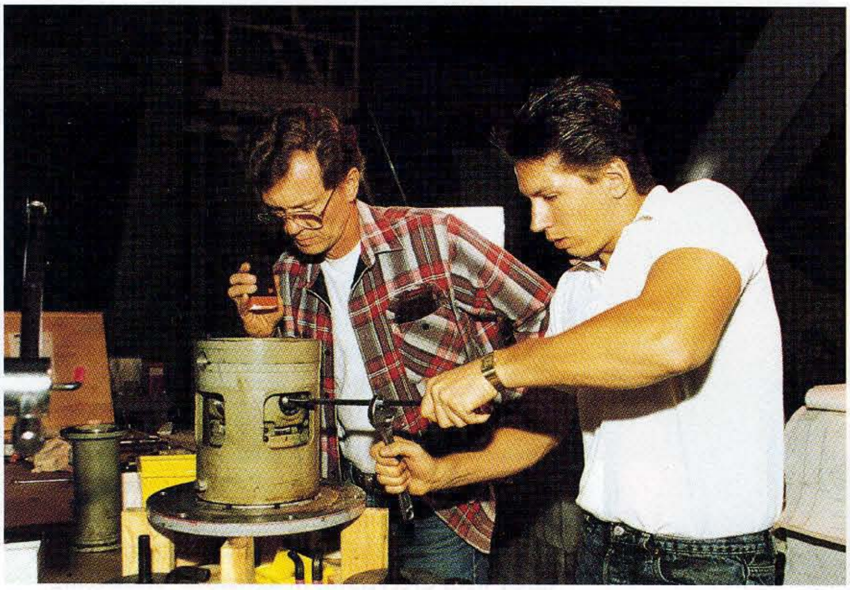
armed with a list of NASA contacts provided by astronomy and space-science photographer Roger Ressmeyer, tracked down the spacecraft-bearing experts at JPL and other NASA centers. Several experts recommended their standard flight-certified spacecraft grease, called BrayCote 601. This stuff doesn't oxidize and its vapor pressure is practically nil—10^{-13} torr. It's also a \$100-per-ounce perfluoropolyether that comes in two-ounce syringes.

The first opportunity to fix a lot of bearings came the week of July 26-31. The plan was to service the 10 worst-performing supports—the ones that went "clunk" when Matthews nudged them. The operation would go like an assembly line—supports would be pulled out, dismantled, bearings replaced, reassembled, and reinstalled in a smooth flow.

Each support comes out in several pieces, which are carried to a temporary work area set up beneath the massive horseshoe girder on the telescope's north pier. There, surrounded by a ring of work lights on stands, the units are placed on long tables covered with brown butcher paper. The area could be an operating theater, with small carts instead of gurneys, and tool chests on wheels instead of heart monitors and anesthesiology equipment. The overall effect is as if the surgical team from M*A*S*H had set up a triage station in the *Enterprise's* shuttle bay.

In the operating suite, the disassembly proceeds amidst a clutter of screwdrivers and wrenches, paintbrushes and cleaning rags, and the ubiquitous $2^1/_2$ pound Folger's coffee cans filled with everything imaginable. Dissecting





Top: Facilities maintenance mechanic Dana Cuney gets the old grease off with paint thinner. Eventually, rags and paintbrushes give way to a wire toothbrush to clean out all the holes and cutouts. Some of the parts have *a lot* of holes and cutouts. Middle: Weber uses the arbor press to seat a bearing. Bottom: Petrie and Weber dismantle a central gimbal assembly, a process not unlike opening a Chinese puzzle box.

some components is easy, as these things go, but disemboweling the upper gimbal assembly isn't. Most of the pieces that hold it together are hidden within a steel sleeve. "You have to work through the windows," says Petrie. There's a lot of twisting, jimmying, and finagling between removal of successive parts, but eventually, everything yields. Nuts, bolts, washers, and other small stuff are consigned to yellow plastic trays labeled with the assembly's name. The larger pieces lie neatly arranged on the butcher paper, awaiting cleaning with paint thinner to get the old grease off.

As the old bearings are removed, they're given a cursory examination and tossed into a corrugated cardboard box by the north pier. The bearings themselves are unremarkable. The larger ones wouldn't look out of place in the wheels of a riding lawnmower.

Pulling the old bearings out goes quickly. Not so getting the new ones in. Several people can be dissecting different parts at once, but there's only one arbor press to seat the snug-fitting bearings. As engineer Bob Weber seats a bearing in a connecting rod in the axial assembly, he explains, "If you don't center the bearing in its hole, one side or the other would rub. And, unfortunately, they didn't put a shoulder on the side of the hole to seat the bearing against." When in doubt, improvise. Someone digs out a dime, which Weber puts under the bearing as a spacer. Petrie remarks, "Money plays a very important role in all this. The first one we did, we had to sacrifice a few pennies to get a bearing out. Now we're making a shim out of a dime." Weber produces a micrometer and announces mournfully, "Dime's too thick." (They need a 0.050 inch spacing between the outer bearing race and the surface, and the dime is 0.052 inch thick.) "So I press the bearing down to the dime, measure it, and then adjust it by feel. The dime gets it really close."

The new bearings were bought off-the-shelf from King Bearing Co. in Commerce. The Aircraft Bearing Corp. in Santa Monica cleaned the manufacturer's grease from the bearings and packed them with BrayCote 601. "It's very time-consuming to clean bearings by hand," Petrie explains. "But they specialize in doing this—they have a clean room, and machines that blow compressed air and solvent through them. Then they use syringes to inject exactly the right amount of grease."

The original plan was to take out 10 supports, but, says Petrie, "we ended up changing out six units. Removal and installation, which had been my big concern, went pretty quickly. More time

Top: "I got into electronics so I wouldn't have to do this kind of stuff," grouses Superintendent Thicksten, only partly in jest.

Middle: Before the bearings reach journey's end at the telescope, they sojourn in Palomar's machine shop, where they take a spin on the lathe. Impaled 20 at a time on the lathe's spindle, the bearings spin against a rubber squeegee pressed gently against their outer races. "Running in" the bearings this way burnishes the balls against the race, substantially reducing their rolling friction, says Petrie. "In normal applications, this happens during the first few minutes of use, but they don't spin in the telescope—they just rotate a few degrees back and forth—so it would never happen."

Bottom: Sometimes parts need a little encouragement to go back together. Electrician Paul Van Ligten lines 'em up while Weber does the honors.

was spent in disassembly, cleaning, reassembly and adjustment than I had anticipated. The last one went in about dinner time on Friday. Saturday morning before the rotating ring and Cassegrain cage were reinstalled, I exercised the mirror for three minutes at each of two defining points. That night, we did a knife-edge test at prime focus. Neither Thicksten nor I could see any mirror problems. A couple of nights later, I talked to Keith Matthews, who was observing at the 200-inch. He indicated that the mirror was pretty good the first night, but there is still some astigmatism. He did not see any of the severe astigmatism he has sometimes seen in the past.

"It will probably be next year before we get all these bearings out. The observing schedules are set each fall for the next year, and certain blocks of time for engineering are taken out. So we're operating right now on the schedule that was set last October. And we can't just arbitrarily bump people to continue doing this work—they've been counting on their observing time for a year. This run right now involved negotiating with a bunch of observers, shifting them to the Fourth of July weekend, and us here. We've attempted work like this in the past in the dead of winter, when observers are more willing to give up time, but discovered that it's really hard on us." As anyone who has ever tried to change a tire in midwinter in the midwest knows, your skin sticks to everything, but gloves really hinder you.

"We can now pull an assembly out, do a quick inspection of it, and put it back in, in one day," Petrie says. (Disassembling one, putting in new bearings, and centering and readjusting everything takes a bit longer.) "We don't intend to let these things go another 50 years. Especially now that we know how to do it.

"Personally, I see this as the beginning of a project in which we get all of the mechanisms working the way they were supposed to, and then, using the wavefront analyzer, we might be able to get some really good response functions by hanging weights on each individual back support and seeing how it changes the mirror. Then if we have a disfigured mirror, we could identify where we should apply forces to fix it." The ultimate end might be to add small, computer-driven actuators to the lever systems, turning the mirror into an "active mirror" that constantly adjusts itself for optimum focus, just as the large telescopes being built today do. But that would take time, engineering, and money, because such a scheme requires a minimum of 108 actuators—three for each back support. For the moment, the folks at Palomar are happy to bring the venerable Hale into the 21st century with perfect vision. □

Letters

What is history of science, and who should speak for the past? In his review of my book, *The Molecular Vision of Life: Caltech, the Rockefeller Foundation, and the Rise of the New Biology,* in the spring issue of *Engineering & Science,* Robert Sinsheimer reaches a curious verdict. He writes, "Kay clearly belongs to the school of historical determinism that maintains the view that the course of scientific progress cannot be autonomous, but is always a response to cultural, usually political and economic, forces." I take this as praise! Historical determinism—the thesis that certain forces shape historical processes—is a fundamental premise of historical scholarship, and demonstrating that the development of science is a genuine historical process is one of the principal challenges to historians. For example, how do intellectual and technocratic elites shape, and how are they shaped by, social and political agendas?

That such a scholarly goal and its attainment constitute a first-order accomplishment in the history profession, while deemed subversive by many scientists, underscores the essential tension between the two professions. To be sure, this tension over who speaks for the past can be healthy and productive, providing it is governed by mutual scholarly respect and authentic interpretations of the arguments.

Thus I cannot fault Sinsheimer (or others from Caltech) for being scandalized given his reading of my interpretive framework. For he drew precisely the conclusions that I warn against in the lengthy introduction to my book: 1) that the Rockefeller Foundation had a hidden agenda of social control; 2) that individual scientists were manipulated and co-opted by the Rockefeller Foundation; and 3) that "human betterment" amounted to conspiracy or a Machiavellian plot. Indeed, as he rightly concludes, such lessons border on the ludicrous.

I do, however, fault Sinsheimer for misreading my thesis. As I make clear:

1) There was nothing covert about the Rockefeller Foundation's interests in social control; it was not a "hidden design." Quite the contrary, the trustees and officers explicitly and openly stated these goals in many of their documents, which I quote verbatim. I do not accuse or condemn but explain how their premises and specific formulations of social control were congruent with their commitment to the political and economic framework of pre-World War II America. (Articulations of social control in 18th-century France or 19th-century China looked quite different.)

2) Individual scientists were not manipulated and science was not co-opted by the Rockefeller Foundation or by Caltech trustees. Throughout the book I show how Millikan, Noyes, Morgan, Pauling, Delbrück, and Beadle "used" the Rockefeller Foundation as much as the Foundation "used" them. These were strong-willed, farsighted individuals who, as Rockefeller advisers, often told the officers how to plan. They were neither helpless pawns nor co-conspirators, as Sinsheimer portrays my account. I clearly say in my book (pages 8–11) that, being cut of the same cultural cloth, the managers of science, Rockefeller Foundation officers, and Caltech's trustees shared a *Weltanschauung*, yet I stress that this did not constitute an explicit agreement on all aspects of programs and policies. The complex problem in political theory of how intellectual elites fit into social agendas has been extensively studied, and it is on this body of knowledge that I base my analysis.

3) It does not take top-down coercion or a Machiavellian plot to get groups of people to cooperate. Any scientist with leadership experience must know that successful power sharing is predicated on compromises—some explicit, some tacit, sometimes unconscious. Scientists have always worked within bounded and negotiated autonomies. Today's constraints are different from those of the 1930s, but there have always been constraints on the course of science. Thus, rather than co-optation, I see the rise of molecular biology as a nuanced co-production of scientific knowledge by patrons and researchers.

That the Rockefeller Foundation had a shaping power in molecular biology is hardly news; there has been excellent scholarship on this topic. There are also outstanding works showing how institutions (including Caltech) and social trends have shaped (though not *determined*) the course of modern science. My work, which links social, institutional, and cognitive agendas, is *not* revisionist; it is squarely within the mainstream of history of science.

Sinsheimer laments that such detailed scholarship should have been placed in the service of a distorting, revisionist ideology. This is strange. Had I written a hagiography of molecular biology at Caltech under the aegis of the Rockefeller Foundation, would scientists view my account as ideology-free? Curiously, a history is pronounced ideological when it challenges the dominant version of the past. I did not come to this subject with an ideological bias. It was the archival documents, primary sources, and earlier historical works that shaped my interpretation.

I genuinely regret that, by the misreading of my thesis, the book has caused Sinsheimer and others dismay and that they feel affronted. It is important to keep in mind that this book is not primarily about individuals but about mechanisms, about how science as a system worked in a specific historical context. Individuals are certainly crucial elements in such a process, but surely the scientific whole is greater than the sum of its individual parts. (The social responsibility of scientists is another topic deserving separate discussion.) I have high regard for the science and scientists at Caltech in the period I have studied. My thesis is not aimed at individuals but at the social processes which, knowingly or not, they helped shape and were shaped by. Does Sinsheimer suggest that science at Caltech has escaped the forces of history?

Lily E. Kay
Associate Professor of History of Science, Program in Science, Technology, and Society, Massachusetts Institute of Technology

Random Walk

Honors and Awards

President Thomas Everhart has been named a Fellow Member of the American Society for Engineering Education (ASEE), a distinction reserved for long-time society members who "have made valuable contributions to the field."

Melany Hunt, assistant professor of mechanical engineering, was awarded the Pi Tau Sigma Medal by the American Society of Mechanical Engineers to recognize her outstanding achievement within 10 years of graduation.

Barbara Imperiali, assistant professor of chemistry, has received a Camille and Henry Dreyfus Teacher-Scholar Award, a grant of $60,000 to support young faculty members who are outstanding teachers as well as researchers.

Hiroo Kanamori, the Smits Professor of Geophysics and director of the Seismological Laboratory, has been named the 1993 California Scientist of the Year by the California Museum of Science and Industry.

Rudy Marcus, the Noyes Professor of Chemistry and Nobel laureate, has been elected the first Foreign Fellow of the Royal Society of Canada.

Stephen Mayo, assistant professor of biology, has been selected as a Rita Allen Foundation Scholar, an honor that includes research support of $30,000 annually for up to five years.

David Rutledge, professor of electrical engineering, has been named a Fellow of the Institute of Electrical and Electronics Engineers, and he and his students have received the 1993 IEEE Microwave Prize for the best paper published in its journals.

The ASCIT Teaching Awards, recognizing extraordinary "enthusiasm, dedication, quality of teaching, and interest in students," were won this year by William Deverell, visiting assistant professor of history; Glen George, lecturer in computer science and electrical engineering; Henry Lester, professor of biology; Mary Lidstrom, professor of applied microbiology; Anthony Readhead, professor of astronomy; Hunter Snevily, Bateman Research Instructor in Mathematics; Edward Zukoski, professor of jet propulsion and mechanical engineering; and graduate teaching assistants Marcia France of chemistry and Sima Setayeshgar of physics.

The Graduate Student Council presented its first Awards of Excellence in Teaching to Jim Knowles, the Kenan Professor and professor of applied mechanics; Charles Peck, professor of physics; David Stevenson, professor of planetary sciences and division chairman; and P. P. Vaidyanathan, professor of electrical engineering.

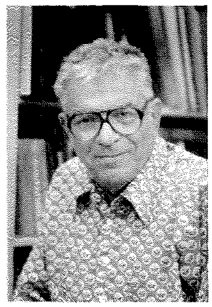
New Director of Development Appointed

J. Ernest "Jerry" Nunnally has been named assistant vice president and director of development and will take up his Caltech fund-raising responsibilities in October. Nunnally comes to the Institute from Harvard University, where he has been on the development staff since 1985, most recently as associate director of university development.

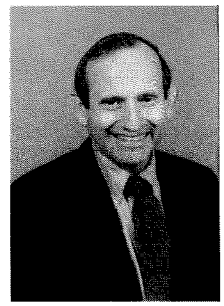
At Harvard, Nunnally has also held the positions of director of school relations and director of corporations and foundations. Previously he had worked at Dartmouth College, Continental Illinois National Bank, and Dillard University. Nunnally received his BA in 1969 from Dillard and a master's of education degree from Harvard in 1984.

Obituaries

Robert V. Langmuir 1912–1993



Robert Langmuir



Edward Posner

Robert V. Langmuir, professor of electrical engineering, emeritus, and the codiscoverer of synchrotron radiation, died May 1. He had been a Caltech faculty member for 45 years.

Langmuir earned his bachelor's degree from Harvard University in 1935 and his PhD from Caltech in 1943. From 1942 till 1948 he worked at the General Electric Company's research laboratory, where he and his colleagues reported the discovery, in 1947, that visible radiation was emitted tangentially from the circular orbits of electrons in the synchrotron, a phenomenon arising from the acceleration of highly relativistic electrons in the synchrotron's magnetic field.

In 1948 Langmuir returned to Caltech as a senior research fellow, where for the next 12 years he played an important role in designing and constructing a much higher energy synchrotron—for many years the highest electron accelerator in the world. Langmuir was mainly responsible for the radio-frequency power systems.

He was named assistant professor of electrical engineering in 1950, associate professor in 1952, and professor in 1957. Until his retirement in 1980 he taught, among others, courses in electricity and magnetism and in electronics, while continuing his research on various topics in applied physics and engineering. He served as head of electrical engineering from 1960 to 1970.

Langmuir's family has suggested that contributions to Caltech in Langmuir's memory may be sent to the Development Office (Caltech 105-40, Pasadena 91125).

Edward C. Posner 1933–1993

Edward C. Posner, visiting professor of engineering at Caltech and the Jet Propulsion Laboratory's chief technologist for the office of telecommunications and data acquisition, was hit by a truck and killed as he bicycled to JPL on the morning of June 15, 1993.

A memorial service will be held on campus September 28, at 2:00 p.m. in Dabney Lounge. Contributions in Posner's memory may be sent to the Director of Special Gifts (Caltech 105-40, Pasadena, CA 91125) for the Edward C. Posner SURF Memorial Fund to provide fellowship support for a student working on a research project in Posner's field.

Posner earned his BA in physics, and his MS and PhD in mathematics at the University of Chicago in 1952, 1953, and 1957 respectively. He taught mathematics at the University of Wisconsin and at Harvey Mudd College before joining JPL as a technologist in 1961. He became chief technologist in 1982. He taught at Caltech as a lecturer in electrical engineering at Caltech from 1970 to 1977, was appointed visiting associate professor in 1977, and had been a visiting professor since 1978.

Posner's specialty was information and communication theory. At JPL, his work on coding theory and data compression has enabled the volume of data returned from spacecraft over the Deep Space Network to be increased by several orders of magnitude. His campus research interests included communication network design, and automatic switching systems for such applications as cellular telephones.

The Caltech electrical engineer was one of the founders of research into neural networks at Caltech and JPL in the early 1980s, and was instrumental in the creation of Caltech's interdisciplinary graduate-study program in Computation and Neural Systems, the first program of its kind in the world.

A dedicated supporter of undergraduate research, Posner was particularly involved with Caltech's SURF (Summer Undergraduate Research Fellowships) program. Since 1984 he had sponsored 13 SURF students, not counting three who had just started this summer's work, and since 1990 had been a member of the SURF administrative committee. In 1986 he cofounded the SURFSAT satellite program and between 1988 and 1991 cosponsored 43 SURFSAT students at JPL.

Take a year...
to make a difference.

If you're interested in an opportunity to work in the White House and deal one-on-one with America's leaders, take a close look at the White House Fellowships.

White House Fellows are a select group of men and women who spend a year early in their careers serving as paid assistants to the President, Vice-President, or cabinet-level officials.

They are people of exceptional abilities, strong motivation and a desire to serve their country.

White House Fellows have gone on to become leaders in many fields... business, politics, science and the arts. And they all agree on one thing: Their year as a Fellow changed their lives.

If it often seems as though you have to choose between helping others, and helping yourself—here's a chance to do both. As a White House Fellow, you can serve America while learning skills you will use your entire life.

And instead of just reading about world events, you can help shape them.

Call (202) 395-4522 for a brochure, application and the opportunity of a lifetime.

THE WHITE HOUSE FELLOWSHIPS

Engineering & Science

California Institute of Technology
Pasadena, California 91125

ADDRESS CORRECTION REQUESTED