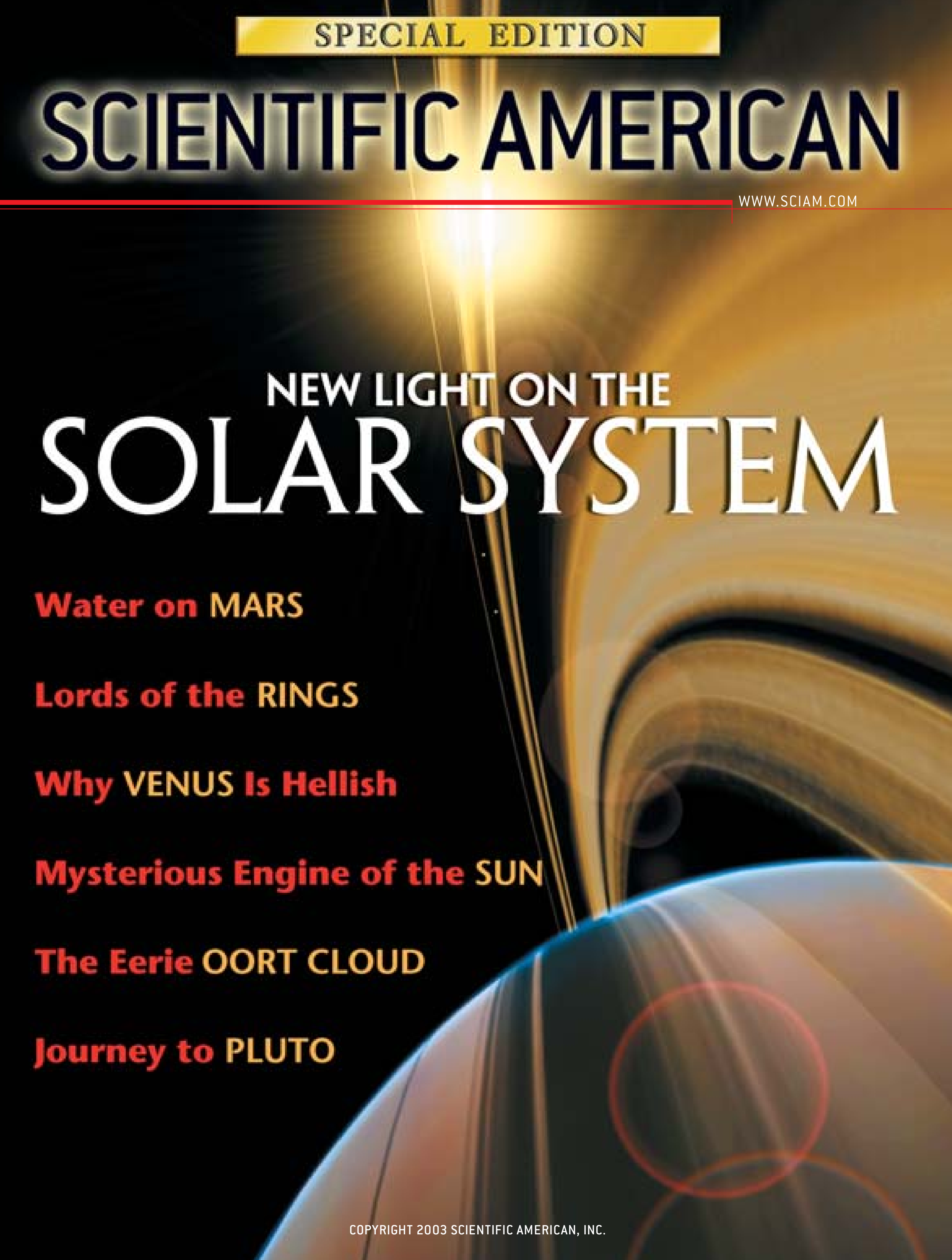


SPECIAL EDITION

SCIENTIFIC AMERICAN

WWW.SCIAM.COM

NEW LIGHT ON THE SOLAR SYSTEM

Water on MARS

Lords of the RINGS

Why VENUS Is Hellish

Mysterious Engine of the SUN

The Eerie OORT CLOUD

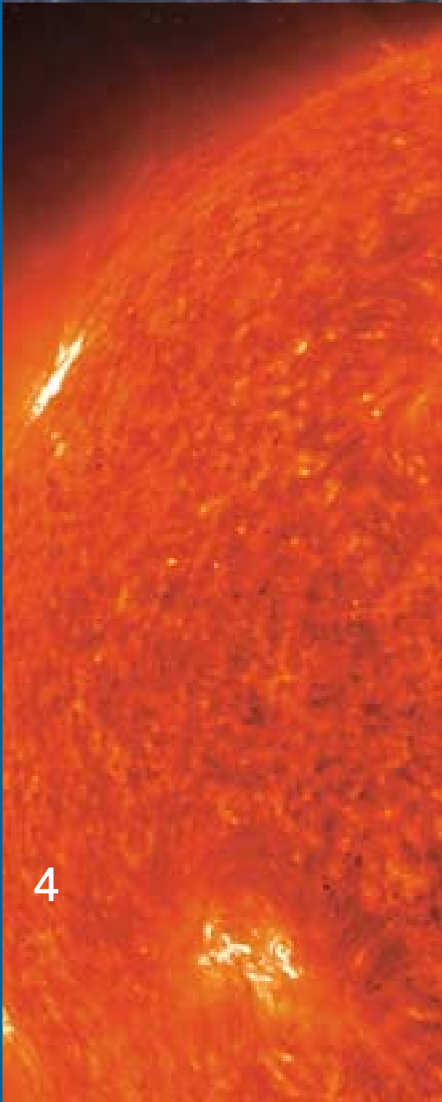
Journey to PLUTO

2003 contents

SCIENTIFIC AMERICAN Volume 13 Number 3

New Light on the Solar System


28


4

12

44 The Small Planets

By ERIK ASPHAUG

Asteroids have become notorious as celestial menaces but are best considered in a positive light, as surreal worlds bearing testimony to the origin of the planets.

54 The Galileo Mission to Jupiter and Its Moons

By TORRENCE V. JOHNSON

Few scientists thought that the Galileo spacecraft could conduct such a comprehensive study of the Jovian system. And few predicted that these worlds would prove so varied.

64 The Hidden Ocean of Europa

By ROBERT T. PAPPALARDO, JAMES W. HEAD AND RONALD GREELEY

Doodles and freckles, creamy plains and crypto-icebergs—the amazing surface of Jupiter's brightest icy moon hints at a global sea underneath.

74 Bejeweled Worlds

By JOSEPH A. BURNS, DOUGLAS P. HAMILTON AND MARK R. SHOWALTER

Small moons sculpt elegant, austere rings around Jupiter, Saturn, Uranus, Neptune and maybe even Mars.

84 Journey to the Farthest Planet

By S. ALAN STERN

Scientists are finally preparing to send a spacecraft to Pluto and the Kuiper belt, the last unexplored region in our planetary system.

92 The Oort Cloud

By PAUL R. WEISSMAN

On the outskirts of the solar system swarms a vast cloud of comets. The dynamics of this cloud may help explain such matters as mass extinctions on Earth.

Cover painting by Don Dixon.

Scientific American Special (ISSN 1048-0943), Volume 13, Number 3, 2003, published by Scientific American, Inc., 415 Madison Avenue, New York, NY 10017-1111. Canadian BN No. 127387652RT; QST No. Q1015332537. To purchase additional quantities: U.S., $10.95 each; elsewhere, $13.95 each. Send payment to Scientific American, Dept. SOL03, 415 Madison Avenue, New York, NY 10017-1111. Inquiries: fax 212-355-0408 or telephone 212-451-8890. Printed in U.S.A.

SCIENTIFIC AMERICAN®
Established 1845

New Light on the Solar System is published by the staff of SCIENTIFIC AMERICAN, with project management by:

EDITOR IN CHIEF: John Rennie
EXECUTIVE EDITOR: Mariette DiChristina
ISSUE EDITOR: Mark Fischetti
ISSUE CONSULTANT: George Musser

ART DIRECTOR: Edward Bell
ISSUE DESIGNER: Jessie Nathans
PHOTOGRAPHY EDITOR: Bridget Gerety
PRODUCTION EDITOR: Richard Hunt

COPY DIRECTOR: Maria-Christina Keller
COPY CHIEF: Molly K. Frances
COPY AND RESEARCH: Daniel C. Schlenoff, Rina Bander, Michael Battaglia, Emily Harrison, David Labrador

EDITORIAL ADMINISTRATOR: Jacob Lasky
SENIOR SECRETARY: Maya Harty

ASSOCIATE PUBLISHER, PRODUCTION: William Sherman
MANUFACTURING MANAGER: Janet Cermak
ADVERTISING PRODUCTION MANAGER: Carl Cherebin
PREPRESS AND QUALITY MANAGER: Silvia Di Placido
PRINT PRODUCTION MANAGER: Georgina Franco
PRODUCTION MANAGER: Christina Hippeli
CUSTOM PUBLISHING MANAGER: Madelyn Keyes-Milch

ASSOCIATE PUBLISHER/VICE PRESIDENT, CIRCULATION: Lorraine Leib Terlecki
CIRCULATION DIRECTOR: Katherine Corvino
CIRCULATION PROMOTION MANAGER: Joanne Guralnick
FULFILLMENT AND DISTRIBUTION MANAGER: Rosa Davis

VICE PRESIDENT AND PUBLISHER: Bruce Brandfon
ASSOCIATE PUBLISHER: Gail Delott
SALES DEVELOPMENT MANAGER: David Tirpack
SALES REPRESENTATIVES: Stephen Dudley, Hunter Millington, Stan Schmidt, Debra Silver

ASSOCIATE PUBLISHER, STRATEGIC PLANNING: Laura Salant
PROMOTION MANAGER: Diane Schube
RESEARCH MANAGER: Aida Dadurian
PROMOTION DESIGN MANAGER: Nancy Mongelli
GENERAL MANAGER: Michael Florek
BUSINESS MANAGER: Marie Maher
MANAGER, ADVERTISING ACCOUNTING AND COORDINATION: Constance Holmes

DIRECTOR, SPECIAL PROJECTS: Barth David Schwartz

MANAGING DIRECTOR, ONLINE: Mina C. Lux
SALES REPRESENTATIVE, ONLINE: Gary Bronson
WEB DESIGN MANAGER: Ryan Reid

DIRECTOR, ANCILLARY PRODUCTS: Diane McGarvey
PERMISSIONS MANAGER: Linda Hertz
MANAGER OF CUSTOM PUBLISHING: Jeremy A. Abbate

CHAIRMAN EMERITUS: John J. Hanley
CHAIRMAN: Rolf Grisebach
PRESIDENT AND CHIEF EXECUTIVE OFFICER: Gretchen G. Teichgraeber
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL: Dean Sanderson
VICE PRESIDENT: Frances Newburg

letter from the editor

the planets in our backyard

LET'S TALK FOR A MOMENT about our immediate neighborhood. A radio signal sweeps from Earth to the moon in just over one and a quarter seconds and from Earth to Mars in as little as three minutes. Even Pluto is only about six hours away at light speed; if you packed a lunch and caught a round-trip sunbeam, you could get to Pluto and back without missing a meal. The gulf to the closest star, Proxima Centauri, however, is a depressingly vast 4.3 light-years.

On the scale of the Milky Way, 100,000 light-years across, our solar system can seem like a puny rut in which to be stuck. Having glimpsed countless exotic stars and galaxies, surely the human imagination will rapidly weary of just one yellow sun, eight or nine planets (depending on your feelings about Pluto), and a loose assortment of moons and debris.

Yet the more we learn about our solar system, the more fascinating it becomes. The sun's atmosphere is hotter than its surface. Venus suffers from a greenhouse effect run amok. On Mars, geologic forces unlike those seen on Earth help to sculpt the landscape. Tiny moons stabilize the ethereal rings around the gas giants. Jupiter's satellite Europa has icy niches where life might evolve. (As this issue goes to press, astronomers are remarking that as Pluto's orbit carries it farther from the sun, the planet's atmosphere is curiously warming up.)

Though astronomers have begun to detect planetary systems around other stars, the uniqueness of ours is so far intact. Many planets in far-off systems seem to be freakishly large and moving in bizarre orbits that would devastate any alien Earths out there. One of the greatest mysteries of our solar system may be why it is so stable.

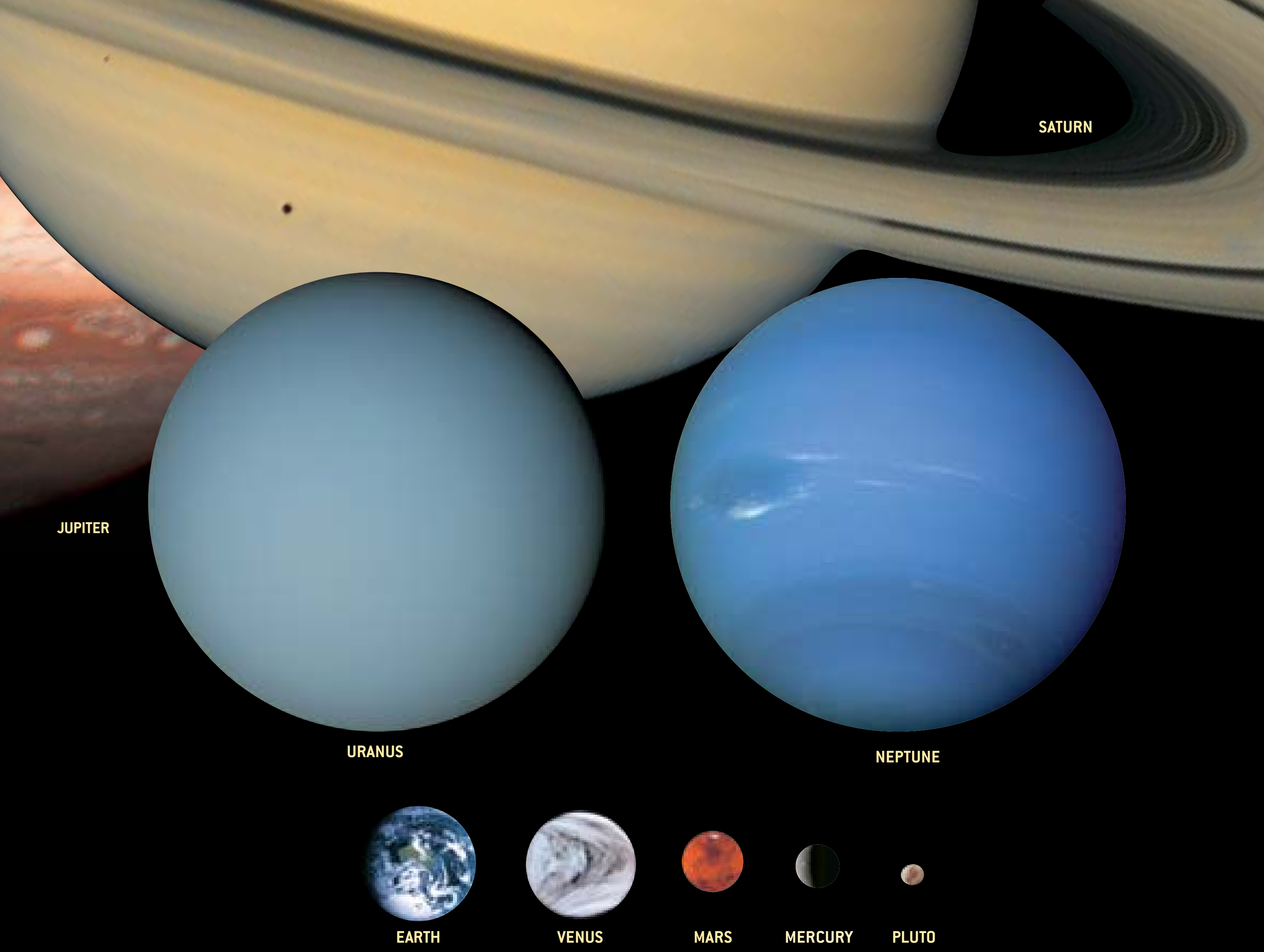
This special edition of *Scientific American* provides the latest developments about our corner of the cosmos, in articles written by the experts who are leading the investigations. Let the pages that follow guide your tour of our solar system, and savor the fact that you can visit these extraordinary nearby worlds and still be home for supper.

John Rennie
Editor in Chief
Scientific American
editors@sciam.com

the planets at a glance

	MERCURY	VENUS
AVERAGE DISTANCE FROM SUN (kilometers)	57.9 million	108.2 million
EQUATORIAL DIAMETER (kilometers)	4,879	12,103.6
MASS (kilograms)	3.3×10^{23}	4.9×10^{24}
DENSITY (grams per cubic centimeter)	5.41	5.25
LENGTH OF DAY (relative to Earth)	58.6 days	243.0 days
LENGTH OF YEAR (relative to Earth)	87.97 days	224.7 days
NUMBER OF KNOWN MOONS	0	0
ATMOSPHERIC COMPOSITION	Traces of sodium, helium and oxygen	96% carbon dioxide, 3.5% nitrogen

JPL/CALTECH/NASA (*all images*); LAURIE GRACE (*table*)



relative sizes of the planets in the solar system

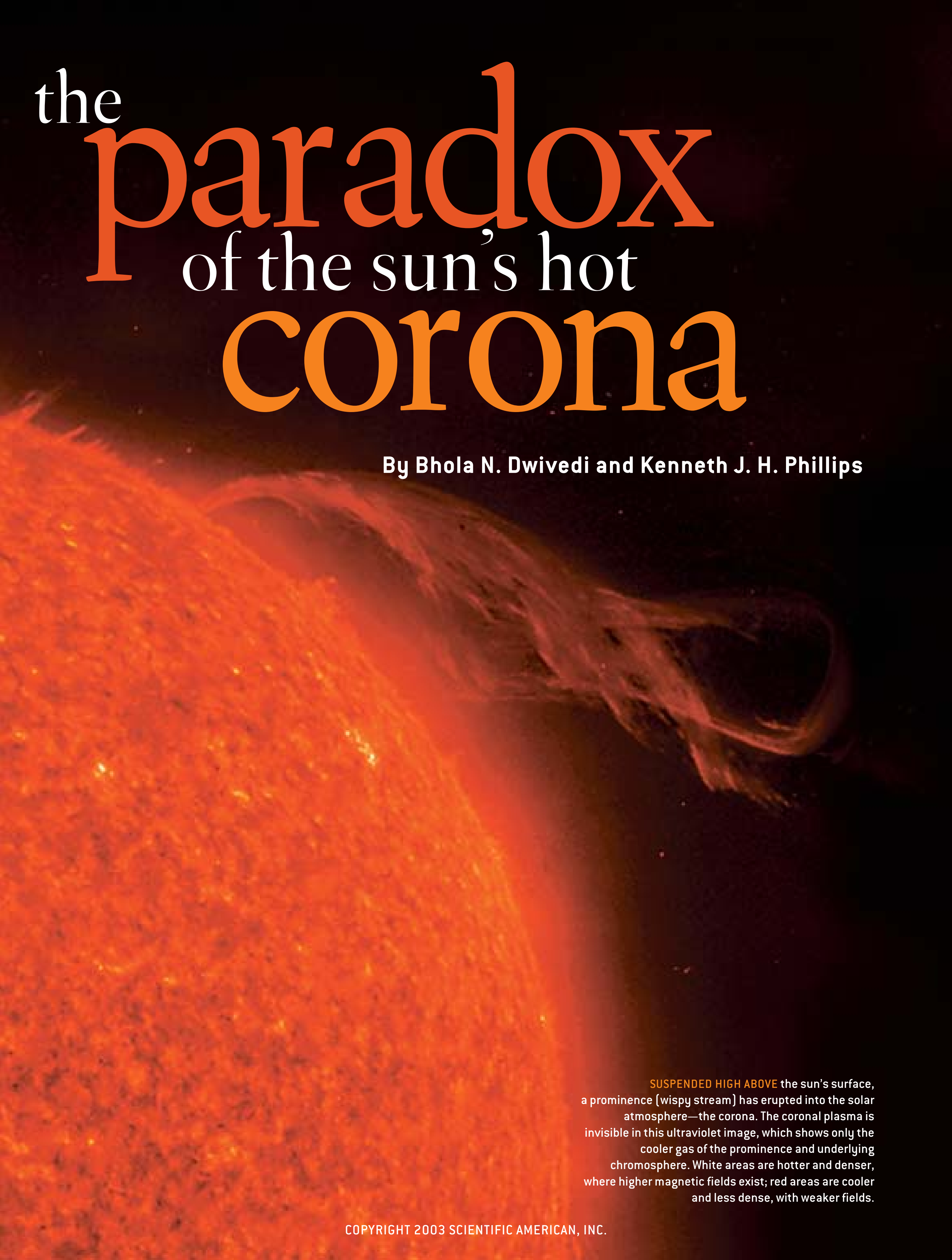
	EARTH	MARS	JUPITER	SATURN	URANUS	NEPTUNE	PLUTO
	149.6 million	227.94 million	778.3 million	1,429.4 million	2,871 million	4,504.3 million	5,913.5 million
	12,756.28	6,794.4	142,984	120,536	51,118	49,492	2,274
	6.0×10^{24}	6.4×10^{23}	1.9×10^{27}	5.7×10^{26}	8.7×10^{25}	1.0×10^{26}	1.3×10^{22}
	5.52	3.9	1.3	0.7	1.3	1.6	2.05
	23.93 hours	24.62 hours	9.92 hours	10.2 hours	17.9 hours	16.1 hours	6.39 days
	365.26 days	686.98 days	11.86 years	29.46 years	84 years	164.8 years	248.5 years
	1	2	16	At least 18	At least 16	8	1
	78% nitrogen, 21% oxygen, 0.9% argon	95% carbon dioxide, 3% nitrogen, 1.6% argon	90% hydrogen, 10% helium, traces of methane	97% hydrogen, 3% helium, traces of methane	83% hydrogen, 15% helium, 2% methane	85% hydrogen, 13% helium, 2% methane	Probably methane, possibly nitrogen and carbon monoxide

Like a boiling teakettle atop a **COLD** stove,
the sun's **HOT** outer layers sit on the relatively cool surface.
And now astronomers are **FIGURING OUT WHY**

Updated from the June 2001 issue

the paradox of the sun's hot corona

By Bhola N. Dwivedi and Kenneth J. H. Phillips



SUSPENDED HIGH ABOVE the sun's surface, a prominence (wispy stream) has erupted into the solar atmosphere—the corona. The coronal plasma is invisible in this ultraviolet image, which shows only the cooler gas of the prominence and underlying chromosphere. White areas are hotter and denser, where higher magnetic fields exist; red areas are cooler and less dense, with weaker fields.

Relatively few people have witnessed a total eclipse of the sun—one of nature's most awesome spectacles. It was therefore a surprise for inhabitants of central Africa to see two total eclipses in quick succession, in June 2001 and December 2002. Thanks to favorable weather along the narrow track of totality across the earth, the 2001 event in particular captivated residents and visitors in Zambia's densely populated capital, Lusaka. One of us (Phillips), with colleagues from the U.K. and Poland, was also blessed with scientific equipment that worked perfectly on location at the University of Zambia. Other scientific teams captured valuable data from Angola and Zimbabwe. Most of us were trying to find yet more clues to one of the most enduring conundrums of the solar system: What is the mechanism that makes the sun's outer atmosphere, or corona, so hot?

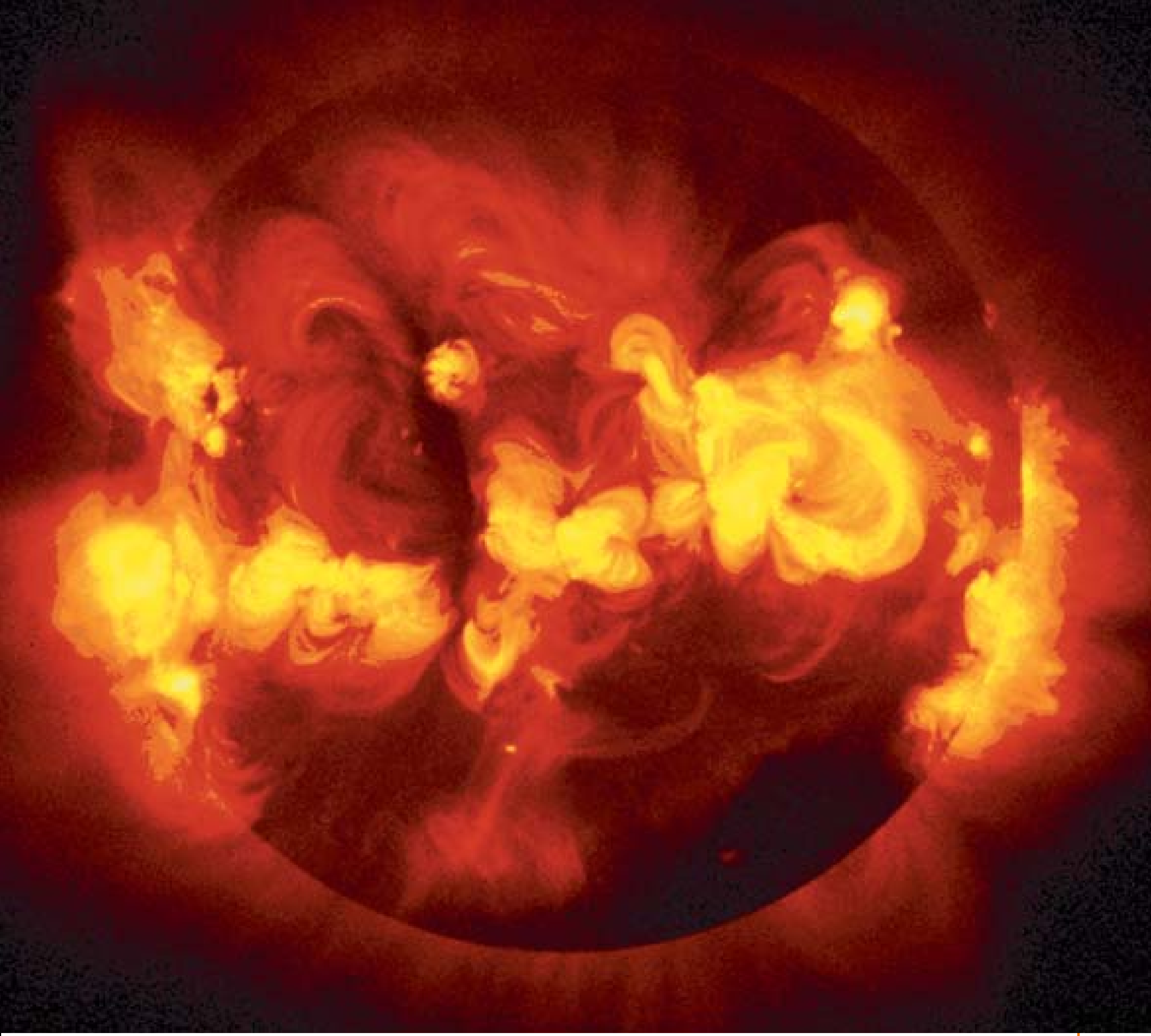
The sun might appear to be a uniform sphere of gas, the essence of simplicity. In actuality it has well-defined layers that can loosely be compared to a planet's solid part and atmosphere. The solar radiation that we receive ultimately derives from nuclear reactions deep in the core. The energy gradually leaks out until it reaches the visible surface, known as the photosphere, and escapes into space. Above that surface is a tenuous atmosphere. The lowest part, the chromosphere, is usually visible only during total eclipses, as a bright red crescent. Beyond it is the pearly white corona, extending millions of kilometers. Further still, the corona becomes a stream of charged particles—the solar wind that blows through our solar system.

Journeying out from the sun's core, an imaginary observer first encounters temperatures of 15 million kelvins, high enough to generate the nuclear reactions that power the sun. Temperatures get progressively cooler en route to the photosphere, a mere 6,000 kelvins. But then an unexpected thing happens: the temperature gradient reverses. The chromosphere's temperature steadily rises to 10,000 kelvins, and going into the corona, the temperature jumps to one million kelvins. Parts of the corona associated with sunspots get even hotter. Considering that the energy must originate below the photosphere, how can this be? It is as if you got warmer the farther away you walked from a fireplace.

The first hints of this mystery emerged in the 19th century when eclipse observers detected spectral emission lines that no known element could account for. In the 1940s physicists associated two of these lines with iron atoms that had lost up to half their normal retinue

NASA GODDARD SPACE FLIGHT CENTER (*preceding pages*); TRACE/NASA (*below*)

CORONAL LOOP, seen in ultraviolet light by the TRACE spacecraft, extends 120,000 kilometers off the sun's surface.



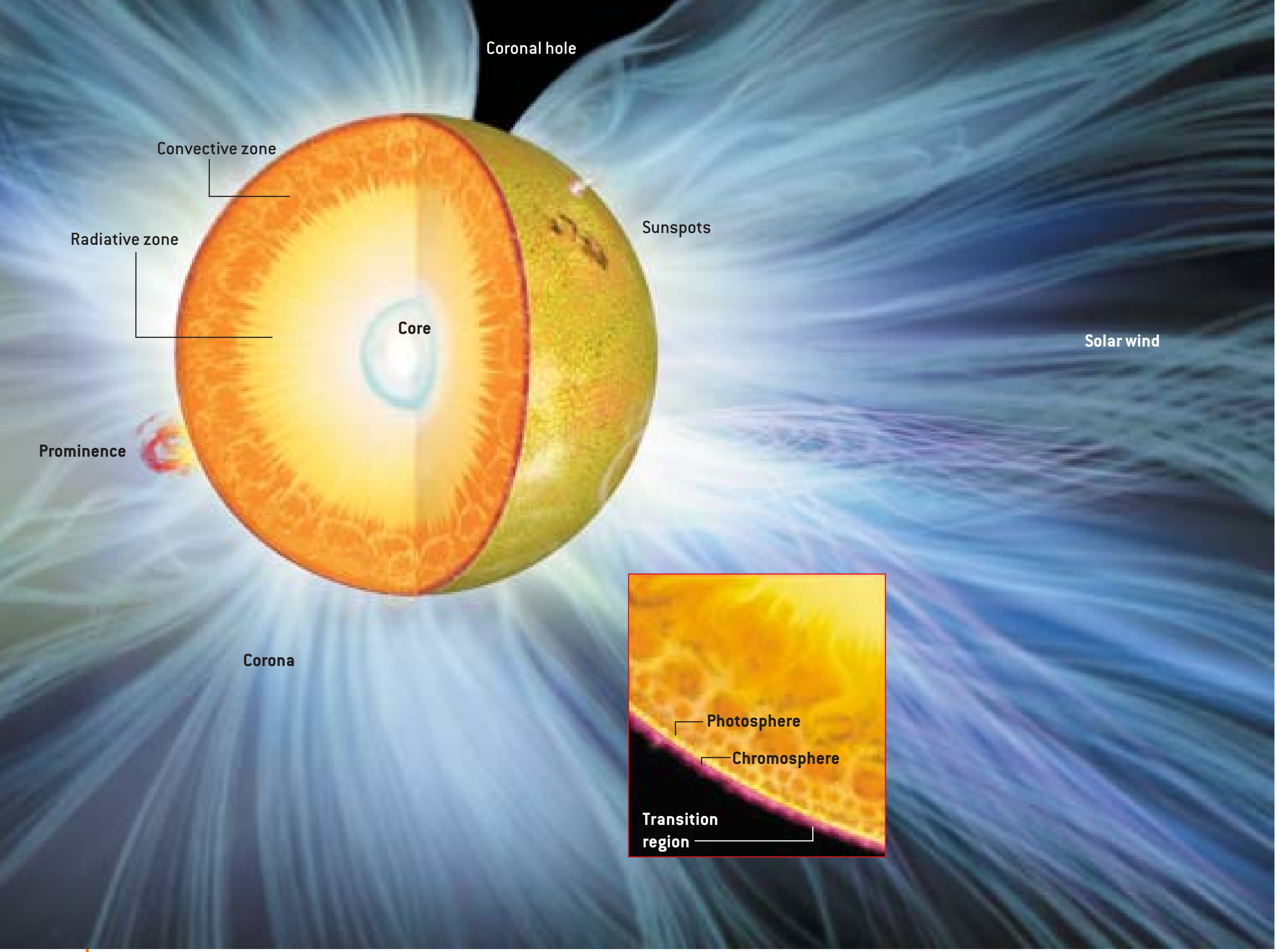
INSTITUTE OF SPACE AND ASTRONAUTICAL SCIENCE, JAPAN; LOCKHEED MARTIN SOLAR AND ASTROPHYSICS LABORATORY; NATIONAL ASTRONOMICAL OBSERVATORY OF JAPAN; UNIVERSITY OF TOKYO; NASA

X-RAY IMAGE from the Yohkoh spacecraft shows structures both bright (associated with sunspots) and dark (polar coronal holes).

of 26 electrons—a situation that requires extremely high temperatures. Later, instruments on rockets and satellites found that the sun emits copious x-rays and extreme ultraviolet radiation—as can be the case only if the coronal temperature is measured in megakelvins. Nor is this mystery confined to the sun: most sunlike stars appear to have x-ray-emitting atmospheres.

At last, however, a solution seems to be within our grasp. Astronomers have long implicated magnetic fields in the coronal heating; where those fields are strongest, the corona is hottest. Such fields can transport energy in a form other than heat, thereby sidestepping the usual thermodynamic restrictions. The energy must still be converted to heat, and researchers are testing two possible theories: small-scale magnetic field reconnections—the same process involved in solar flares—and magnetic waves. Important clues have come from complementary observations: spacecraft can observe at wavelengths inaccessible from the ground, while ground-based telescopes can gather reams of data unrestricted by the bandwidth of orbit-to-Earth radio links. The findings may be crucial to understanding how events on the sun affect the atmosphere of Earth [see "The Fury of Space Storms," by James L. Burch; Scientific American, April 2001].

The first high-resolution images of the corona came from the ultraviolet and x-ray telescopes on board Skylab, the American space station inhabited in 1973 and



FAR FROM A UNIFORM BALL of gas, the sun has a dynamic interior and atmosphere that heat and light our solar system.

DON DIXON

1974. Pictures of active regions of the corona, located above sunspot groups, revealed complexes of loops that came and went in a matter of days. Much larger but more diffuse x-ray arches stretched over millions of kilometers, sometimes connecting sunspot groups. Away from active regions, in the "quiet" parts of the sun, ultraviolet emission had a honeycomb pattern related to the large convection granules in the photosphere. Near the solar poles and sometimes in equatorial locations were areas of very faint x-ray emission—the so-called coronal holes.

Connection to the Starry Dynamo

EACH MAJOR SOLAR SPACECRAFT since Skylab has offered a distinct improvement in resolution. From 1991 to late 2001, the x-ray telescope on the Japanese Yohkoh spacecraft routinely imaged the sun's corona, tracking the evolution of loops and other features through one complete 11-year cycle of solar activity. The Solar and Heliospheric Observatory (SOHO), a joint European-American satellite launched in 1995, orbits a point 1.5 million kilometers from Earth on its sunward side, giving the spacecraft the advantage of an uninterrupted view of the sun [see "SOHO Reveals the Secrets of the Sun," by Kenneth R. Lang; SCIENTIFIC AMERICAN, March 1997]. One of its instruments, called the Large Angle and Spectroscopic Coronagraph (LASCO), observes in visible light using an opaque disk to mask out the main part of the sun. It has tracked large-scale coronal structures as they rotate with the rest of the sun (a period of about 27 days as seen from Earth). The images show huge bubbles of plasma known as coronal mass ejections, which move at up to 2,000 kilometers a second, erupting from the corona and occasionally colliding with Earth and other planets. Other SOHO instruments, such as the Extreme Ultraviolet Imaging Telescope, have greatly improved on Skylab's pictures.

The Transition Region and Coronal Explorer (TRACE) satellite, operated by the Stanford-Lockheed Institute for Space Research, went into a polar orbit around Earth in 1998. With unprecedented resolution, its ultraviolet telescope has revealed a vast wealth of detail. The active-region loops are now known to be threadlike features no more than a few hundred kilometers wide. Their incessant flickering and jouncing hint at the origin of the corona's heating mechanism.

The latest spacecraft dedicated to the sun is the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI), launched in 2002, which is providing images and spectra in the x-ray region of wavelengths less than four nanometers. Because solar activity has been high, much of its early attention was focused on intense flares, but as the solar minimum approaches, investigators will increasingly be interested in tiny microflares, a clue to the corona's heating mechanism.

The loops, arches and coronal holes trace out the sun's magnetic fields. The fields are thought to originate in the upper third of the solar interior, where energy is transported mostly by convection rather than radiation. A combination of convection currents and differential rotation—whereby low latitudes rotate slightly faster than higher latitudes—twist the fields to form ropelike or other tightly bound configurations that eventually emerge at the photosphere and into the solar atmosphere. Particularly intense fields are marked by sunspot groups and active regions.

For a century, astronomers have measured the magnetism of the photosphere using magnetographs, which observe the Zeeman effect: in the presence of a magnetic field, a spectral line can split into two or more lines with slightly different wavelengths and polarizations. But Zeeman observations for the corona have yet to be done. The spectral splitting is too small to be detected with present instruments, so astronomers have had to resort to mathematical extrapolations from the photospheric field. These predict that the magnetic field of the corona generally has a strength of about 10 gauss, 20 times Earth's magnetic field strength at its poles. In active regions, the field may reach 100 gauss.

Space Heaters

THESE FIELDS ARE WEAK compared with those that can be produced with laboratory magnets, but they have a decisive influence in the solar corona. This is because the corona's temperature is so high that it is almost fully ionized: it is a plasma, made up not of neutral atoms but of electrons, protons and other atomic nuclei. Plasmas undergo a wide range of phenomena that neutral gases do not. The magnetic fields of the corona are strong enough to bind the charged particles to the field lines. Particles move in tight helical paths up and down these field lines like very small beads on very long strings. The limits on their motion explain the sharp boundaries of features such as coronal holes. Within the tenuous plasma, the magnetic pressure (proportional to the strength squared) exceeds the thermal pressure by a factor of at least 100.

One of the main reasons astronomers are confident that magnetic fields energize the corona is the clear relation between field strength and temperature. The bright loops of active regions, where there are extremely strong fields, have a temperature of about four million kelvins. But the giant arches of the quiet-sun corona, characterized by weak fields, have a temperature of about one million kelvins.

Until recently, however, ascribing coronal heating to magnetic fields ran into a serious problem. To convert field energy to heat energy, the fields must be able to diffuse through the plasma, which requires that the corona have a certain amount of electrical resistivity—in other words, that it not be a perfect conductor. A perfect conductor cannot sustain an electric field, because charged particles instantaneously reposition themselves to neutralize it. And if a plasma cannot sustain an electric field, it cannot move relative to the magnetic field (or vice versa), because to do so would induce an electric field. This is why astronomers talk about magnetic fields being "frozen" into plasmas.

This principle can be quantified by considering the time it takes a magnetic field to diffuse a certain distance through a plasma. The diffusion rate is inversely proportional to resistivity. Classical plasma physics assumes that electrical resistance arises from so-called Coulomb collisions: electrostatic forces from charged particles deflect the flow of electrons. If so, it should take about 10 million years to traverse a distance of 10,000 kilometers, a typical length of active-region loops.

Events in the corona—for example, flares, which may last for only a few minutes—far outpace that rate. Either the resistivity is unusually high or the diffusion distance is extremely small, or both. A distance as short as a few meters could occur in certain structures, accompanied by a steep magnetic gradient. But researchers have come to realize that the resistivity could be higher than they traditionally thought.

Raising the Mercury

ASTRONOMERS HAVE TWO basic ideas for coronal heating. For years, they concentrated on heating by

THE AUTHORS

BHOLA N. DWIVEDI and *KENNETH J. H. PHILLIPS* began collaborating on solar physics a decade ago. Dwivedi teaches physics at Banaras Hindu University in Varanasi, India. He has been working with SUMER, an ultraviolet telescope on the SOHO spacecraft, for more than 10 years; the Max Planck Institute for Aeronomy near Hannover, Germany, recently awarded him one of its highest honors, the Gold Pin. As a boy, Dwivedi studied by the light of a homemade burner and became the first person in his village ever to attend college. Phillips recently left the Rutherford Appleton Laboratory in England to become a senior research associate in the Reuven Ramaty High Energy Solar Spectroscopic Imager group at the NASA Goddard Space Flight Center in Greenbelt, Md. He has worked with x-ray and ultraviolet instruments on numerous spacecraft—including OSO-4, SolarMax, IUE, Yohkoh, Chandra and SOHO—and has observed three solar eclipses using CCD cameras.

waves. Sound waves were a prime suspect, but in the late 1970s researchers established that sound waves emerging from the photosphere would dissipate in the chromosphere, leaving no energy for the corona itself. Suspicion turned to magnetic waves. Such waves might be purely magnetohydrodynamic (MHD)—so-called Alfvén waves—in which the field lines oscillate but the pressure does not. More likely, however, they share characteristics of both sound and Alfvén waves.

MHD theory combines two theories that are challenging in their own right—ordinary hydrodynamics and electromagnetism—although the broad outlines are clear. Plasma physicists recognize two kinds of MHD pressure waves, fast and slow mode, depending on the phase velocity relative to an Alfvén wave—around 2,000 kilometers a second in the corona. To traverse a typical active-region loop requires about five seconds for an Alfvén wave, less for a fast MHD wave, but at least half a minute for a slow wave. MHD waves are set into motion by convective perturbations in the photosphere and transported out into the corona via magnetic fields. They can then deposit their energy into the plasma if it has sufficient resistivity or viscosity.

A breakthrough occurred in 1998 when the TRACE spacecraft observed a powerful flare that triggered waves in nearby fine loops. The loops oscillated back and forth several times before settling down. The damping rate was millions of times as fast as classical theory predicts. This landmark observation of "coronal seismology" by Valery M. Nakariakov, then at the University of St. Andrews in Scotland, and his colleagues has shown that MHD waves could indeed deposit their energy into the corona.

An intriguing observation made with the ultraviolet coronagraph on the SOHO spacecraft has shown that highly ionized oxygen atoms have temperatures in coronal holes of more than 100 million kelvins, much higher than those of electrons and protons in the plasma. The temperatures also seem higher perpendicular to the magnetic field lines than parallel to them. Whether this is important for coronal heating remains to be seen.

Despite the plausibility of energy transport by waves, a second idea has been ascendant: that coronal heating is caused by very small, flarelike events. A flare is a sudden release of up to 10^{25} joules of energy in an active region of the sun. It is thought to be caused by reconnection of magnetic field lines, whereby oppositely directed lines cancel each other out, converting magnetic energy into heat. The process requires that the field lines be able to diffuse through the plasma.

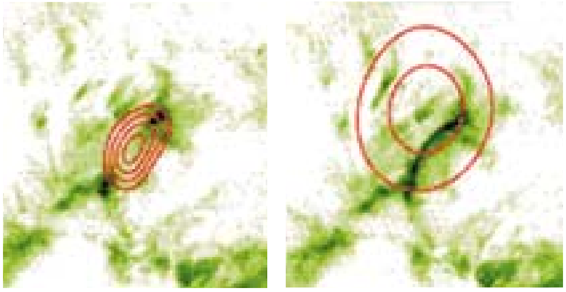
A flare sends out a blast of x-rays and ultraviolet radiation. At the peak of the solar cycle (reached in 2000), several flares an hour may burst out across the sun. Spacecraft such as Yohkoh and SOHO have shown that much smaller but more frequent events take place not only in active regions but also in regions otherwise deemed quiet. These tiny events have about a millionth the energy of a full-blown flare and so are called microflares. They were first detected in 1980 by Robert P. Lin of the University of California at Berkeley and his colleagues with a balloon-borne hard x-ray detector. During the solar minimum in 1996, Yohkoh also recognized events with energy as small as 0.01 of a microflare.

Early results from the RHESSI measurements indicate more than 10 hard x-ray microflares an hour. In addition, RHESSI can produce images of microflares, which was not possible before. As solar activity declines, RHESSI should be able to locate and characterize very small flares.

Flares are not the only type of transient phenomena. X-ray and ultraviolet jets, representing columns of coronal material, are often seen spurting up from the lower corona at a few hundred kilometers a second. But tiny x-ray flares are of special interest because they reach the megakelvin temperatures required to heat the corona. Several researchers have attempted to extrapolate the microflare rates to even tinier nanoflares, to test an idea raised some years ago by Eugene Parker of the University of Chicago that numerous nanoflares occurring outside of active regions could account for the entire energy of the corona. Results remain confusing, but perhaps the combination of RHESSI, TRACE and SOHO data during the forthcoming minimum can provide an answer.

Which mechanism—waves or nanoflares—dominates? It depends on the photospheric motions that perturb the magnetic field. If these motions operate on timescales of half a minute or longer, they cannot trigger MHD waves. Instead they create narrow current sheets in which reconnections can occur. Very high resolution optical observations of bright filigree structures by the Swedish Vacuum Tower Telescope on La Palma in the Canary Islands—as well as SOHO and

X-RAY IMAGE taken by the RHESSI spacecraft outlines the progression of a microflare on May 6, 2002. The flare peaked (*left*), then six minutes later (*right*) began to form loops over the original flare site.



COURTESY OF SAM KRUCKER *University of California, Berkeley*

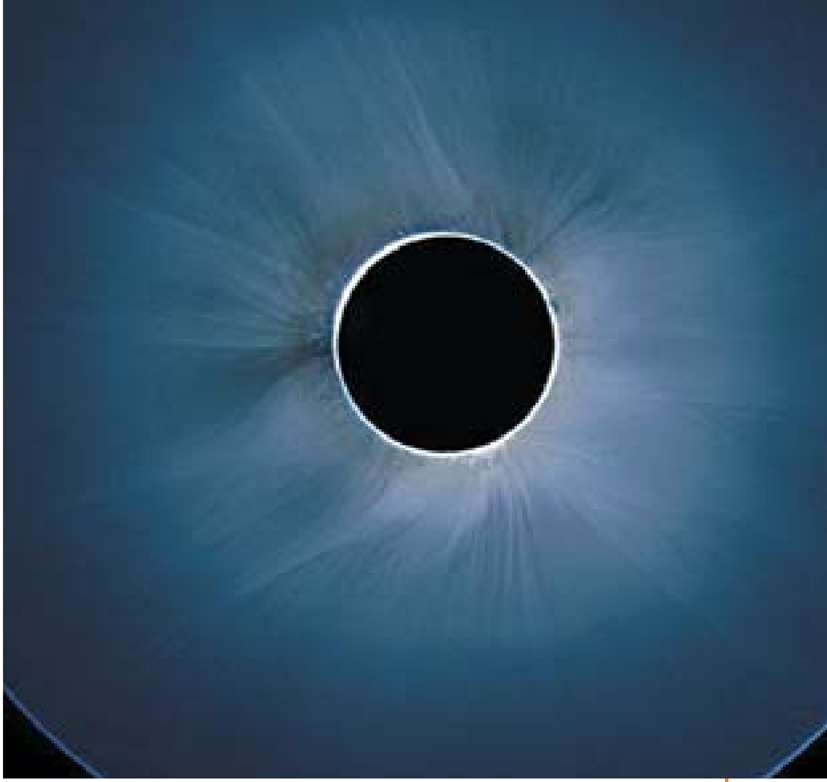
TRACE observations of a general, ever changing "magnetic carpet" on the surface of the sun—demonstrate that motions occur on a variety of timescales. Although the evidence now favors nanoflares for the bulk of coronal heating, waves may also play a role.

Fieldwork

IT IS UNLIKELY, for example, that nanoflares have much effect in coronal holes. In these regions, the field lines open out into space rather than loop back to the sun, so a reconnection would accelerate plasma out into interplanetary space rather than heat it. Yet the corona in holes is still hot. Astronomers have scanned for signatures of wave motions, which may include periodic fluctuations in brightness or Doppler shift. The difficulty is that the MHD waves involved in heating probably have very short periods, perhaps just a few seconds. At present, spacecraft imaging is too sluggish to capture them.

For this reason, ground-based instruments remain important. A pioneer in this work has been Jay M. Pasachoff of Williams College. He and his students have used high-speed detectors and CCD cameras to look for modulations in the coronal light during eclipses. Analyses of his best results indicate oscillations with periods of one to two seconds. Serge Koutchmy of the Institute of Astrophysics in Paris, using a coronagraph, has found evidence of periods equal to 43, 80 and 300 seconds.

The search for those oscillations is what led Phillips and his colleagues to Bulgaria in 1999 and Zambia in 2001. Our instrument consists of a pair of fast-frame CCD cameras that observe both white light and the green spectral line produced by highly ionized iron. A tracking mirror, or heliostat, directs sunlight into a horizontal beam that passes into the instrument. At our observing sites, the 1999 eclipse totality lasted two minutes and 23 seconds, the 2001 totality three minutes and 38 seconds. Analyses of the 1999 eclipse by David A. Williams, now at University College London, reveal the possible presence of an MHD wave with fast-mode characteristics moving down a looplike structure. The CCD signal for this eclipse is admittedly weak, however, and Fourier analysis by Pawel Rudawy of the University of Wroclaw in Poland fails to find significant periodicities in the 1999 and 2001 data. We continue to try to determine if there are other, nonperiodic changes.



ORDINARY LIGHT, EXTRAORDINARY SIGHT: The corona is photographed in visible light on August 11, 1999, from Chadegan in central Iran.

JEAN MOUETTE AND SERGE KOUTCHMY, © INSTITUT D'ASTROPHYSIQUE DE PARIS

Insight into coronal heating has also come from observations of other stars. Current instruments cannot see surface features of these stars directly, but spectroscopy can deduce the presence of starspots, and ultraviolet and x-ray observations can reveal coronae and flares, which are often much more powerful than their solar counterparts. High-resolution spectra from the Extreme Ultraviolet Explorer and the latest x-ray satellites, Chandra and XMM-Newton, can probe temperature and density. For example, Capella—a stellar system consisting of two giant stars—has photospheric temperatures like the sun's but coronal temperatures that are six times higher. The intensities of individual spectral lines indicate a plasma density of about 100 times that of the solar corona. This high density implies that Capella's coronae are much smaller than the sun's, stretching out a tenth or less of a stellar diameter. Apparently, the distribution of the magnetic field differs from star to star. For some stars, tightly orbiting planets might even play a role.

Even as one corona mystery begins to yield to our concerted efforts, additional ones appear. The sun and other stars, with their complex layering, magnetic fields and effervescent dynamism, still manage to defy our understanding. In an age of such exotica as black holes and dark matter, even something that seems mundane can retain its allure.

MORE TO EXPLORE

Guide to the Sun. Kenneth J. H. Phillips. Cambridge University Press, 1992.

The Solar Corona above Polar Coronal Holes as Seen by SUMER on SOHO. Klaus Wilhelm et al. in *Astrophysical Journal,* Vol. 500, No. 2, pages 1023–1038; June 20, 1998.

Today's Science of the Sun, Parts 1 and 2. Carolus J. Schrijver and Alan M. Title in *Sky & Telescope,* Vol. 101, No. 2, pages 34–39; February 2001; and No. 3, pages 34–40; March 2001.

Glorious Eclipses: Their Past, Present and Future. Serge Brunier and Jean-Pierre Luminet. Cambridge University Press, 2001.

Probing the Sun's Hot Corona. K.J.H. Phillips and B. N. Dwivedi in *Dynamic Sun.* Edited by B. N. Dwivedi. Cambridge University Press, 2003.

Mercury:

Although one of Earth's nearest neighbors, this

The planet closest to the sun, Mercury is a world of extremes. Of all the objects that condensed from the presolar nebula, it formed at the highest temperatures. The planet's dawn-to-dawn day, equal to 176 Earth-days, is the longest in the solar system, longer even than its own year. When Mercury is at perihelion (the point in its orbit closest to the sun), it moves so swiftly that, from the vantage of someone on the surface, the sun would appear to stop in the sky and go backward—until the planet's rotation catches up and makes the sun appear to go forward again. During daytime, its ground temperature reaches 700 kelvins (more than enough to melt lead); at night, it plunges to a mere 100 kelvins (enough to freeze krypton).

By Robert M. Nelson

Such oddities make Mercury exceptionally intriguing to astronomers. The planet, in fact, poses special challenges to scientific investigation. Its extreme properties make Mercury difficult to fit into any general scheme for the evolution of the solar system. In a sense, its unusual attributes provide an exacting

Updated from the November 1997 issue

DAWN ON MERCURY, 10 times as brilliant as on Earth, is heralded by flares from the sun's corona snaking over the horizon. They light up the slopes of Discovery scarp (*cliffs at right*). In the sky, a blue planet and its moon are visible. (This artist's conception is based on data from the Mariner 10 mission.)

DON DIXON

the forgotten planet

strange world remains, for the most part, unknown

and sensitive test for astronomers' theories. Yet even though Mercury ranks after Mars and Venus as one of Earth's nearest neighbors, distant Pluto is the only planet we know less about. Much about Mercury—its origin and evolution, its puzzling magnetic field, its tenuous atmosphere, its possibly liquid core and its remarkably high density—remains obscure. Mercury shines brightly, but it is so far away that early astronomers could not discern any details of its terrain; they could map only its motion in the sky. As the innermost planet, Mercury (as seen from Earth) never wanders more than 27 degrees from the sun. This angle is less than that made by the hands on a watch at one o'clock. It can thus be observed only during the day, when scattered sunlight makes it difficult to see, or shortly before sunrise and after sunset, with the sun hanging just over the horizon. At dawn or dusk, however, Mercury is very low in the sky, and the light from it must pass through about 10 times as much turbulent air as when it is directly overhead. The best Earth-based telescopes can see only those features on Mercury that are a few hundred kilometers across or wider—a resolution far worse than that for the moon seen with the unaided eye.

Despite these obstacles, terrestrial observation has yielded some interesting results. In 1955 astronomers were able to bounce radar waves off Mercury's surface. By measuring the so-called Doppler shift in the frequency of the reflections, they learned of Mercury's 59-day rotational period. Until then, Mercury had been thought to have an 88-day period, identical to its year, so that one side of the planet always faced the sun. The simple two-to-three ratio between the planet's day and year is striking. Mercury, which initially rotated much faster, probably dissipated energy through tidal flexing and slowed down, becoming locked into this ratio by an obscure process.

Modern space-based observatories, such as the Hubble Space Telescope, are not limited by atmospheric distortion. Unfortunately, the Hubble, like many other sensors in space, cannot point at Mercury, because the rays of the nearby sun might accidentally damage its sensitive optical instruments.

The only other way to investigate Mercury is to send a spacecraft. Only once has a probe made the trip: Mariner 10 flew by in the 1970s as part of a larger mission to explore the inner solar system. Getting the spacecraft there was not trivial. Falling directly into the gravitational potential well of the sun was impossible; the spacecraft had to ricochet around Venus to relinquish gravitational energy and thus slow down for a Mercury encounter. Mariner's orbit around the sun provided three close flybys of Mercury: on March 29, 1974; September 21, 1974; and March 16, 1975. The spacecraft returned images of 40 percent of the planet, showing a heavily cratered surface that, at first glance, appeared similar to that of the moon.

The pictures, sadly, led to the mistaken impression that Mercury differs very little from the moon and just happens to occupy a different region of the solar system. As a result, Mercury has become the neglected planet of the American space program. There have been 38 U.S. missions to the moon, eight to Venus and 17 to Mars. In the next five years, an armada of spacecraft will be in orbit around Venus, Mars, Jupiter and Saturn, returning detailed information about these planets and their environs for many years to come. But Mercury will remain largely unexplored.

The Iron Question

IT WAS THE MARINER MISSION that elevated scientific understanding of Mercury from almost nothing to most of what

THE AUTHOR

ROBERT M. NELSON is senior research scientist at the Jet Propulsion Laboratory in Pasadena, Calif., where he has worked since 1979. Nelson was co-investigator for the Voyager spacecraft's photopolarimeter and is on the science team for the Visual and Infrared Mapping Spectrometer of the Cassini Saturn Orbiter mission. He was also the principal investigator on the Hermes '94 and '96 proposals for a Mercury orbiter and was the project scientist for the Deep Space 1 mission, which flew past Comet Borrelly in 2001. The author expresses his gratitude to the Hermes team members for their enlightening contributions.

VITAL STATISTICS

MERCURY IS THE INNERMOST PLANET and has a highly inclined and eccentric orbit. It rotates about its own axis very slowly, so that one Mercury-day equals 176 Earth-days—longer than its year of 88 Earth-days. Proximity to the sun combined with elongated days gives Mercury the highest daytime temperatures in the solar system.

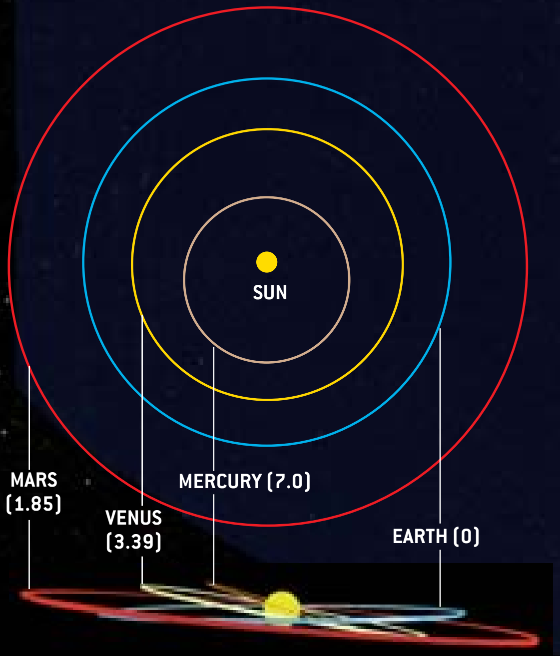
The planet has a rocky and cratered surface and is somewhat larger than the Earth's moon. It is exceptionally dense for its size, implying a large iron core. In addition, it has a strong magnetic field, which suggests that parts of the core are liquid. Because the small planet should have cooled fast enough to have entirely solidified, these findings raise questions about the planet's origins—and even about the birth of the solar system.

Mercury's magnetic field forms a magnetosphere around the planet, which partially shields the surface from the powerful wind of protons emanating from the sun. Its tenuous atmosphere consists of particles recycled from the solar wind or ejected from the surface.

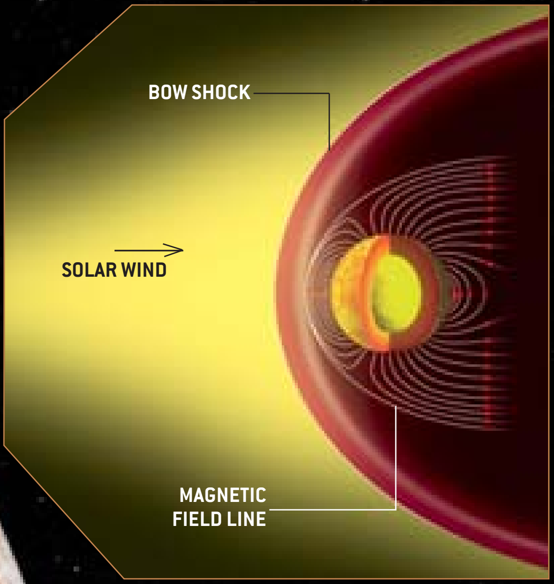
Despite the planet's puzzling nature, only one spacecraft, Mariner 10, has ever flown by Mercury. *—R.M.N.*

RELATIVE SIZES OF TERRESTRIAL BODIES

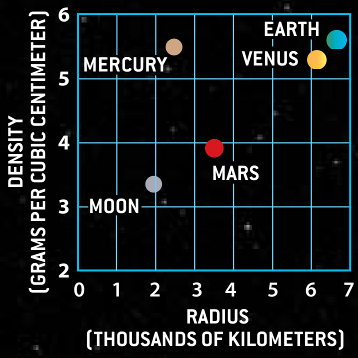
RELATIVE ORBITS OF TERRESTRIAL BODIES (DEGREE OF INCLINATION TO ECLIPTIC)



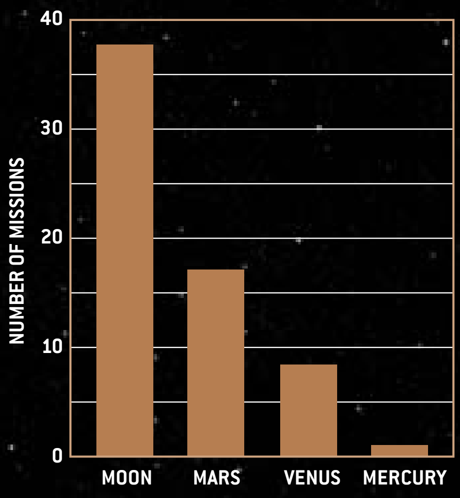
MERCURY'S MAGNETOSPHERE



DENSITY OF TERRESTRIAL BODIES



U.S. MISSIONS TO TERRESTRIAL BODIES

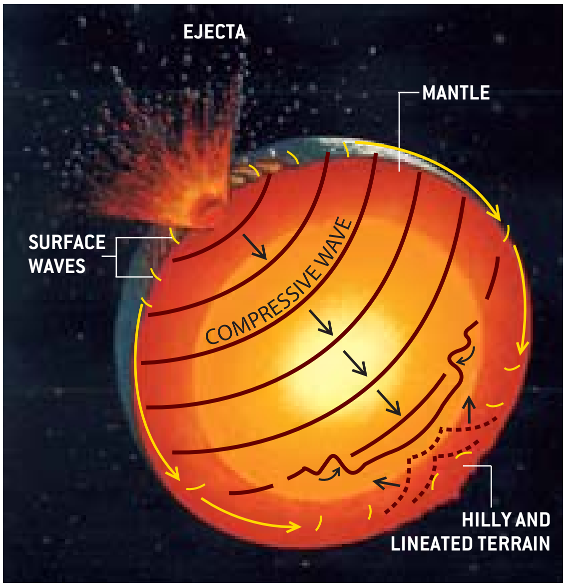


PHOTOGRAPHS BY NASA; ILLUSTRATIONS BY SLIM FILMS

we currently know. The ensemble of instruments carried on that probe sent back about 2,000 images, with an effective resolution of about 1.5 kilometers, comparable to shots of the moon taken from Earth through a large telescope. Yet those many pictures captured only one face of Mercury; the other side has never been seen.

By measuring the acceleration of Mariner in Mercury's surprisingly strong gravitational field, astronomers confirmed one of the planet's most unusual characteristics: its high density. The other terrestrial (that is, nongaseous) bodies—Venus, the moon, Mars and Earth—exhibit a fairly linear relation between density and size. The largest, Earth and Venus, are quite dense, whereas the moon and Mars have lower densities. Mercury is not much bigger than the moon, but its density is typical of a far larger planet, such as Earth.

This observation provides a fundamental clue about Mercury's interior. The outer layers of a terrestrial planet consist of lighter materials such as silicate rocks. With depth, the density increases, because of compression by the overlying rock layers and the different composition of the interior materials. The high-density cores of the terrestrial planets are probably made mostly of iron. This can be inferred because iron is the only element that has both the requisite density and cosmic abundance to sustain the great density of planetary interiors. Other high-density elements are not plentiful enough.



CALORIS CRATER was formed when a giant projectile hit Mercury 3.6 billion years ago (*above*). Shock waves radiated through the planet, creating hilly and lineated terrain on the opposite side. The rim of Caloris itself (*below*) consists of concentric waves that froze in place after the impact. The flattened bed of the crater, 1,300 kilometers across, has since been covered with smaller craters.

Mercury may therefore have the largest metallic core, relative to its size, of all the terrestrial planets. This finding has stimulated a lively debate on the origin and evolution of the solar system. Astronomers assume that all the planets condensed from the solar nebula at about the same time. If this premise is true, then one of three possible circumstances may explain why Mercury is so special. First, the composition of the solar nebula might have been dramatically different in the vicinity of Mercury's orbit—much more so than theoretical models would predict. Or, second, the sun may have been so energetic early in the life of the solar system that the more volatile, low-density elements on Mercury were vaporized and driven off. Or, third, a very massive object might have collided with Mercury soon after the planet's formation, vaporizing the less dense materials. The current body of evidence is not sufficient to discriminate among these possibilities.

Oddly enough, careful analyses of the Mariner findings, along with laborious spectroscopic observations from Earth, have failed to detect even trace amounts of iron in Mercury's crustal rocks. Iron occurs on Earth's crust and has been detected by spectroscopy on the rocks of the moon and Mars. So Mercury may be the only planet in the inner solar system with all its high-density iron concentrated in the interior and only low-density silicates in the crust. It may be that Mercury was

ALFRED T. KAMAJIAN; COURTESY OF P. H. SCHULTZ AND D. E. GAULT (*top*); NASA (*bottom*)

molten for so long that the heavy substances settled at the center, just as iron drops below slag in a smelter.

Mariner 10 also found that Mercury has the most powerful magnetic field of all the terrestrial planets except Earth. The magnetic field of Earth is generated by electrically conductive molten metals circulating in the core, a process called the self-sustaining dynamo. If Mercury's magnetic field has a similar source, then that planet must have a liquid interior.

But there is a problem with this hypothesis. Small objects like Mercury have a high proportion of surface area compared with volume. Therefore, other factors being equal, smaller bodies radiate their energy to space faster. If Mercury has a purely iron core, as its large density and strong magnetic field imply, then the core should have cooled and solidified eons ago. But a solid core cannot support a self-sustaining magnetic dynamo.

This contradiction suggests that other materials are present in the core. These additives may depress the freezing point of iron, so that it remains liquid even at relatively low temperatures. Sulfur, a cosmically abundant element, is a possible candidate. Recent models, in fact, assume Mercury's core to be made of solid iron but surrounded by a liquid shell of iron and sulfur, at 1,300 kelvins. But this solution to the paradox remains a surmise.

Once a planetary surface solidifies sufficiently, it may bend when stress is applied steadily over long periods, or it may crack on sudden impact. After Mercury was born four billion years ago, it was bombarded with huge asteroids that broke through its fragile outer skin and released torrents of lava. More recently, smaller collisions have caused lava to flow. These impacts must have either released enough energy to melt the surface or tapped deeper, liquid layers. Mercury's surface is stamped with events that occurred after its outer layer solidified.

Planetary geologists have tried to sketch Mercury's history using these features—and without accurate knowledge of the surface rocks. The only way to determine absolute age is by radiometric dating of returned samples. But geologists have ingenious ways of assigning relative ages, mostly based on the principle of superposition: any feature that overlies or cuts across another is the younger. This principle is particularly helpful in establishing the relative ages of craters.

A Fractured History

MERCURY HAS SEVERAL large craters that are surrounded by multiple concentric rings of hills and valleys. The rings probably originated when an asteroid hit, causing shock waves to ripple outward like waves from a stone dropped into a pond, and then froze in place. Caloris, a behemoth 1,300 kilometers in diameter, is the largest of these craters. The impact that created it established a flat basin—wiping the slate clean, so to speak—on which a fresh record of smaller impacts has built up. Given an estimate of the rate at which projectiles hit the planet, the size distribution of these craters indicates that the Caloris impact probably occurred around 3.6 billion years ago; it serves as a reference point in time. The collision was so violent that it disrupted the surface on the opposite side of Mercury, where the antipode of Caloris shows many cracks and faults.

ANTIPODE OF CALORIS contains highly chaotic terrain, with hills and fractures that resulted from the impact on the other side of the planet. Petrarch crater (*at center*) was created by a far more recent impact, as evinced by the paucity of smaller craters on its smooth bed. But that collision was violent enough to melt rock, which flowed through a 100-kilometer-long channel and flooded a neighboring crater.

NASA

Mercury's surface is also crosscut by linear features of unknown origin that are preferentially oriented north-south, northeast-southwest and northwest-southeast. These lineaments are called the Mercurian grid. One explanation is that the crust solidified when the planet was rotating much faster, perhaps with a day of only 20 hours. Because of its rapid spin, the planet would have had an equatorial bulge; after it slowed to its present period, gravity pulled it into a more spherical shape. The lineaments may have arisen as the surface accommodated this change. The wrinkles do not cut across the Caloris crater, indicating that they were established before that impact.

While Mercury's rotation was slowing, the planet was also cooling, so that the outer regions of the core solidified. The accompanying shrinkage probably reduced the planet's surface area by about a million square kilometers, producing a network of faults that are evident as a series of curved scarps, or cliffs, crisscrossing Mercury's surface.

Compared with Earth, where erosion has smoothed out most craters, Mercury, Mars and the moon have heavily cratered surfaces. The craters also show a similar distribution of sizes, except that Mercury's tend to be somewhat larger. The objects striking Mercury most likely had higher velocity. Such a pattern is to be expected if the projectiles were in elliptical orbits around the sun: they would have been moving faster in the region of Mercury's orbit than if they were farther out. So these rocks may have been all from the same family, one that probably originated in the asteroid belt. In contrast, the moons of Jupiter have a different distribution of crater sizes, indicating that they collided with a different group of objects.

A Tenuous Atmosphere

MERCURY'S MAGNETIC FIELD is strong enough to trap charged particles, such as those blowing in with the solar wind (a stream of protons ejected from the sun). The magnetic field forms a shield, or magnetosphere, that is a miniaturized version of the one surrounding Earth. Magnetospheres change constantly in response to the sun's activity; Mercury's magnetic shield, because of its smaller size, can change much faster than Earth's. Thus, it responds quickly to the solar wind, which is 10 times as dense at Mercury as at Earth.

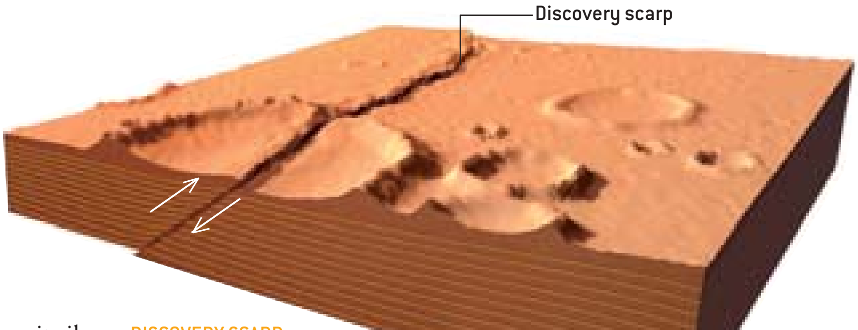
The fierce solar wind steadily bombards Mercury on its illuminated side. The magnetic field is just strong enough to prevent the wind from reaching the planet's surface, except when the sun is very active or when Mercury is at perihelion. At these times, the solar wind reaches all the way down to the surface, and its energetic protons knock material off the crust. The particles ejected during this process can then get trapped by the magnetosphere.

Objects as hot as Mercury do not, however, retain appreciable atmospheres around them, because gas molecules tend to move faster than the escape velocity of the planet. Any significant amount of volatile material on Mercury should soon be lost to space. For this reason, it had long been thought that Mercury did not have an atmosphere. But the ultraviolet spectrometer on Mariner 10 detected small amounts of hydrogen, helium and oxygen, and subsequent Earth-based observations have found traces of sodium and potassium.

The source and ultimate fate of this atmospheric material is a subject of animated argument. Unlike Earth's gaseous cloak, Mercury's atmosphere is constantly evaporating and being replenished. Much of the atmosphere is probably created, directly or indirectly, by the solar wind. Some components may come from the magnetosphere or from the direct infall of cometary material. And once an atom is "sputtered" off the surface by the solar wind, it adds to the tenuous atmosphere. It is even possible that the planet is still outgassing the last remnants of its primordial inventory of volatile substances.

An additional component of Mercury's complex atmosphere-surface dynamics arises from the work of astronomers at Caltech and the Jet Propulsion Laboratory, both in Pasadena, Calif.,who observed the circular polarization of a radar beam reflected from Mercury's polar areas. Those results suggest the presence of water ice. The prospect of a planet as hot as Mercury having ice caps—or any water at all—is intriguing. It may be that the ice resides in permanently shaded regions near Mercury's poles and is left over from primordial water that condensed on the planet when it formed.

If so, Mercury must have stayed in a remarkably stable orientation for the entire age of the solar system, never tipping either pole to the sun—despite devastating events such as the Caloris impact. Such stability would be highly remarkable. Another possible source of water might be the comets that are continually falling into Mercury. Ice landing at a pole may remain in the shade, evaporating very slowly; such water deposits may be a source of Mercury's atmospheric oxygen and hydrogen. On the other hand, astronomers at the University of Arizona have suggested that the shaded polar regions may contain other volatile species such as sulfur, which mimics the radar reflectivity of ice but has a higher melting point.



DISCOVERY SCARP (*crooked line seen in inset above and on opposite page*) stretches for 500 kilometers and in places is two kilometers high. It is a thrust fault, one of many riddling the surface of Mercury. These faults were probably created when parts of Mercury's core solidified and shrank. In consequence, the crust had to squeeze in to cover a smaller area. This compression is achieved when one section of crust slides over another—generating a thrust fault.

NASA; SLIM FILMS

Obstacles to Exploration

WHY HAS MERCURY been left out of the efforts to explore the solar system for more than a quarter century? One possibility, as mentioned, is the superficial similarity between Mercury and the moon. A second is that NASA attaches a high priority to missions that study environments in which life may exist or is believed to have evolved; Mercury is a poor candidate for this. Another, more subtle factor arises from the way planetary missions are devised. The members of peer-review panels for NASA have generally been involved in the agency's most recent missions. The preponderance of missions has been to other planets, so these scientists have developed preferential interests.

Another consideration is economics. NASA's research program has undergone a profound transition since the Apollo days. After the lunar landings, political interest in NASA waned, and its budgets became tight. Nevertheless, robotic missions to explore the solar system continued successfully. Voyager examined the giant planets, and Galileo orbited Jupiter; the Cassini and Huygens probes, which will interrogate the Saturnian system, were launched. Though much less costly than manned spacecraft, robotic missions were still expensive. Each one was in the billion-dollar class, and many encountered cost overruns, often as a result of initial underestimates by industrial suppliers. Therefore, NASA could afford only about one mission a decade. The prospect of a project dedicated to Mercury was bleak.

To address this situation, in the early 1990s NASA inaugurated the Discovery program. In this scheme, scientists with a common interest team with industry and propose a low-cost mission concept with a limited set of high-priority scientific objectives that can be attained with a minimal instrument ensemble. NASA attempts to select a mission every 18 months or so. The awards contain strict cost caps, currently $325 million to $350 million, including the launch vehicle.

A mission to orbit Mercury poses a special technical hurdle. The spacecraft must be protected against the intense energy radiating from the sun and also against the solar energy reflected off Mercury. Because the spacecraft will be close to the planet, at times "Mercury-light" can become a greater threat than the direct sun itself. Despite all the challenges, NASA received one

Discovery mission proposal for a Mercury orbiter in 1994 and two in 1996.

The 1994 proposal, called Hermes ’94, employed a traditional hydrazine–nitrogen tetroxide propulsion system, requiring as much as 1,145 kilograms of propellants. Much of this fuel is needed to slow the spacecraft as it falls toward the sun. The mission’s planners, who included myself, could have reduced the fuel mass only by increasing the number of planetary encounters (to remove gravitational energy). Unfortunately, these maneuvers would have increased the time spent in space, where exposure to radiation limits the lifetime of critical solid-state components.

The instrument complement would have permitted Mercury’s entire surface to be mapped at a resolution of one kilometer or better. These topographic maps could be correlated with charts of Mercury’s magnetic and gravitational fields. NASA initially selected the mission as a candidate for study but ultimately rejected it because of the high cost and risk.

In 1996 the Hermes team, JPL and Spectrum Astro Corp. in Gilbert, Ariz., proposed a new technology that permitted the same payload while slashing the fuel mass, cost and travel time. Their design called for a solar-powered ion-thruster engine, requiring only 295 kilograms of fuel. This revolutionary engine would propel the spacecraft by using the sun’s energy to ionize atoms of xenon and accelerate them to high velocity via an electrical field directed out of the rear of the spacecraft. This innovation would have made the interplanetary cruise time of Hermes ’96 a year shorter than that for Hermes ’94. Yet NASA did not consider Hermes ’96 for further study, because it regarded solar-electric propulsion without full backup from chemical propellant to be too experimental. NASA did subsequently fly a solar-electric-powered craft as a technology validation concept. Deep Space 1 was launched in October 1998 and culminated in a dramatic flyby of Comet Borrelly in September 2001, returning the best close-up images of a comet ever taken.

NASA did actually select one proposal for a Mercury orbiter in the 1996 cycle of Discovery missions. This design, called Messenger, was developed by engineers at the Johns Hopkins Applied Physics Laboratory. It relies on traditional chemical propulsion and has two large devices that can determine the proportions of the most abundant elements of the crustal rocks. The devices’ mass requires that the spacecraft swoop by Venus twice and Mercury three times before it goes into orbit. This trajectory will lengthen the journey to Mercury to more than four years (about twice that of Hermes ’96). Messenger is also the most costly Discovery mission yet attempted. It has pressed its budget cap, and assembly of the vehicle has not been completed. Under the Discovery rules, the only recourse is to reduce the craft’s capability, which would reduce scientific return; the ambitious payload exceeds Discovery’s program limits.

Fortunately, NASA’s Messenger is not the only planned mission to Mercury. The European Space Agency has teamed with the Japanese space agency to develop an ambitious exploration called BepiColombo, to be launched in 2011. It is named after Giuseppe Colombo, an Italian engineer and mathematics professor who in the 1970s made key insights into the complexities of Mercury’s orbital dynamics. The BepiColombo mission comprises three spacecraft delivered by one or two vehicles powered by an ion drive similar to that of Deep Space 1; such systems are no longer considered experimental. The vehicle will take less than 3.5 years to reach a Mercury orbit, so its electronics will be spared excessive exposure to the ravaging deep-space environment.

BepiColombo will have two orbiters and a surface lander, each with a magnetometer and the ability to analyze material immediately around it. One of the orbiters will direct remote sensing instruments at Mercury’s surface; the other provides simultaneous measurements of the planet’s particle and field environments from another location in the magnetosphere. Both spacecraft will be in elliptical orbit and will reach within 400 kilometers at closest approach. One will move out as far as 12,000 kilometers, the other to 1,500 kilometers. The surface lander will touch down on Mercury’s unlit side, where temperature extremes are less, minimizing thermal stress on the instruments. It will have a small camera and gear to measure the chemical composition of surface rocks.

Mercury has presented science with a host of interesting and mysterious questions. The upcoming missions will make the measurements necessary to answer these questions, improving our knowledge of the sun’s nearest neighbor. In learning more about Mercury, we will discover more about our entire solar system, its origin and evolution, and we will be better able to project those evolutionary trends into the future. SA

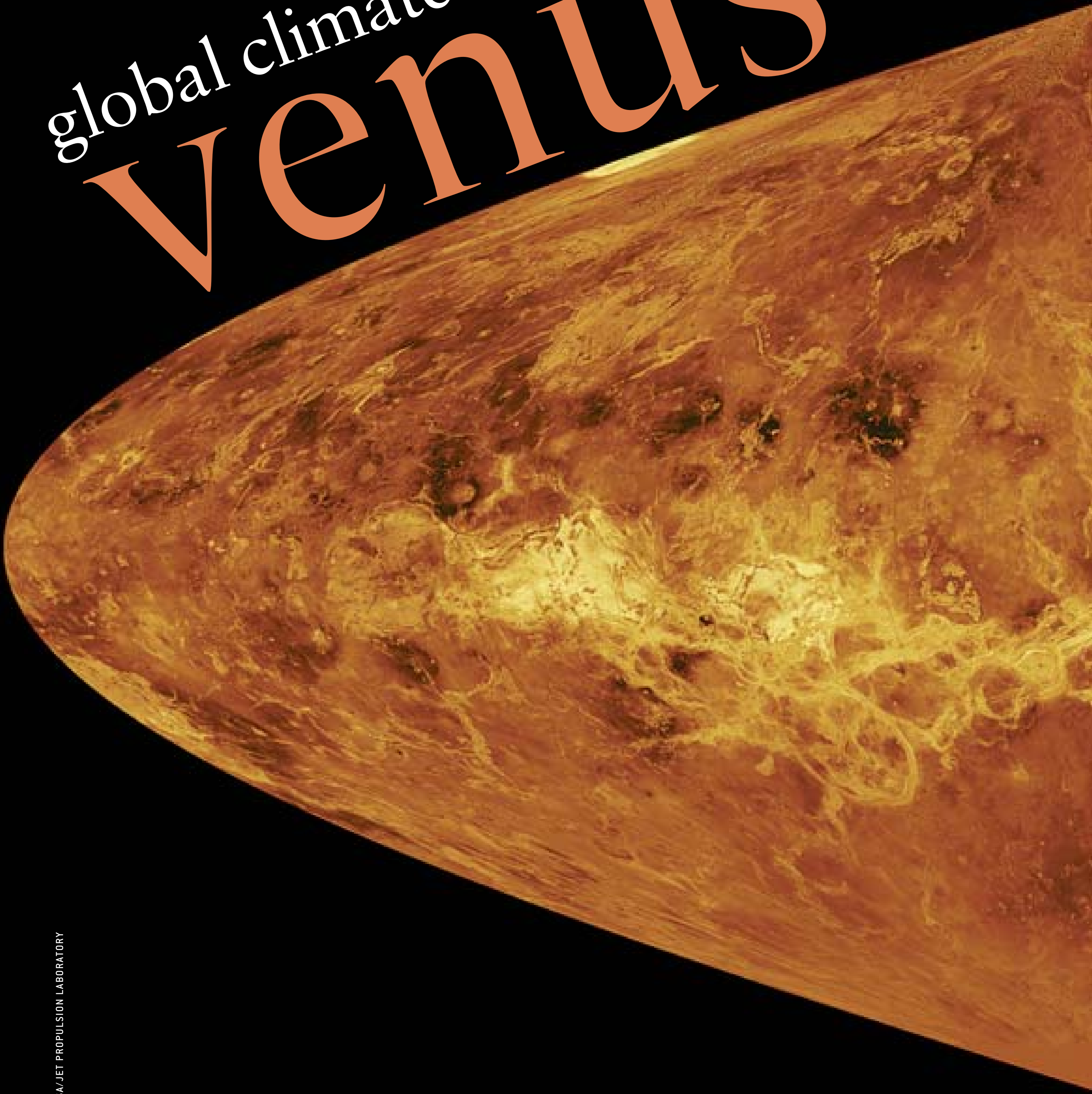
MORE TO EXPLORE

Mercury. Edited by F. Vilas, C. R. Chapman and M. S. Matthews. University of Arizona Press, 1988.

The New Solar System. Edited by J. K. Beatty and A. Chaikin. Cambridge University Press and Sky Publishing Corporation, 1990.

Mercury. Robert M. Nelson in *Encyclopedia of Space Science and Technology*. John Wiley & Sons, 2003.

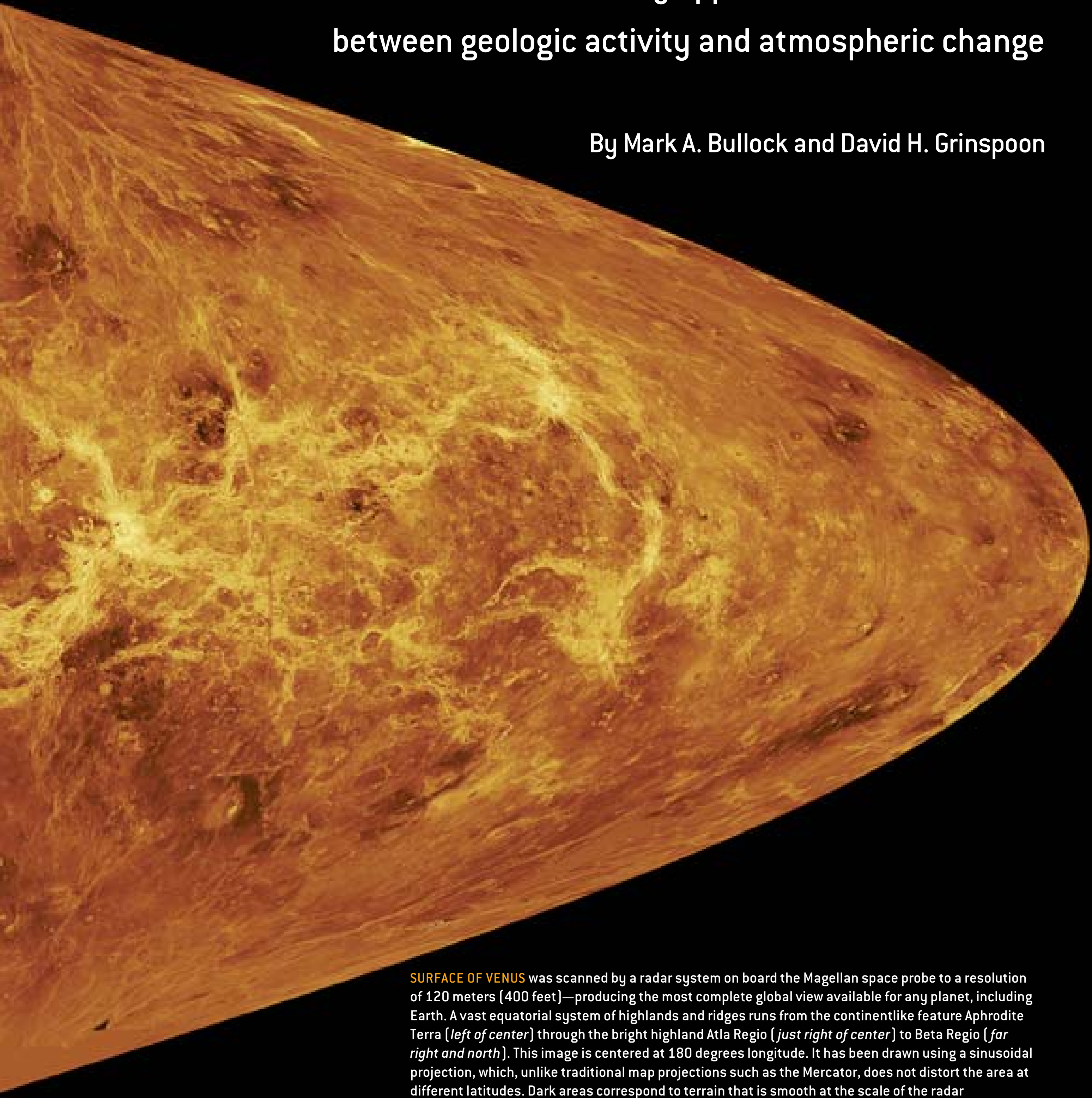
global climate change on venus



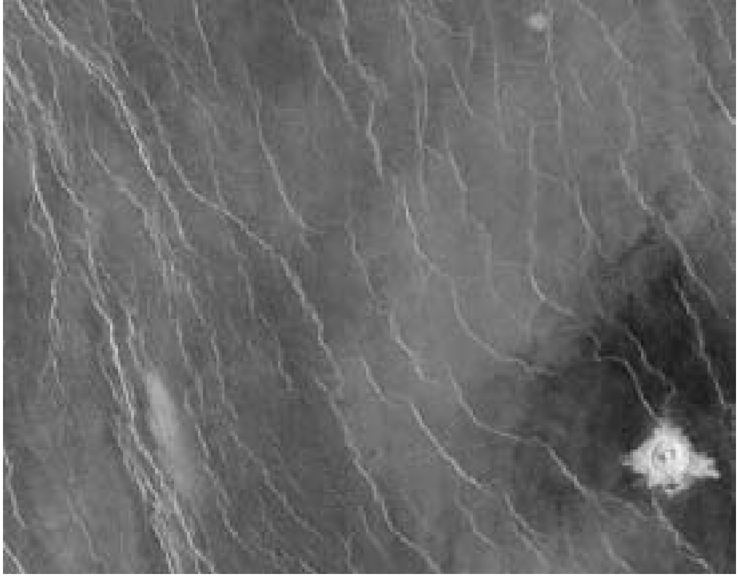
NASA/JET PROPULSION LABORATORY

Venus's climate, like Earth's, has varied over time—the result of newly appreciated connections between geologic activity and atmospheric change

By Mark A. Bullock and David H. Grinspoon



SURFACE OF VENUS was scanned by a radar system on board the Magellan space probe to a resolution of 120 meters (400 feet)—producing the most complete global view available for any planet, including Earth. A vast equatorial system of highlands and ridges runs from the continentlike feature Aphrodite Terra (*left of center*) through the bright highland Atla Regio (*just right of center*) to Beta Regio (*far right and north*). This image is centered at 180 degrees longitude. It has been drawn using a sinusoidal projection, which, unlike traditional map projections such as the Mercator, does not distort the area at different latitudes. Dark areas correspond to terrain that is smooth at the scale of the radar wavelength (13 centimeters); bright areas are rough. The meridional striations are image artifacts.



WRINKLE RIDGES are the most common feature on the volcanic plains of Venus. They are parallel and evenly spaced, suggesting that they formed when the plains as a whole were subjected to stress—perhaps induced by a dramatic, rapid change in surface temperature. This region, which is part of the equatorial plains known as Rusalka Planitia, is approximately 300 kilometers across.

NASA/JPL

Emerging together from the presolar cauldron, Earth and Venus were endowed with nearly the same size and composition. Yet they have developed into radically different worlds. The surface temperature of Earth's sister planet is about 460 degrees Celsius—hot enough for rocks to glow visibly to any unfortunate carbon-based visitors. A deadly efficient greenhouse effect prevails, sustained by an atmosphere whose major constituent, carbon dioxide, is a powerful insulator. Liquid water is nonexistent. The air pressure at the surface is almost 100 times that on Earth; in many ways it is more an ocean than an atmosphere. A mélange of gaseous sulfur compounds, along with what little water vapor there is, provides chemical fodder for the globally encircling clouds of sulfuric acid.

This depiction of hell has been brought to us by an armada of 22 robotic spacecraft that have photographed, scanned, analyzed and landed on Venus over the past four decades. Throughout most of that time, however, Venus's obscuring clouds hindered a full reconnaissance of its surface. Scientists' view of the planet remained static because they knew little of any dynamic processes, such as volcanism or tectonism, that might have occurred there. The Magellan spacecraft changed that perspective. From 1990 to 1994 it mapped the entire surface of the planet at high resolution by peering through the clouds with radar. It revealed a planet that has experienced massive volcanic eruptions in the past and is almost surely active today. Coupled with this probing of Venusian geologic history, detailed computer simulations have attempted to reconstruct the past billion years of the planet's climate history. The intense volcanism, researchers are realizing, has driven large-scale climate change. Like Earth but unlike any other planet astronomers know, Venus has a complex, evolving climate.

Earth's other neighbor, Mars, has also undergone dramatic changes in climate. Its atmosphere today, however, is a relic of its past. The interior of Mars is too cool now for active volcanism, and the surface rests in a deep freeze. Although variations in Mars's orbital and rotational motions can induce climate change there, volcanism will never again participate. Earth and Venus have climates that are driven by the dynamic interplay between geologic and atmospheric processes.

From our human vantage point next door in the solar system, it is sobering to ponder how forces similar to those on Earth have had such a dissimilar outcome on Venus. Studying that planet has broadened research on climate evolution beyond the single example of Earth and given scientists new approaches for answering pressing questions: How unique is Earth's climate? How stable is it? Humankind is engaged in a massive, uncontrolled experiment on the terrestrial climate brought on by the growing effluent from a technological society. Discerning the factors that affect the evolution of climate on other planets is crucial to understanding how natural and anthropogenic forces alter the climate on Earth.

To cite one example, long before the ozone hole became a topic of household discussion, researchers were trying to come to grips with the exotic photochemistry of Venus's upper atmosphere. They found that chlorine reduced the levels of free oxygen above the planet's clouds. The elucidation of this process for Venus eventually shed light on an analogous one for Earth, whereby chlorine from artificial sources destroys ozone in the stratosphere.

Climate and Geology

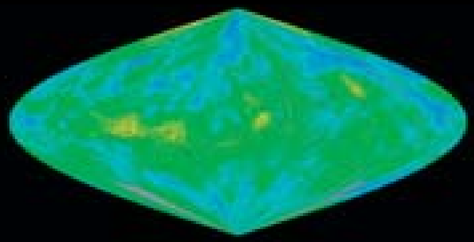
THE CLIMATE OF EARTH is variable partly because its atmosphere is a product of the ongoing shuffling of gases among the crust, the mantle, the oceans, the polar caps and outer space. The driver of geologic processes, geothermal energy, is also an impetus for the evolution of the atmosphere. Geothermal en-

THE AUTHORS

MARK A. BULLOCK and *DAVID H. GRINSPOON* are planetary scientists at the Southwest Research Institute in Boulder, Colo., and have served on national committees that advise NASA on space exploration policy. Bullock began his career studying Mars and now analyzes the evolution of atmospheric conditions on Venus. He co-directs a summer research program for undergraduates on global climate change and society (http://sciencepolicy.colorado.edu/gccs/). Grinspoon studies the evolution of atmospheres and environments on Earth-like planets. His new book, *Lonely Planets: The Natural Philosophy of Alien Life,* will be published in November 2003 by HarperCollins.

Updated from the March 1999 issue

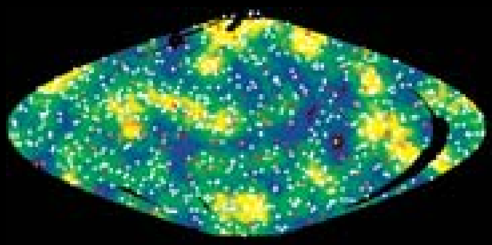
TOPOGRAPHY



The topography of Venus spans a wide range of elevations, about 13 kilometers from low (*blue*) to high (*yellow*). But three fifths of the surface lies within 500 meters of the average elevation, a planetary radius of 6,051.9 kilometers. In contrast, topography on Earth clusters around two distinct elevations, which correspond to continents and ocean floors.

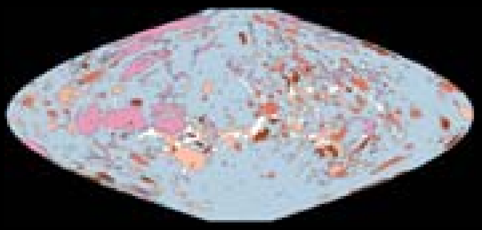
NASA/JPL

IMPACT CRATERS



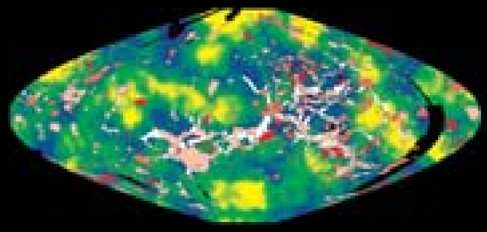
Impact craters are randomly scattered all over Venus. Most are pristine (*white dots*). Those modified by lava (*orange dots*) or by faults (*red triangles*) are concentrated in places such as Aphrodite Terra. Areas with a low density of craters (*blue background*) are often located in highlands. Higher crater densities (*yellow background*) are usually found in the lowland plains.

TYPES OF TERRAIN



The terrain of Venus consists predominately of volcanic plains (*gray*). Within the plains are deformed areas such as tesserae (*pink*) and rift zones (*white*), as well as volcanic features such as coronae (*peach*), lava floods (*red*) and volcanoes of various sizes (*orange*). Volcanoes are not concentrated in chains as they are on Earth, indicating that plate tectonics does not operate.

AGES OF TERRAIN



This geologic map shows the different terrains and their relative ages, as inferred from the crater density. Volcanoes and coronae tend to clump along equatorial rift zones, which are younger (*blue*) than the rest of the Venusian surface. The tesserae, ridges and plains are older (*yellow*). In general, however, the surface lacks the extreme variation in age that is found on Earth and Mars.

MARIBETH PRICE *South Dakota School of Mines and Technology (bottom three images)*

RIVER ON VENUS? This delta exists at the terminus of a narrow channel that runs for 800 kilometers through the northern volcanic plains. Water could not have carved it; Venus is too hot and dry. Instead it was probably the work of lavas rich in carbonate and sulfate salts—which implies that the average temperature used to be several tens of degrees higher than it is today. The region shown here is approximately 40 by 90 kilometers.

NASA/JPL

ergy is primarily a result of the decay of radioactive elements in the interior, and a central problem in studying solid planets is understanding how they lose their heat. Two mechanisms are chiefly responsible: volcanism and plate tectonics.

The interior of Earth cools mainly by means of its plate-tectonic conveyor-belt system, whose steady recycling of gases has exerted a stabilizing force on Earth's climate [*see box on page 26*]. Whereas volcanoes pump gases into the atmosphere, the subduction of lithospheric plates returns them to the interior. Most volcanoes are associated with plate tectonic activity, but some of the largest volcanic edifices on Earth (such as the Hawaiian Islands) have developed as "hot spots" independent of plate boundaries. Historically, the formation of immense volcanic provinces—regions of intense eruptions possibly caused by enormous buoyant plumes of magma within the underlying mantle—may have spewed large amounts of gases and led to periods of global warming.

What about Venus? Plate tectonics is not in evidence, except possibly on a limited scale. It appears that heat was transferred, at least in the relatively recent past, by the eruption of vast plains of basaltic lava and later by the volcanoes that grew on top of them. Understanding the effects of volcanoes is the starting point for any discussion of climate.

A striking feature of Magellan's global survey is the paucity of impact craters. Although Venus's thick atmosphere can stop meteoroids smaller than a kilometer in diameter, which would otherwise gouge craters up to 15 kilometers (nine miles) across, there is a shortage of larger craters as well. Observations of the number of asteroids and comets in the inner solar system, as well as crater counts on the moon, give a rough idea of how quickly Venus should have collected impact scars: about 1.2 craters per million years. Magellan saw only, by the latest count, 963 craters spread randomly over its surface. Somehow impacts from the first 3.7 billion years of the planet's history have been eradicated.

A sparsity of craters is also evident on Earth, where old craters are eroded by wind and water. Terrestrial impact sites are found in a wide range of altered states, from the nearly pristine bowl of Meteor Crater in Arizona to the barely discernible outlines of buried Precambrian impacts in the oldest continental crust. Yet the surface of Venus is far too hot for liquid water to exist, and surface winds are mild. In the absence of erosion, the chief processes altering craters should be volcanic and tectonic activity. That is the paradox. Most of the Venusian craters look fresh: only 6 percent of them have lava lapping their rims, and only 12 percent have been disrupted by folding and cracking of the crust. So where did all the old ones go, if most of those that remain are unaltered? If they have been covered up by lava, why do we not see more craters that are partially covered? And how have they been removed so that their initial random placement has been preserved?

To some researchers, the random distribution of the observed craters and the small number of partially modified ones imply that a geologic event of global proportions abruptly wiped out all the old craters some 800 million years ago. In this scenario, proposed in 1992 by Gerald G. Schaber of the U.S. Geological Survey and Robert G. Strom of the University of Arizona, impacts have peppered the newly formed surface ever since.

But the idea of paving over an entire planet is unpalatable to many geologists. It has no real analogue on Earth. Roger J. Phillips of Washington University proposed an alternative model the same year, known as equilibrium resurfacing, which hypothesized that steady geologic processes continually eradicate craters in small patches, preserving an overall global distribution that appears random. But some geologic features on Venus are immense, suggesting that geologic activity would not wipe craters out cleanly and randomly everywhere.

These two views grew into a classic scientific debate as the analysis of Magellan data became more sophisticated. The truth is probably somewhere in the middle. Elements of both models have been incorporated into the prevailing interpretation of the past billion years of Venus's geologic history: globally extensive volcanism wiped out most impact craters and created the vast volcanic plains 800 million years ago, and it has been followed by a reduced level of continued volcanic activity.

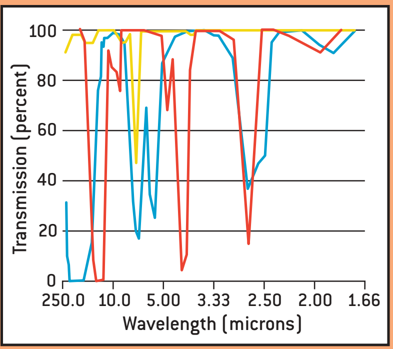
Chocolate-Covered Caramel Crust

ALTHOUGH THERE IS NO DOUBT that volcanism has shaped Venus's surface, the interpretation of some enigmatic geologic features has until recently resisted integration into a coherent picture of the planet's evolution. Some of these features hint that the planet's climate may have changed drastically.

First, several striking lineaments resemble water-carved landforms. Up to 7,000 kilometers long, they are similar to meandering rivers and floodplains on Earth. Many end in outflow channels that look like river deltas. The extreme dryness of the environment makes it highly unlikely that water carved these features. So what did? Perhaps calcium carbonate, calcium sulfate and other salts are the culprit. The surface, which is in equi-

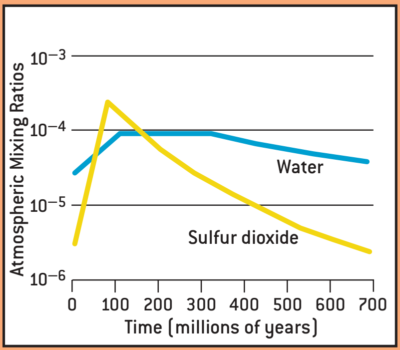
GREENHOUSE EFFECT

Greenhouse gases let sunlight reach the Venusian surface but block outgoing infrared light. Carbon dioxide (*red*), water (*blue*) and sulfur dioxide (*yellow*) each absorb certain wavelengths. Were it not for these gases, the sunlight and infrared light would balance at a surface temperature of –20 degrees Celsius.



GAS CONCENTRATIONS

Water and sulfur dioxide are removed from the atmosphere after they are belched out by volcanic activity. Sulfur dioxide (*yellow*) reacts relatively quickly with carbonates at the surface, whereas water (*blue*) is slowly broken apart by solar ultraviolet radiation.



MARK A. BULLOCK AND DAVID H. GRINSPOON

librium with a hefty carbon dioxide atmosphere laced with sulfur gases, should be replete with these substances. Indeed, the Soviet Venera landers found that surface rocks are about 7 to 10 percent calcium minerals (almost certainly carbonates) and 1 to 5 percent sulfates.

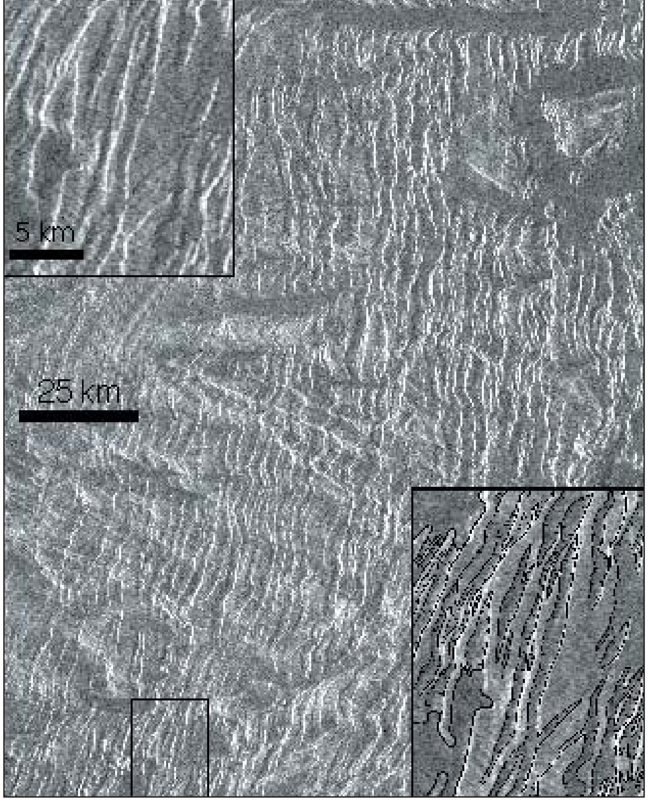
Lavas laden with these salts melt at temperatures of a few tens to hundreds of degrees higher than Venusian surface temperatures today. Jeffrey S. Kargel of the USGS and his co-workers have hypothesized that vast reservoirs of molten carbonatite (salt-rich) magma, analogous to water aquifers on Earth, may exist a few hundred meters to several kilometers under the surface. Moderately higher surface temperatures in the past could have spilled salt-rich fluid lavas onto the surface, where they were stable enough to carve the features we see today.

Second, the mysterious tesserae—the oldest terrain on Venus—also hint at higher temperatures in the past. These intensely crinkled landscapes are located on continentlike crustal plateaus that rise several kilometers above the lowland lava plains. Analyses by Phillips and by Vicki L. Hansen of Southern Methodist University indicate that the plateaus were formed by extension of the lithosphere (the rigid exoskeleton of the planet, consisting of the crust and upper mantle). The process was something like stretching apart a chocolate-covered caramel that is gooey on the inside with a thin, brittle shell. Today the outer, brittle part of the lithosphere is too thick to behave this way. At the time of tessera formation, it must have been thinner, which implies that the surface was much hotter.

Finally, cracks and folds crisscross the planet. At least some of these patterns, particularly the so-called wrinkle ridges, may be related to temporal variations in climate. We and Sean C. Solomon of the Carnegie Institution of Washington have argued that the plains preserve global episodes of deformation that may have occurred over short geologic intervals. That is, the entire lithosphere seems to have been stretched or compressed all at the same time. It is hard to imagine a mechanism internal to the solid planet that could do that. But what about global climate change? Solomon calculated that stresses induced in the lithosphere by fluctuations in surface temperature of about 100 degrees C would have been as high as 1,000 bars—comparable to those that form mountain belts on Earth and sufficient to deform Venus's surface in the observed way.

Around the time that the debate over Venus's recent geologic history was raging, we were working on a detailed model of its atmosphere. Theory reveals that the alien and hostile conditions are maintained by the complementary properties of Venus's atmospheric constituents. Water vapor, even in trace amounts, absorbs infrared radiation at wavelengths that carbon dioxide does not. Sulfur dioxide and other sulfur gases block still other infrared wavelengths. Together these greenhouse gases conspire to make the atmosphere of Venus partially transparent to incoming solar radiation but nearly completely opaque to outgoing thermal radiation. Consequently, the surface temperature is three times what it would be without an atmosphere. On Earth, by comparison, the greenhouse effect currently boosts the surface temperature by only about 15 percent.

If volcanoes really did repave the Venusian surface 800 million years ago, they should have also injected a great deal of greenhouse gases into the atmosphere in a relatively short time. A reasonable estimate is that enough lava erupted to cover the planet with a layer one to 10 kilometers thick. In that case, the amount of carbon dioxide in the atmosphere would have hardly changed—there is already so much of it. But the abundances of water vapor and sulfur dioxide would have increased 10- and 100-fold, respectively. Fascinated by the possible implications,

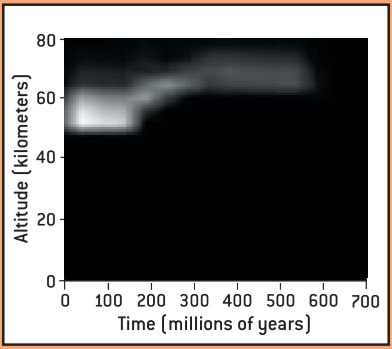


VICKI L. HANSEN *Southern Methodist University* AND ROGER J. PHILLIPS *Washington University*

RIBBON TERRAIN consists of steep-sided, flat-bottomed, shallow (400-meter) troughs. These features may have resulted from fracturing of a thin, brittle layer of rock above a weaker, ductile substrate. The insets magnify the region in the box; troughs are marked on the bottom right.

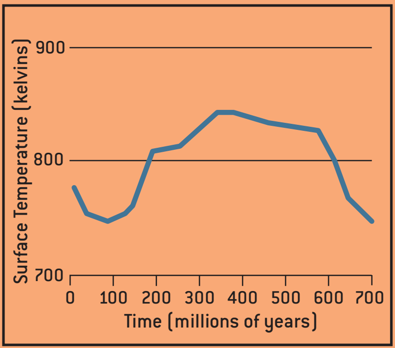
CLOUD COVER

The sulfuric acid clouds vary in thickness after a global series of volcanic eruptions. The clouds first thicken as water and sulfur dioxide pour into the air. Then they dissipate as these gases thin out. About 400 million years after the onset of volcanism, the acidic clouds are replaced by thin, high water clouds.



TEMPERATURE

The surface temperature depends on the relative importance of clouds and the greenhouse effect. Initially volcanism produces thick clouds that cool the surface. But because water is lost more slowly from the planet's atmosphere than sulfur dioxide is, a greenhouse effect subsequently warms the surface.



MARK A. BULLOCK AND DAVID H. GRINSPOON

we modeled the planet's climate as an interconnected system of processes, including volcanic outgassing, cloud formation, the loss of hydrogen from the top of the atmosphere, and reactions of atmospheric gases with surface minerals.

The interaction of these processes can be subtle. Although carbon dioxide, water vapor and sulfur dioxide all warm the surface, the last two also have a countervailing effect: the production of clouds. Higher concentrations of water vapor and sulfur dioxide would not only enhance the effect but also thicken the clouds, which reflect sunlight back into space and thereby cool the planet. Because of these competing effects, it is not obvious what the injection of the two gases did to the climate.

The Planetary Perspective

OUR SIMULATIONS SUGGEST that the clouds initially won out, so that the surface cooled by about 100 degrees C. But then the clouds were slowly eaten away. Water diffused higher in the atmosphere, where it was dissociated by solar ultraviolet radiation. The hydrogen slowly escaped into space; half of it was lost within 200 million years. The sulfur dioxide, meanwhile, was taken up in carbonate rocks.

As the clouds thinned, more solar energy reached the surface, heating it. After 200 million years or so, temperatures were high enough to start evaporating the clouds from below. A positive feedback ensued: the more the clouds eroded, the less sunlight was reflected, the hotter the surface became, the more the clouds were evaporated from below, and so on. The magnificent cloud decks rapidly disappeared. For about 400 million years, all that remained of them was a wispy, high stretch of clouds composed mostly of water. Surface temperatures were 100 degrees C higher than at present, because the atmospheric abundance of water vapor was still fairly high and because the thin clouds contributed to the greenhouse effect without reflecting much solar energy. Eventually, 600 million years after the onset of volcanism but in the absence of any further volcanic activity, the clouds would have dissipated completely.

Because sulfur dioxide and water vapor are continuously lost, clouds require ongoing volcanism for their maintenance. We calculated that volcanism must have been active within the past 30 million years to support the thick clouds observed today. The interior processes that generate surface volcanism occur over many tens of millions of years, so volcanoes are prob-

WHY IS VENUS A HELLHOLE?

THE STUNNING DIFFERENCES between the climates of Earth and Venus today are intimately linked to the history of water on these two worlds. Liquid water is the intermediary in reactions of carbon dioxide and surface rocks that can form minerals. In addition, water mixed into the underlying mantle is probably responsible for the low-viscosity layer, or asthenosphere, on which Earth's lithospheric plates slide. The formation of carbonate minerals and their subsequent descent on tectonic plates prevent carbon dioxide from building up.

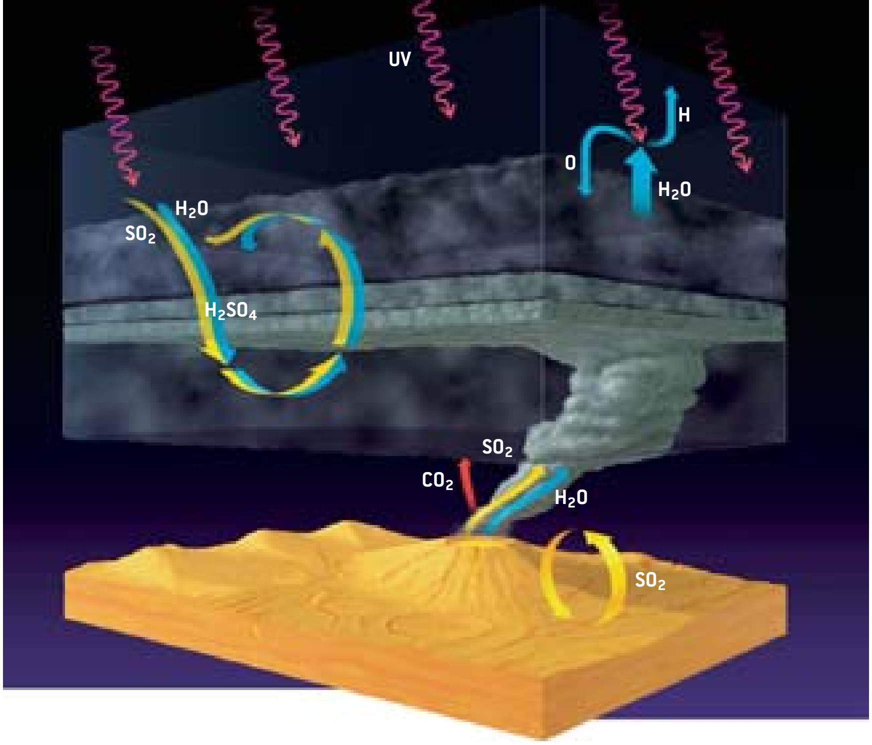
Models of planet formation predict that the two worlds should have been endowed with roughly equal amounts of water, delivered by the impact of icy bodies from the outer solar system. But, when the Pioneer Venus mission went into orbit in 1978, it measured the ratio of deuterium to ordinary hydrogen within the water of Venus's clouds. The ratio was an astonishing 150 times the terrestrial value. The most likely explanation is that Venus once had far more water and lost it. When water vapor drifted into the upper atmosphere, solar ultraviolet radiation decomposed it into oxygen and either hydrogen or deuterium. Because hydrogen, being lighter, escapes to space more easily, the relative amount of deuterium increased.

Why did this process occur on Venus but not on Earth? In 1969 Andrew P. Ingersoll of the California Institute of Technology showed that if the solar energy available to a planet were strong enough, any water at the surface would rapidly evaporate. The added water vapor would further heat the atmosphere and set up what he called the runaway greenhouse effect. The process would transport the bulk of the planet's water into the upper atmosphere, where it would ultimately be decomposed and lost. Later James F. Kasting of Pennsylvania State University and his co-workers developed a more detailed model of this effect. They estimated that the critical solar flux required to initiate a runaway greenhouse was about 40 percent larger than the present flux on Earth. This value corresponds roughly to the solar flux expected at the orbit of Venus shortly after it was formed, when the sun was 30 percent fainter. An Earth ocean's worth of water could have fled Venus in the first 30 million years of its existence.

A shortcoming of this model is that if Venus had a thick carbon dioxide atmosphere early on, as it does now, it would have retained much of its water. The amount of water that is lost depends on how much of it can rise high enough to be decomposed—which is less for a planet with a thick atmosphere. Furthermore, any clouds that developed during the process would have reflected sunlight back into space and shut off the runaway greenhouse.

So Kasting's group also considered a solar flux slightly below the critical value. In this scenario, Venus had hot oceans and a humid stratosphere. The seas kept levels of carbon dioxide low by dissolving the gas and promoting carbonate formation. With lubrication from water in the asthenosphere, plate tectonics might have operated. In short, Venus possessed climate-stabilizing mechanisms similar to those on Earth today. But the atmosphere's lower density could not prevent water from diffusing to high altitudes. Over 600 million years, an ocean's worth of water vanished. Any plate tectonics shut down, leaving volcanism and heat conduction as the interior's ways to cool. Thereafter carbon dioxide accumulated in the air.

This picture, termed the moist greenhouse, illustrates the intricate interaction of solar, climate and geologic change. Atmospheric and surface processes can preserve the status quo, or they can conspire in their own destruction. If the theory is right, Venus once had oceans—perhaps even life, although it may be impossible to know. *—M.A.B. and D.H.G.*



ATMOSPHERE OF VENUS suffers from ovenlike temperatures, oceanic pressures and sulfuric acid clouds (H_2SO_4) because Venus lacks the cycles that stabilize conditions on Earth. Its atmospheric processes are one-way. Carbon dioxide (CO_2), once injected by volcanoes, stays in the atmosphere; water (H_2O), once destroyed by ultraviolet light, is lost forever to the depths of space; sulfur dioxide (SO_2), once locked up in minerals, piles up on the surface (although a small amount does recycle).

TOM MOORE; SOURCE: MARK A. BULLOCK AND DAVID H. GRINSPOON; BASED ON DIAGRAM BY CARTER EMMART *Hayden Planetarium*

ably still active. This finding accords with observations of varying amounts of sulfur dioxide on Venus. Surface temperature fluctuations, precipitated by volcanism, are also a natural explanation for many of the enigmatic features found by Magellan.

Fortunately, Earth's climate has not experienced quite the same extremes in the geologically recent past. Although it is also affected by volcanism, the oxygen-rich atmosphere—provided by biota and plentiful water—readily removes sulfur gases. Therefore, water clouds are key to the planet's heat balance. The amount of water vapor available to these clouds is determined by the evaporation of the oceans, which in turn depends on surface temperature. A slightly enhanced greenhouse effect on Earth puts more water into the atmosphere and results in more cloud cover. The higher reflectivity reduces the incoming solar energy and hence the temperature. This negative feedback acts as a thermostat, keeping the surface temperature moderate over short intervals (days to years). An analogous feedback, the carbonate-silicate cycle, also stabilizes the abundance of atmospheric carbon dioxide. Governed by the slow process of plate tectonics, this mechanism operates over timescales of about half a million years.

These remarkable cycles, intertwined with water and life, have saved Earth's climate from the wild excursions its sister planet has endured. Anthropogenic influences, however, operate on intermediate timescales. The abundance of carbon dioxide in Earth's atmosphere has risen by a quarter since 1860. Although nearly all researchers agree that global warming is occurring, debate continues on how much of it is caused by the burning of fossil fuels and how much stems from natural variations. Whether there is a critical amount of carbon dioxide that overwhelms Earth's climate regulation cycles is not known. But one thing is certain: the climates of Earth-like planets can undergo abrupt transitions because of interactions among planetary-scale processes. In the long run, Earth's fate is sealed. As the sun ages, it will brighten. In about a billion years, the oceans will begin to evaporate rapidly and the climate will succumb to a runaway greenhouse. Earth and Venus, having started as nearly identical twins and diverged, may one day look alike.

Many of us recall the utopian view that science and technology promised in the 1960s. Earth's capacity to supply materials and absorb refuse seemed limitless. For all the immense change that science has wrought since, one of the most powerful is the acquired sense of Earth as a generous but finite home. That perspective has been gained from the growing awareness that by-products from a global technological society have the power to alter the planetary climate. Studying the dynamics of Venus, however alien the planet may seem, is essential to the quest for the general principles of climate variation—and thus to understanding the frailty or robustness of our home world.

The potential for learning more about Earth from Venus has touched off a new era of exploration. The European Space Agency plans to launch the Venus Express spacecraft in November 2005, with arrival in 2006. The orbiter will explore the atmosphere for two Venusian days (486 Earth days). Cameras and spectrometers will give scientists an unprecedented view, from the overheated surface to the thin plasma far above the clouds. The questions of whether there is active volcanism on Venus today, what drives the striking atmospheric rotation, and what mysterious atmospheric constituent is responsible for absorbing most of the sunlight will be addressed by Europe's first Venusian explorer.

In the meantime, Japan's Institute of Space and Aeronautical Science plans to launch a Venus climate orbiter in February 2007. The robot probe's main purpose will be to survey the clouds at various depths to better understand the general atmospheric circulation and superrotation.

Finally, an independent National Academy of Sciences review board recommended to NASA in its report on solar system exploration priorities that a U.S. lander on Venus should be high on the list. Such a mission would have to overcome the extremely harsh conditions on the Venusian surface, as did the Soviet landers of the 1970s and 1980s, and return useful data on the rocks and atmosphere of this incredible world. If selected, this ambitious craft could launch as early as 2009. SA

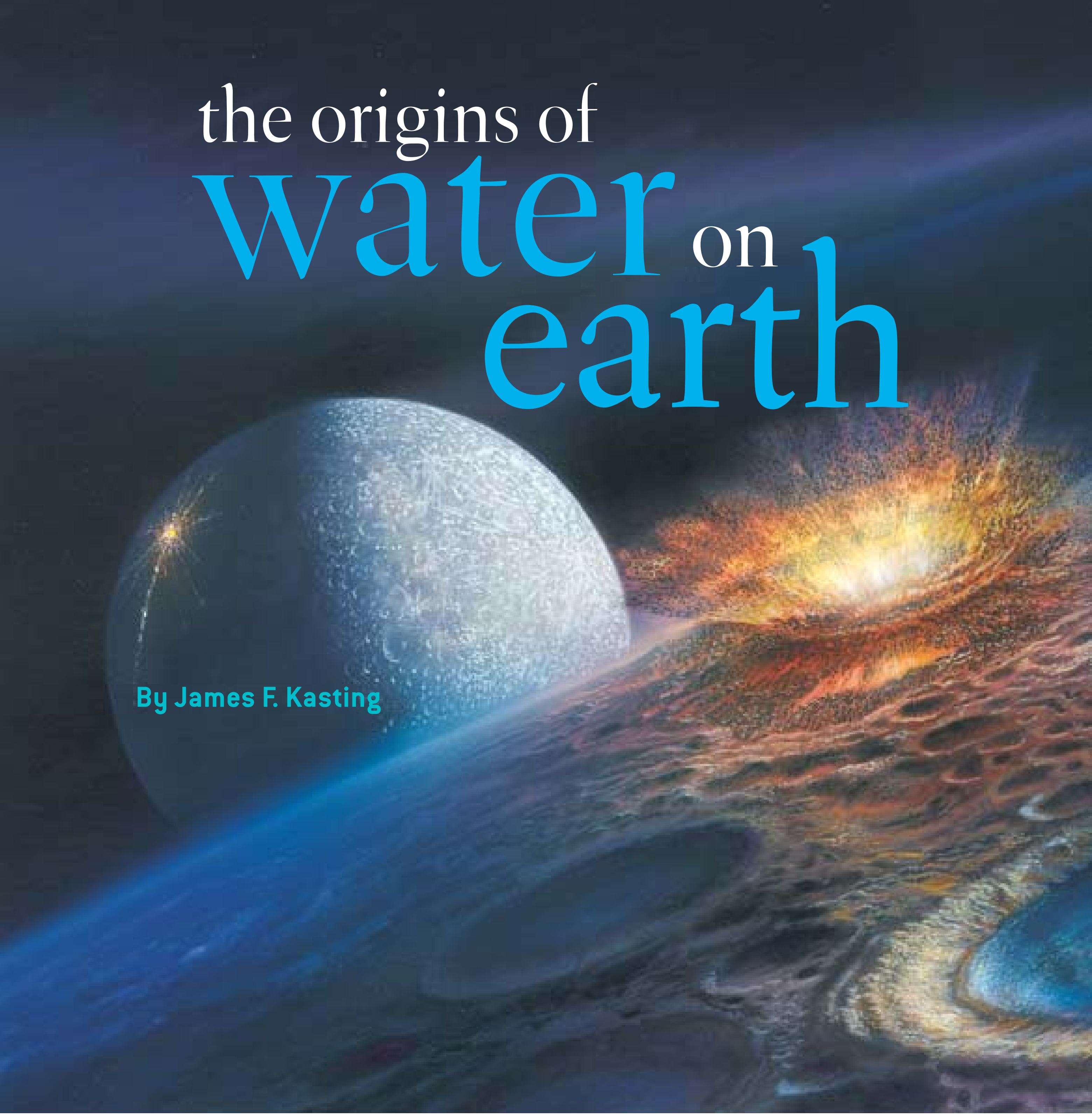
MORE TO EXPLORE

The Stability of Climate on Venus. Mark A. Bullock and David H. Grinspoon in *Journal of Geophysical Research,* Vol. 101, No. E3, pages 7521–7530; March 1996.

Venus Revealed: A New Look below the Clouds of Our Mysterious Twin Planet. David H. Grinspoon. Perseus Books, 1997.

the origins of water on earth

By James F. Kasting



ICE-LADEN COMET crashes into a primitive Earth, which is accumulating its secondary atmosphere (the original having been lost in the catastrophic impact that formed the moon). Earth appears moonlike, but its higher gravity allows it to retain most of the water vapor liberated by such impacts, unlike the newly formed moon in the background. A cooler sun illuminates three additional comets hurtling toward Earth, where they will also give up their water to the planet's steamy, nascent seas.

DON DIXON

Of all the planets, why is Earth the only one fit for life? Simple: because Earth has a surface that supports liquid water, the magic elixir required by all living things. Some scientists speculate that forms of life that do not require water might exist elsewhere in the universe. But I would guess not. The long molecular chains and complex branching structures of carbon make this element the ideal chemical backbone for life, and water is the ideal solvent in which carbon-based chemistry can proceed.

Updated from the Fall 1998 issue of Scientific American Presents

Evidence is mounting that other planets hosted oceans at one time, but **ONLY EARTH** has maintained its watery endowment

Given this special connection between water and life, many investigators have lately focused their attention on one of Jupiter's moons, Europa. Astronomers believe this small world may possess an ocean of liquid water underneath its globe-encircling crust of ice. NASA researchers are making plans to measure the thickness of ice on Europa using radar and, eventually, to drill through that layer should it prove thin enough.

The environment of Europa differs dramatically from conditions on Earth, so there is no reason to suppose that life must have evolved there. But the very existence of water on Europa provides sufficient motivation for sending a spacecraft to search for extraterrestrial organisms. Even if that probing finds nothing alive, the effort may help answer a question closer to home: Where did water on Earth come from?

Water from Heaven

CREATION OF THE MODERN OCEANS required two obvious ingredients: water and a container in which to hold it. The ocean basins owe their origins, as well as their present configuration, to plate tectonics. This heat-driven convection churns

BARRAGE OF COMETS nears an end as a late-arriving body hits at the horizon, sending shocks through the planet and stirring up this primordial sea.

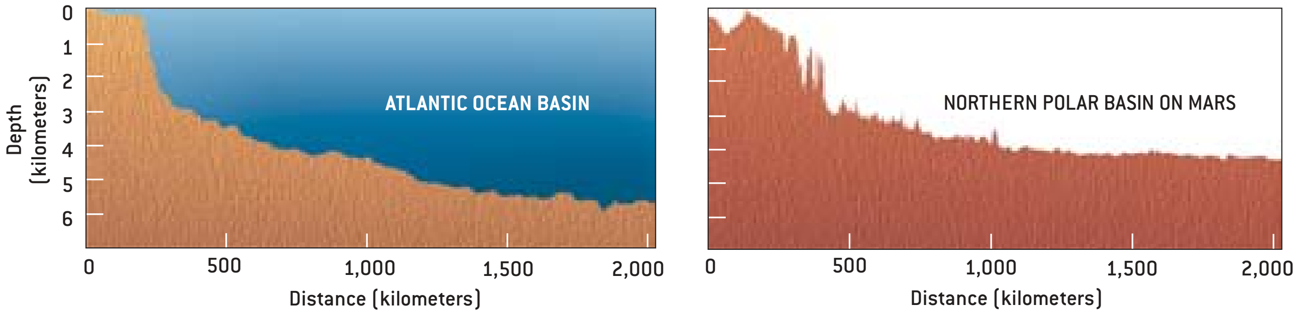
the mantle of Earth—the region between the crust and core—and results in the separation of two kinds of material near the surface. Lighter, less dense granitic rock makes up the continents, which float like sponges in the bath over denser, heavier basalt, which forms the ocean basins.

Scientists cannot determine with certainty exactly when these depressions filled or from where the water came, because there is no geologic record of the formative years of Earth. Dating of meteorites shows that the solar system is about 4.6 billion years old, and Earth appears to be approximately the same age. The oldest sedimentary rocks—those that formed by processes requiring liquid water—are only about 3.9 billion years old. But there are crystals of zirconium silicate, called zircons, that formed 4.4 billion years ago and whose oxygen isotopic composition indicates that liquid water was present then. So water has been on Earth's surface throughout most of its history.

Kevin J. Zahnle, an astronomer at the NASA Ames Research Center, suggests that the primordial Earth was like a bucket. In his view, water was added not with a ladle but with a firehose. He proposes that icy clumps of material collided with Earth during the initial formation of the planet, injecting huge quantities of water into the atmosphere in the form of steam.

Much of this water streamed skyward through holes in the atmosphere blasted open by these icy planetesimals themselves. Many of the water molecules (H_2O) were split apart by ultraviolet radiation from the sun. But enough of the initial steam in the atmosphere survived and condensed to form sizable oceans when the planet eventually cooled.

No one knows how much water rained down on the planet at the time. But suppose the bombarding planetesimals resembled the most abundant type of meteorite (called ordinary chondrite), which contains about 0.1 percent water by weight. An Earth composed entirely of this kind of rubble would therefore have started with 0.1 percent water—at least four times the amount now held in the oceans. So three quarters of this wa-



TOPOGRAPHIC MAPPING of Mars has recently revealed remarkable similarities to the ocean basins on Earth. For example, the western Atlantic near Rio de Janeiro (*left*) presents a similar profile to that of the northern polar basin on Mars (*right*).

DON DIXON (*top*); DAVID SCHNEIDER (*bottom*)

ter has since disappeared. Perhaps half an ocean of the moisture became trapped within minerals of the mantle. Water may also have taken up residence in Earth's dense iron core, which contains some relatively light elements, including, most probably, hydrogen.

The initial influx of meteoric material probably endowed Earth with more than enough water for the oceans. Indeed, that bombardment lasted a long time: from 4.5 billion to 3.8 billion years ago, a time called, naturally enough, the heavy bombardment period.

Where these hefty bodies came from is still a mystery. They may have originated in the asteroid belt, which is located between the orbits of Mars and Jupiter. The rocky masses in the outer parts of the belt may hold up to 20 percent water. Alternatively, if the late-arriving bodies came from beyond the orbit of Jupiter, they would have resembled another water-bearing candidate—comets.

Comets are often described as dirty cosmic snowballs: half ice, half dust. Christopher F. Chyba, a planetary scientist at the University of Arizona, estimates that if only 25 percent of the bodies that hit Earth during the heavy bombardment period were comets, they could have accounted for all the water in the modern oceans. This theory is attractive because it could explain the extended period of heavy bombardment: bodies originating in the Uranus-Neptune region would have taken longer to be swept up by planets, so the volley of impacts on Earth would have stretched over hundreds of millions of years.

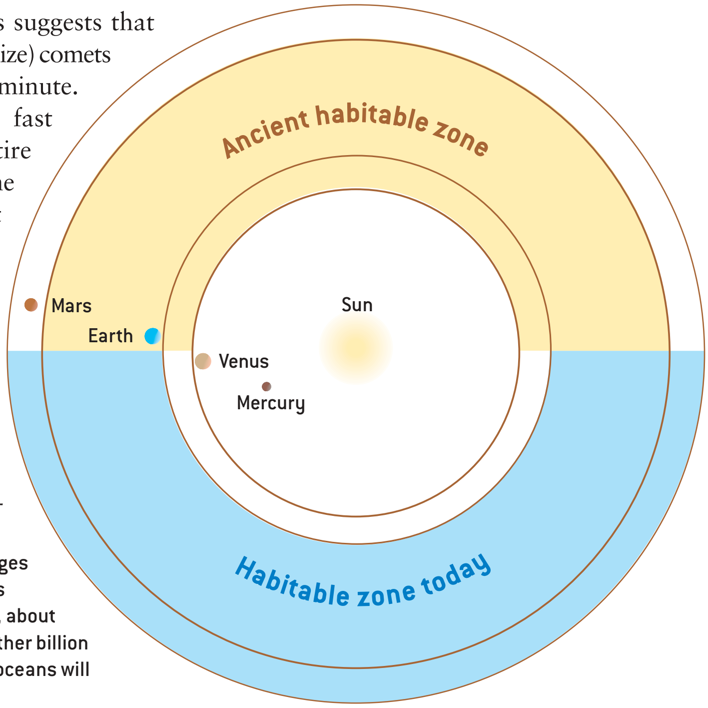
Alternatively, the impactors may have come from the asteroid belt region between 2.0 astronomical units (AU, the mean distance from Earth to the sun) and 3.5 AU. Alessandro Morbidelli of the Observatory of the Côte d'Azur in France and his co-workers have shown that asteroids whose orbits were highly inclined to the plane of the solar system could have continued to pelt Earth for a similar period.

THE AUTHOR

JAMES F. KASTING **received his bachelor's degree in chemistry and physics from Harvard University. He went on to graduate studies in physics and atmospheric science at the University of Michigan at Ann Arbor, where he obtained a doctorate in 1979. Kasting worked at the National Center for Atmospheric Research and for the NASA Ames Research Center before joining Pennsylvania State University, where he now teaches in the departments of geosciences and meteorology. Kasting's research focuses on the evolution of habitable planets around the sun and other stars.**

This widely accepted theory of an ancient cometary firehose has recently hit a major snag. Astronomers have found that three comets—Halley, Hyakutake and Hale-Bopp—have a high percentage of deuterium, a form of hydrogen that contains a neutron as well as a proton in its nucleus. Compared with normal hydrogen, deuterium is twice as abundant in these comets as it is in seawater. One can imagine that the oceans might now hold proportionately more deuterium than the cometary ices from which they formed, because normal hydrogen, being lighter, might escape the tug of gravity more easily and be lost to space. But it is difficult to see how the oceans could contain proportionately less deuterium. If these three comets are representative of those that struck here in the past, then most of the water on Earth must have come from elsewhere.

A controversial idea based on observations from satellites suggests that about 20 small (house-size) comets bombard Earth every minute. This rate, which is fast enough to fill the entire ocean over the lifetime of Earth, implies that the ocean is still growing. This much debated theory, championed by Louis A. Frank of the University of Iowa, raises many unanswered questions, among them: Why do the objects not show up on radar? Why do they break up at high altitude? And the deuterium paradox remains, unless these "cometesimals" contain less deuterium than their larger cousins.



HABITABLE ZONE, where liquid water can exist on the surface of a planet, now ranges from just inside the orbit of Earth to beyond the orbit of Mars (*blue*). This zone has migrated slowly outward from its position when the planets first formed (*yellow*), about 4.6 billion years ago, because the sun has gradually brightened over time. In another billion years, when Earth no longer resides within this expanding zone, the water in the oceans will evaporate, leaving the world as dry and lifeless as Venus is today.

EDWARD BELL

More recently, Morbidelli has argued convincingly that most of Earth's water came from the asteroid belt. The ordinary chondrites are thought to come from the inner part of this region (2.0 to 2.5 AU). But outer-belt asteroids (2.5 to 3.5 AU) are thought to be water-rich. According to Morbidelli, as Earth formed it collided with one or more large planetesimals from the outer belt. Gravitational perturbations caused by Jupiter elongated the planetesimal's orbit, allowing it to pass within Earth's orbit. Earth may have picked up additional water from asteroids on highly inclined orbits that arrived during the heavy bombardment period. In this scheme, no more than 10 percent of Earth's water came from comets that originated farther out in the solar system. This theory is consistent with deuterium-hydrogen ratios, which indicate that the comets' watery contributions were small.

The Habitable Zone

WHATEVER THE SOURCE, plenty of water fell to Earth early in its life. But simply adding water to an evolving planet does not ensure the development of a persistent ocean. Venus was probably also wet when it formed, but its surface is completely parched today.

How that drying came about is easy to understand: sunshine on Venus must have once been intense enough to create a warm, moist lower atmosphere and to support an appreciable amount of water in the upper atmosphere as well. As a result, water on the surface of Venus evaporated and traveled high into the sky, where ultraviolet light broke the molecules of H_2O apart and allowed hydrogen to escape into space. Thus, this key component of water on Venus took a one-way route: up and out.

This sunshine-induced exodus implies that there is a critical inner boundary to the habitable zone around the sun, which lies beyond the orbit of Venus. Conversely, if a planet does not receive enough sunlight, its oceans may freeze by a process called runaway glaciation. Suppose Earth somehow slipped slightly farther from the sun. As the solar rays faded, the climate would get colder and the polar ice caps would expand. Because snow and ice reflect more sunlight back to space, the climate would become colder still. This vicious cycle could explain in part why Mars, which occupies the next orbit out from Earth, is frozen today.

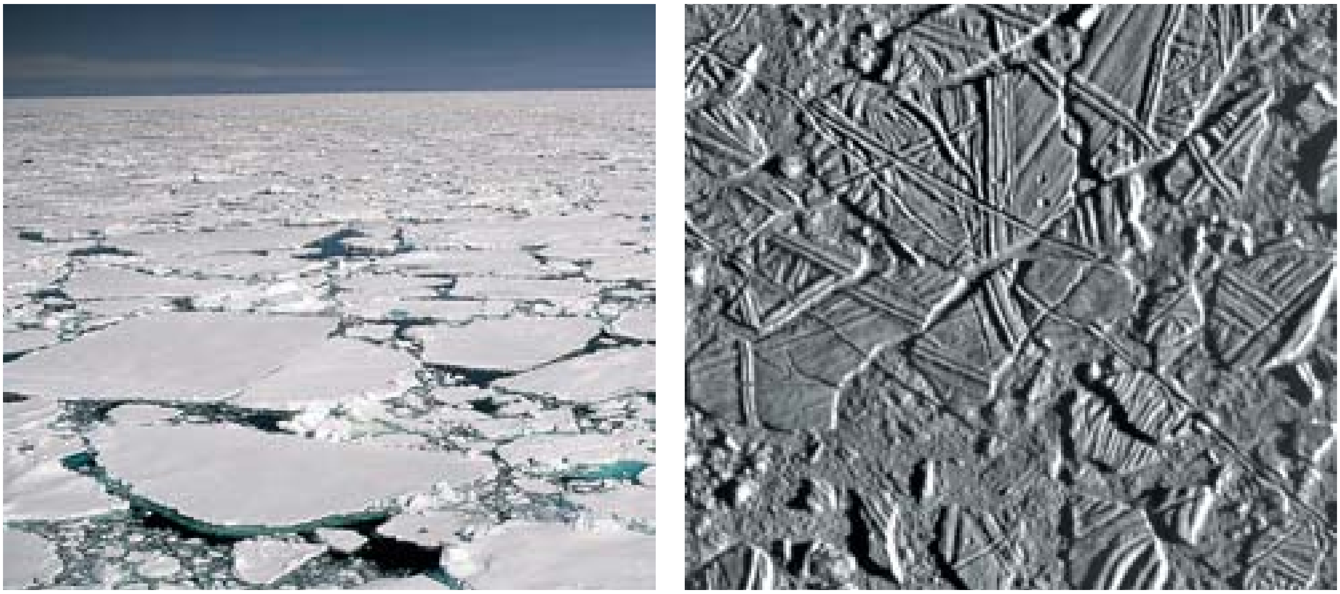
The actual story of Mars is probably more complicated. Pictures taken from the Mariner and Viking probes and from the Global Surveyor spacecraft show that older parts of the Martian surface are laced with channels carved by liquid water. Measurements from the laser altimeter on board the Global Surveyor indicate that the vast northern plains of Mars are exceptionally flat. The only correspondingly smooth surfaces on Earth lie on the seafloor, far from the midocean ridges. Thus, many scientists are now even more confident that Mars once had an ocean. Mars, it would seem, orbits within a potentially habitable zone around the sun. But somehow, eons ago, it plunged into its current chilly state.

A Once Faint Sun

UNDERSTANDING THAT dramatic change on Mars may help explain nagging questions about the ancient oceans of Earth. Theories of solar evolution predict that when the sun first became stable, it was 30 percent dimmer than it is now. The smaller solar output would have caused the oceans to be completely frozen before about two billion years ago. But the geologic record tells a different tale: liquid water and life were both present as early as 3.8 billion years ago. The disparity between this prediction and fossil evidence has been termed the faint young sun paradox.

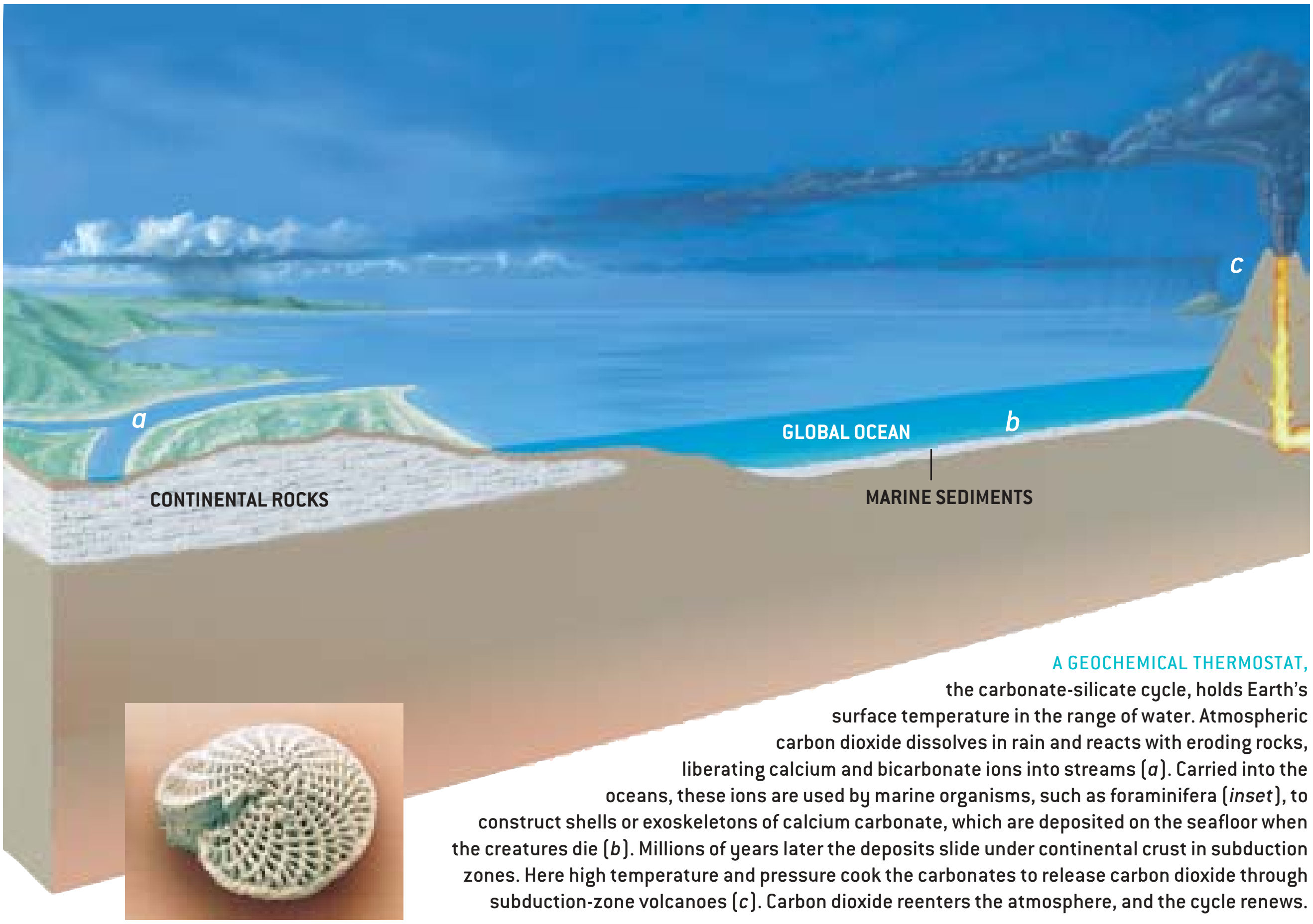
The paradox disappears only when one recognizes that the composition of the atmosphere has changed considerably over time. The early atmosphere probably contained much more carbon dioxide than at present and perhaps more methane. Both these gases enhance the greenhouse effect because they absorb infrared radiation; their presence could have kept the early Earth warm, despite less heat coming from the sun.

The greenhouse phenomenon also helps to keep Earth's climate in a dynamic equilibrium through a process called the carbonate-silicate cycle. Volcanoes continually belch carbon dioxide into the atmosphere. But silicate minerals on the continents absorb much of this gas as they erode from crustal rocks and wash out to sea. The carbon dioxide then sinks to the bottom of the ocean in the form of solid calcium carbonate. Over millions of years, plate tectonics drives this carbonate down into the up-



ICY BLOCKS cover the surface of the Weddell Sea off Antarctica (*left*); similarly shaped blocks blanket the surface of Europa, a moon of Jupiter (*right*). This resemblance, and the lack of craters on Europa, suggests that liquid water exists below the frozen surface of that body.

GALEN ROWELL *Mountain Light* (*left*); NASA/JET PROPULSION LABORATORY (*right*)



A GEOCHEMICAL THERMOSTAT, the carbonate-silicate cycle, holds Earth's surface temperature in the range of water. Atmospheric carbon dioxide dissolves in rain and reacts with eroding rocks, liberating calcium and bicarbonate ions into streams (*a*). Carried into the oceans, these ions are used by marine organisms, such as foraminifera (*inset*), to construct shells or exoskeletons of calcium carbonate, which are deposited on the seafloor when the creatures die (*b*). Millions of years later the deposits slide under continental crust in subduction zones. Here high temperature and pressure cook the carbonates to release carbon dioxide through subduction-zone volcanoes (*c*). Carbon dioxide reenters the atmosphere, and the cycle renews.

TOM MOORE (*illustration*); ANDREW SYRED *SPL/Photo Researchers, Inc.* (*inset*)

per mantle, where it reacts chemically and is spewed out as carbon dioxide again through volcanoes.

If Earth had ever suffered a global glaciation, silicate rocks, for the most part, would have stopped eroding. But volcanic carbon dioxide would have continued to accumulate in the atmosphere until the greenhouse effect became large enough to melt the ice. And eventually the warmed oceans would have released enough moisture to bring on heavy rains and to speed erosion, in the process pulling carbon dioxide out of the atmosphere and out of minerals. Thus, Earth has a built-in thermostat that automatically maintains its surface temperature within the range of liquid water.

Paul Hoffman and Daniel Schrag of Harvard University have argued that Earth did freeze over at least twice during the Late Precambrian era, 600 to 750 million years ago. Earth recovered with a buildup of volcanic carbon dioxide. This theory remains controversial because scientists do not fully understand how the biota would have survived, but I am convinced it happened. There is no other good way to explain the evidence for continental-scale, low-latitude glaciation. Six hundred million years ago, Australia straddled the equator, and it was glaciated from one end to the other.

The same mechanism may have operated on Mars. Although the planet is now volcanically inactive, it once had many eruptions and could have maintained a vigorous carbonate-silicate cycle. If Mars has sufficient stores of carbon—one question that NASA scientists hope to answer with the Global Surveyor—it could also have had a dense shroud of carbon dioxide at one time. Clouds of carbon dioxide ice, which scatter infrared radiation, and perhaps a small amount of methane would have generated enough greenhouse heating to maintain liquid water on the surface.

Mars is freeze-dried today not because it is too far from the sun but because it is a small planet and therefore cooled off comparatively quickly. It was unable to sustain the volcanism necessary to maintain balmy temperatures. Over the eons, the water ice that remained probably mixed with dust and is now trapped in the uppermost few kilometers of the Martian crust.

The conditions on Earth that formed and maintain the oceans—an orbit in the habitable zone, plate tectonics creating ocean basins, volcanism driving a carbonate-silicate cycle, and a stratified atmosphere that prevents loss of water or hydrogen—are unique among the planets in our solar system. But other planets are known to orbit other stars, and the odds are good that similar conditions may prevail, creating other brilliantly blue worlds, with oceans much like ours. SA

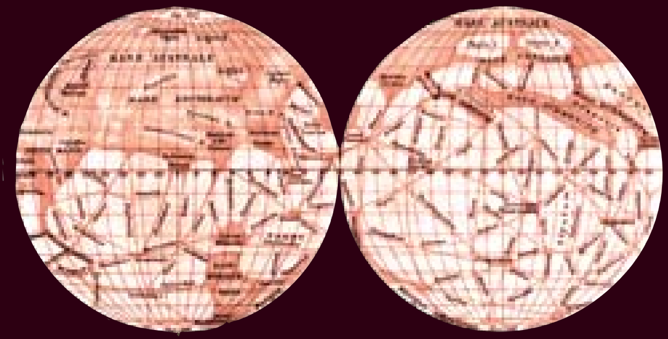
MORE TO EXPLORE

How Climate Evolved on the Terrestrial Planets. James F. Kasting, Owen B. Toon and James B. Pollack in *Scientific American,* Vol. 258, No. 2, pages 90–97; February 1988.

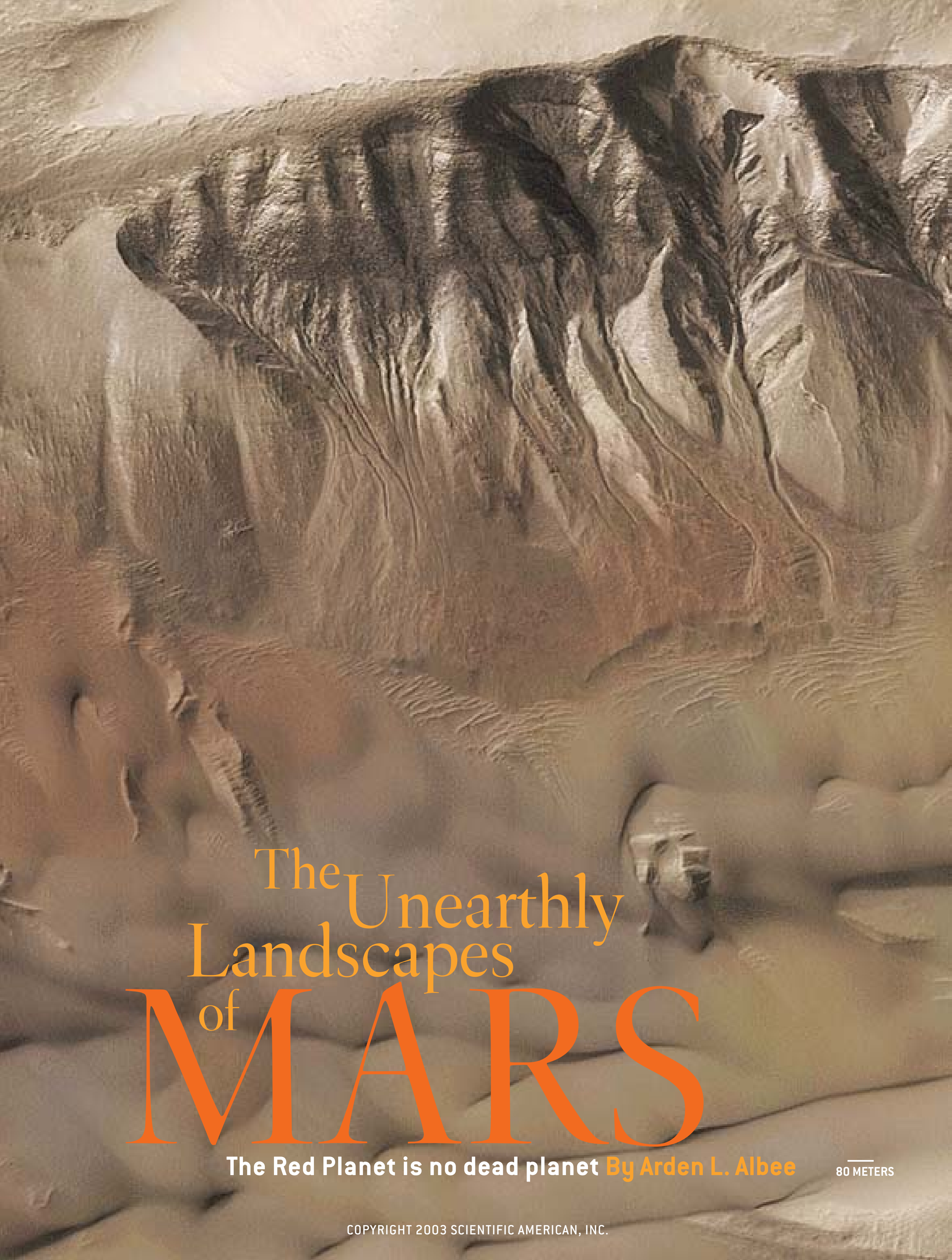
Source Regions and Timescales for the Delivery of Water to Earth. Alessandro Morbidelli et al. in *Meteoritics and Planetary Science,* Vol. 35, Issue 6, pages 1309–1320; November 2000.

A Plausible Cause of the Late Heavy Bombardment. Alessandro Morbidelli, J.-M. Petit, B. Gladman and J. Chambers in *Meteoritics and Planetary Science,* Vol. 36, Issue 3, pages 371–380; March 2001.

Captain John Carter, the hero of the adventure novels of Edgar Rice Burroughs, was a gentleman of Virginia and an officer of the Confederacy. Impoverished after the Civil War, he went looking for gold in Arizona and, while being chased by Apache warriors, fell and struck his head. He returned to consciousness on an arid planet with twin moons, populated by six-legged creatures and beautiful princesses who knew the place as "Barsoom." The landscape bore an uncanny resemblance to southern Arizona. It was not entirely dissimilar to Earth, only older and decayed. "Theirs is a hard and pitiless struggle for existence upon a dying planet," Burroughs wrote in the first novel.



IT WOULD TAKE YOU ABOUT HALF AN HOUR to hike across the area shown in this image, on the north side of Newton Crater in the southern hemisphere of Mars. You would leave your footprints on lightly frosted soil (*bright areas*), clamber over windblown features such as sand dunes and jump across possibly water-carved features such as gullies. These landforms probably continue to form even today. Like other Mars Global Surveyor images, this one is a composite of high-resolution grayscale and low-resolution color; the colors are only approximate. It is a far cry from the vague (and often fanciful) view of Mars a century ago (*above*).



The Unearthly Landscapes of MARS

The Red Planet is no dead planet By Arden L. Albee

In science as well as science fiction, Mars is usually depicted as a version of Earth in its extreme—smaller, colder, drier, but sculpted by basically the same processes. Even well into the 20th century, many thought the planet had flowing water and proliferating plants. The resemblance to Earth fell apart when spacecraft in the late 1960s revealed a barren, cratered world, more like the moon. But it quickly returned with the subsequent discoveries of giant mountains, deep canyons and complex weather patterns. The Viking and Mars Pathfinder images from the surface look eerily Earth-like. Like Burroughs, researchers compare the equatorial regions of Mars to the American Southwest. For the polar regions, the model is the Dry Valleys of Antarctica, a frozen desert in a landscape of endless ice.

But if there is one thing researchers have learned from recent Mars exploration, it is to be careful about drawing such comparisons. In the past five years, spacecraft have collected more information about the Red Planet than in all previous missions combined. Mars has proved to be a very different and more complicated planet than scientists thought beforehand. Even the single biggest question—Was Mars once warm and wet, possibly hospitable to the evolution of life?—is more nuanced than people have tended to assume. To make sense of Mars, investigators cannot be blinded by their experience of Earth. The Red Planet is a unique place.

Mars as the Abode of Dust

MARS EXPLORATION has certainly had its up and downs. In the past decade NASA has lost three spacecraft at Mars: Mars Observer, Mars Climate Orbiter (intended as a partial replacement for Mars Observer) and Mars Polar Lander. Lately, though, the program has had a run of successes. Mars Global Surveyor has been taking pictures and collecting infrared spectra and other data continuously since 1997. It is now the matriarch of a veritable family of Mars spacecraft. Another, Mars Odyssey, has been orbiting the planet for more than a year, mapping the water content of the subsurface and making infrared images of the surface. This summer NASA launched the Mars Exploration Rovers, successors to the famous Sojourner rover of Mars Pathfinder [*see box on page 42*]. Around the same time, the European Space Agency launched the Mars Express orbiter, with its Beagle 2 lander. The Nozomi orbiter, sent by the Institute of Space and Astronautical Science in Japan, should arrive at Mars in December.

Never before have scientists had such a comprehensive record of the processes that operate on the surface and in the atmosphere [*see box on page 44*]. They have also studied the craters, canyons and volcanoes that are dramatic relics of the distant past. But there is a huge gap in our knowledge. Between ancient Mars and modern Mars are billions of missing years. No one is sure of the conditions and the processes that sculpted Mars during most of its history. Even less is known about the subsurface geology, which will have to be the subject of a future article.

Present-day Mars differs from Earth in a number of broad respects. First, it is enveloped in dust. Much of Earth's surface consists of soil derived by chemical weathering of the underlying bedrock and, in some regions, glacial debris. But much of Mars's surface consists of dust—very fine grained material that has settled out of the atmosphere. It drapes over all but the steepest features, smothering the ancient landscape. It is thick even on the highest volcanoes. The dustiest areas correspond to the bright areas of Mars long known to telescope observers.

Dust produces otherworldly landscapes, such as distinctively pitted terrain. As dust settles through the atmosphere, it traps volatile material, forming a mantle of icy dust. Later on, the volatile ices turn to gas, leaving pits. Intriguingly, the thickness of the icy, dusty mantle on Mars varies with latitude; near the poles, Mars Odyssey has shown, as much as 50 percent of the upper meter of soil may be ice. On slopes, the icy mantle shows signs of having flowed like a viscous fluid, much in the manner of a terrestrial glacier. This mantle is becoming the focus of intense scientific scrutiny.

Second, Mars is extremely windy. It is dominated by aeolian activity in much the way that Earth is dominated by the action of liquid water. Spacecraft have seen globe-encircling dust storms, huge dust devils and dust avalanches—all wrought by the wind. Dust streaks behind obstacles change with the seasons, presumably because of varying wind conditions.

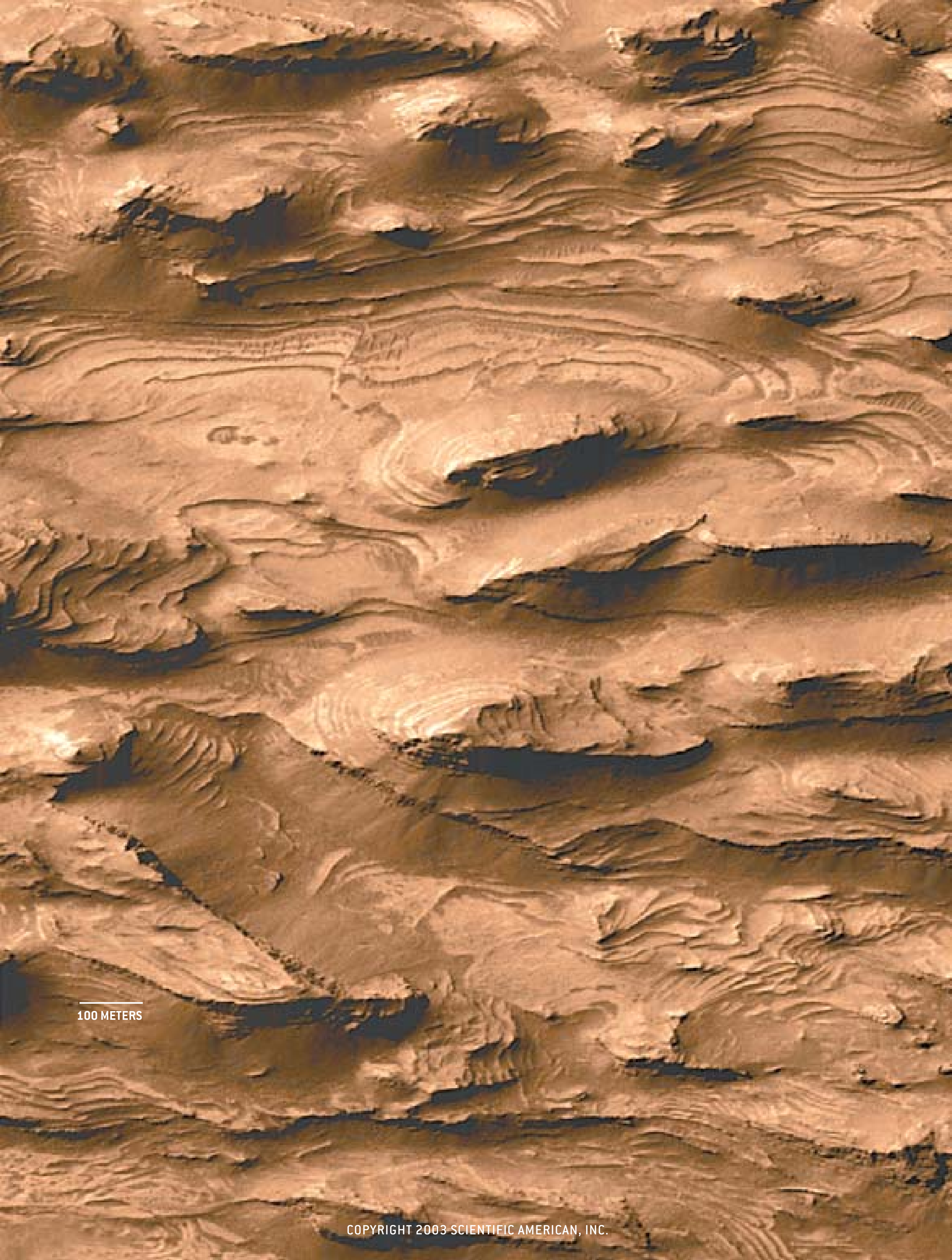
Where not dust-covered, the surface commonly shows aeolian erosion or deposition. Evidence for erosion shows up in craters, from which material appears to have been removed by wind, and in yardangs, bedrock features that clearly have been carved by windblown sand. Evidence for deposition includes sand sheets and moving sand dunes. The lat-

LAYERED TERRAIN looks surreal, almost like a topographic map, but is quite real. It covers the floor of western Candor Chasma, a ravine that is part of the Valles Marineris canyon system. Scientists have identified 100 distinct layers, each about 10 meters thick. They could be sedimentary rock originally laid down by water, presumably before the canyon cut through the terrain. Alternatively, the layers could be dust deposited by cyclic atmospheric processes. This image was taken by Mars Global Surveyor.

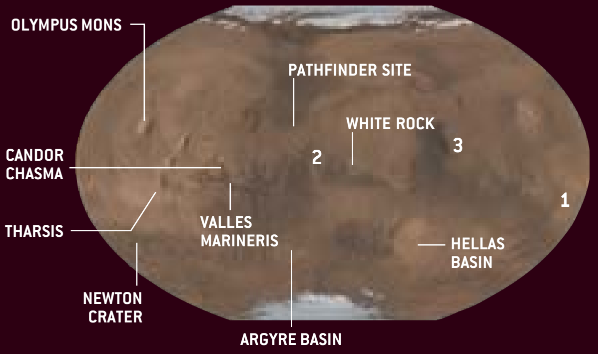
Overview/*The Martian Surface*

- Two ongoing missions to Mars, Mars Global Surveyor and Mars Odyssey, are raising difficult questions about the Red Planet. Flowing water, ice and wind have all helped to carve the landscape over the past several billion years. The processes are both similar and dissimilar to those acting on Earth's surface. Scientists' experience of Earth has sometimes led them astray.
- The question of whether Mars was once hospitable is more confusing than ever. Spacecraft have gathered evidence both for and against the possibility. Three landers now en route—two American and one European—could prove crucial to resolving the matter.

COURTESY OF www.nirgal.net/schemas/schiaparelli_map.html (*page 36*); NASA/JPL/MALIN SPACE SCIENCE SYSTEMS MOC 2–320 (*page 37*); NASA/JPL/MALIN SPACE SCIENCE SYSTEMS MOC 2–259 (*opposite page*)

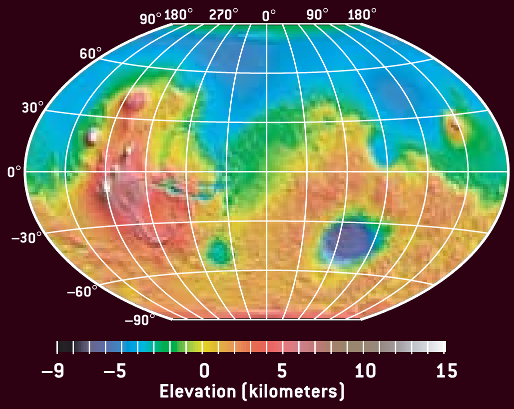

100 METERS

GLOBAL VIEWS of MARS

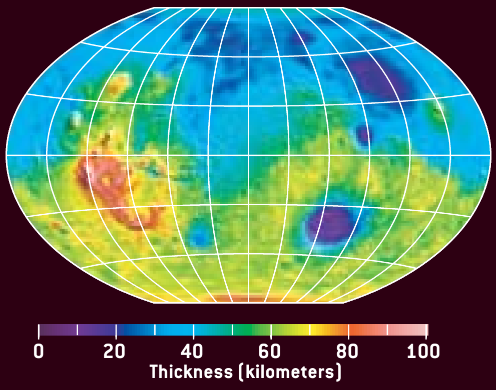


Landing Sites: Gusev Crater (1), *Meridiani Planum* (2), *Isidis Planitia* (3)

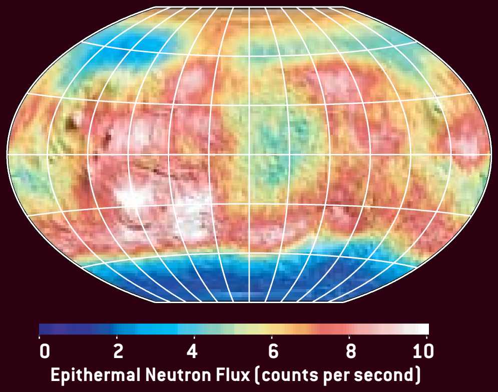
TRUE COLOR Mars is four worlds in one: the heavily cratered southern hemisphere (with riverlike valley networks), the smoother northern hemisphere (with hints of an ancient shoreline), the equatorial region (with giant volcanoes and canyons), and the polar caps (with bizarre, protean terrain). This map combines wide-angle camera images with altimetry, which brings out details. The color is realistic.



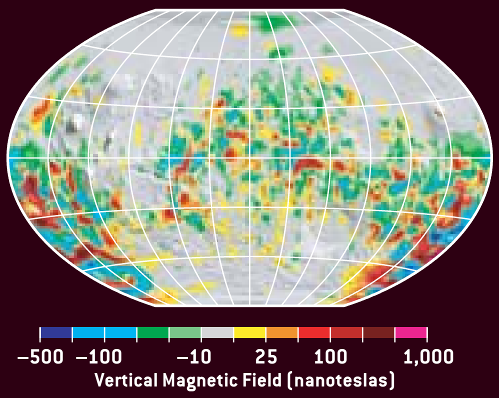
TOPOGRAPHY The elevation spans 30 kilometers from the lowest basins (*dark blue*) to the highest volcanoes (*white*). For comparison, the range of elevation on Earth is only 20 kilometers. The large blue circle in the southern hemisphere represents the Hellas impact basin, one of the biggest craters in the solar system. Girdling Hellas is a vast ring of highlands about two kilometers in elevation.



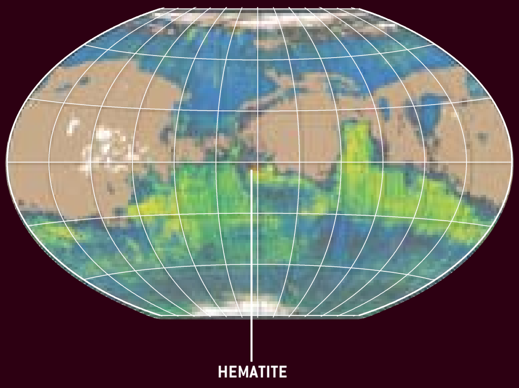
CRUSTAL THICKNESS Combining the topographic map with measurements of Mars's gravity, researchers have deduced the thickness of the Martian crust: roughly 40 kilometers under the northern plains and 70 kilometers under the far southern highlands. The crust is especially thick (*red*) under the giant Tharsis volcanoes and thin (*purple*) under the Hellas impact basin.



WATER Neutrons reveal the presence of water in the top meter of soil. The energy of these particles, which are produced when cosmic radiation bombards the soil, is sapped by the hydrogen within water molecules. A dearth of medium-energy ("epithermal") neutrons means water-rich soil (*blue*). The implied amount of water, most of it in the far south, would fill two Lake Michigans. More may lie deeper underground.



MAGNETISM Mars lacks a global magnetic field, yet areas of its crust are magnetized up to 10 times as strongly as Earth's crust. In these areas, iron-rich rocks have become bar magnets, suggesting that Mars had a global field at the time the rocks solidified from a molten state. The east-west banding resembles patterns produced by plate tectonics on Earth, but its origin is unknown.



GEOLOGY Infrared spectral measurements reveal rock types. Basalt (*green*), a primitive volcanic rock, dominates the southern hemisphere. Andesite (*blue*), a more complex volcanic rock, seems to be common in the north. Near the equator is an outcrop of hematite (*red*), a mineral typically produced in the presence of water. In large regions, dust (*tan*) or clouds (*white*) hide the underlying rock types.

MARS ORBITER LASER ALTIMETER SCIENCE TEAM, MALIN SPACE SCIENCE SYSTEMS (*true color mosaic*); ODED AHARONSON *California Institute of Technology* (*all other mosaics*);
SOURCE: MARS ODYSSEY MISSION (*water*); PHILIP CHRISTENSEN *Arizona State University* (*geology*); MARS GLOBAL SURVEYOR MISSION (*all other mosaics*)

ter are composed of sand-size grains, which move around by saltation—multiple bounces of windblown grains along the ground. It takes a stronger wind to loft dust directly than to initiate saltation, so this phenomenon accounts for most of the dust kicked into the atmosphere.

Aeolian activity seems to have persisted since the time of heavy cratering, back when the solar system was still young. Many images show craters with varying degrees of erosion: some are shallow and partially filled with deposits and sand dunes, whereas others are pristine—deeper and bowl-shaped. Michael C. Malin and Kenneth S. Edgett of Malin Space Science Systems in San Diego, the research firm that operates the Mars Global Surveyor camera, have inferred a sequence of processes: Sand was blown through the region, and some of it got trapped in craters; other craters formed later. Where and how such a volume of sand was produced and how it was blown around remain a mystery, however.

The Angry Skies of Mars

A THIRD WAY in which Mars differs from Earth is in its amazing variety of weather and climate cycles, many of which are similar to those on Earth, many like nothing on Earth. The Martian day is almost the same as an Earth day, but the Martian year is 687 Earth days. The tilt of Mars's rotation axis, which produces seasons, is very close to that of Earth's. Mars lacks the precipitation and oceans that are so crucial to weather on Earth. But the atmospheric pressure (less than 1 percent of that on Earth) varies seasonally by about 25 percent, driven by the condensation and sublimation of carbon dioxide frost at the poles. The thin atmosphere has a very low heat capacity, so the surface temperature swings by more than 100 degrees Celsius from day to night. The thermal properties of the thin atmosphere are dramatically affected by dust and ice particles suspended in the air. The upshot is that, despite being so thin, the atmosphere has complex circulation patterns and dynamics. A daily weather report might talk of strong winds, high-level ice clouds, low-level fogs, seasonal frost, dust devils and massive dust storms.

As on Earth, storm systems often spiral southward from the northern polar regions. But the largest dust storms typically start during the southern spring as the planet rapidly heats up. Periodically they coalesce and come to encircle the entire planet. Mars Global Surveyor closely followed the evolution of a four-month global dust storm that started in June 2001. Contrary to scientists' expectations, it was not, in fact, one single global storm, but the confluence of several regional storms. Malin has compared the climatic effect of the dust raised by this storm with the aftermath of Mount Pinatubo's eruption on Earth in 1991—namely, a brief but widespread cooling.

The polar ice caps play a key role in the atmospheric cycles. Their size and shape, as shown by topographic measurements, indicate that the caps are predominantly water ice, as opposed to so-called dry ice, made of carbon dioxide: dry ice is not as rigid as water ice, and it could not support the observed domelike shape. A major new discovery has been that the layer of dry ice that covers much of the south polar cap is being eroded away at a high rate. Clearly, the erosion cannot go on forever. Nor can the current dust sinks and sources remain in their current states indefinitely. To replenish the ice and dust, other cycles must be occurring, perhaps tied to orbital variations. Malin and Edgett have suggested that wind conditions may be less intense now than in the fairly recent past, another hint that the Martian climate changes with time.

A fourth major difference between Earth and Mars is the behavior of liquid water. Liquid water is unstable at the surface under present pressure and temperature conditions. It does not rain. Still, water ice can—and does—persist at some depth within the Martian soil during all or much of the year. On Mars, as on Earth, several types of patterned ground mark the presence of ice-rich soil. Mars Odyssey has detected ground ice over most of the planet outside the equatorial regions, and models predict that the ice extends to considerable depths.

Liquid water can sometimes leak onto the surface. In 2000 Malin and Edgett described fresh gullies that look like water-carved features on Earth [see "Gully Gee Whiz," by George Musser; SCIENTIFIC AMERICAN, News & Analysis, September 2000].

In the ensuing excitement, scientists advanced many theories to explain them: leaking aquifers (which would be inexplicably perched high on crater rims); pressurized geysers of water; high-pressure outbursts of carbon dioxide gas; volcanic heat sources at depth. Finally, earlier this year Philip Christensen of Arizona State University discovered gullies that clearly emerge from underneath a bank of snow and ice. He concluded that they are related to Martian climate cycles. In colder periods, slopes become blanketed with a mixture of snow and dust. Sunlight penetrates this insulating blanket, heating it enough for water to melt under the snow and to run down the slope, creating gullies. In warmer periods, the blanket melts or evaporates entirely, exposing the gullies.

Layer upon Layer

DESPITE THE ABUNDANCE of water, however, Mars is arid. It has the mineralogy of a nearly waterless surface. On Earth, the action of warm liquid water produces weathered, quartz-rich soils, hydrated clays, and salts such as calcium carbonate and sulfate. Beach sand and sand dunes are largely quartz. On Mars, spacecraft have yet to find any deposits of these minerals. The darker Martian

THE AUTHOR

ARDEN L. ALBEE is the project scientist—that is, overall leader of the science team—for the Mars Global Surveyor mission. He had the same role for the ill-fated Mars Observer mission, for which Surveyor is a partial replacement. Albee is emeritus professor of geology and planetary science at the California Institute of Technology and served as chief scientist of NASA's Jet Propulsion Laboratory from 1978 until 1984. His research interests run from field geology to compositional analysis of rocks, meteorites, comets and lunar samples. He still makes time for his eight children and 11 grandchildren.

dunes are basaltic, consisting mainly of minerals such as pyroxene and plagioclase, which on Earth would readily weather away. It follows that the present cold and dry atmospheric conditions have persisted since a time far back in the planet's history.

Has Mars always differed so much from Earth? Below the mantles of dust and sand are numerous signs that the Red Planet has transformed over time. To begin with, there is a striking dichotomy in landscape between the planet's northern and southern hemispheres. The southern hemisphere is high in altitude and heavily cratered (indicating an ancient surface). The northern one is a vast, low-lying plain with fewer craters (indicating a younger age). In between is the immense Tharsis Plateau, intermediate in age and capped by volcanoes that dwarf any on Earth. Using the new high-resolution data on these volcanoes, James W. Head III of Brown University has found flow patterns that look strikingly like mountain glaciers—and that may suggest the presence of ice under a blanket of rock and dust.

The northern lowlands are exceedingly level, leading to speculation that they were lake beds during a significant chunk of Martian history. They appear to be covered with multiple layers of volcanic flows and sedimentary debris that originated in the south. Detailed new topographic maps have unveiled "stealth craters"—faint circular expressions, evidently part of an ancient cratered surface that has been buried by a thin layer of younger deposits.

Along the edge of the southern highlands are features that could only have been carved by liquid water. These features are tremendously larger than their counterparts on Earth. The famous canyon Valles Marineris would run from Los Angeles to New York with a width extending from New York to Boston and a depth similar to the elevation of Mount McKinley. No terrestrial canyon comes close. At its head is a jumbled terrain, intimating that water flowed not in a steady trickle but in concentrated, catastrophic outflows, scouring the surface along its path. Other Martian outflow channels have similar features, which, because they are carved into the Tharsis Plateau, must have an intermediate age.

Streamlined islands and other features in these channels look much like the scablands in the northwestern U.S., which were gouged by the Spokane Flood toward the end of the last ice age, about 10,000 years ago. During the massive deluge, a lake roughly the size of one of the Great Lakes burst its ice dam and rushed out within just a few days. On Mars, such calamities were 10 to 100 times as devastating. They may have been triggered by volcanic heat sources or by the general heat flow from the interior of the planet. Heat would have melted ice underneath the thick permafrost layer, building up tremendous pressures until the water finally burst out.

The most contentious water-related features of all are the valley networks. Located throughout the southern highlands, they have a branching, dendritic pattern reminiscent of rivers on Earth—suggesting that they were formed by surface runoff from rainfall or snowfall. They are the strongest hint that Mars was once as warm as Earth. But these networks look rather different from rain-fed rivers on Earth. They more closely resemble river networks in desert areas, which are fed by water that slowly seeps from subterranean sources. Such streams typically originate in steep-walled amphitheaters rather than in ever smaller tributaries. Heated debates have been taking place at scientific meetings over the crucial question: Did it rain on early Mars?

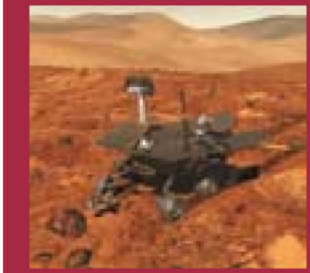
The timing of the water networks could be the key to making sense of them. Recent detailed studies of the northern

Going for a Drive

THIS SUMMER NASA LAUNCHED twin rovers to the Red Planet, and the European Space Agency has sent off a lander, too. The three of them, scheduled to arrive next January, will be robotic geologists—studying the geologic history of landing sites, investigating what role water played there and determining how suitable past conditions would have been for life.

The rovers deserve particular mention, because they will give scientists unprecedented mobility. Each rover can travel 100 meters a day. For comparison, the Sojourner rover on the 1997 Mars Pathfinder lander traversed 100 meters over its entire mission [see "The Mars Pathfinder Mission," by Matthew Golombek; Scientific American, July 1998]. A mast, about 1.5 meters high, supports a binocular camera and a thermal-emission spectrometer, one of many instruments that can analyze the composition of materials. A robot arm holds the other instruments: a Mossbauer spectrometer, an alpha-particle detector, an x-ray spectrometer and a microscope. The arm also carries a scraper to clean off rock surfaces for study. A dish antenna beams signals directly to Earth, and a black rod antenna relays data through the Mars Global Surveyor or Mars Odyssey orbiter.

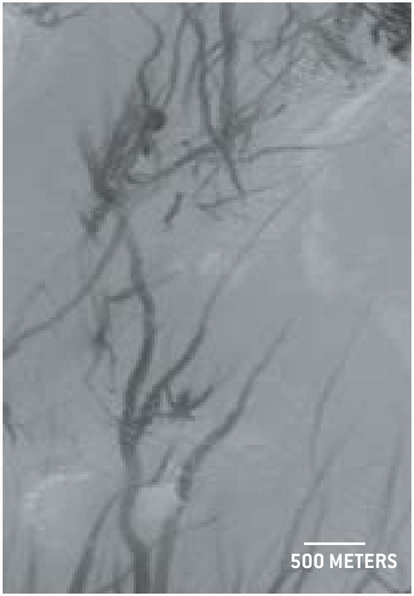
Deciding on their landing sites has been one of the most exciting uses of data from the orbiters. Unlike the selection of sites for the Viking landers in 1976—which was a matter of a few people's gut instinct—the choice for the rovers involved a long deliberation among dozens of scientists and engineers. Weighing tantalizing geology (such as suspected water-related features) against potential dangers (such as steep slopes and high winds), they winnowed an initial list of more than 150 possible sites down to seven, then four and finally, on April 11, two: Gusev Crater, whose layered deposits might be lake-bed sediments, and Meridiani Planum, which is rich in coarse-grained hematite, a mineral typically formed in association with liquid water. (The European Beagle 2 lander will touch down on Isidis Planitia, a possible sedimentary basin.) —*A.L.A.*

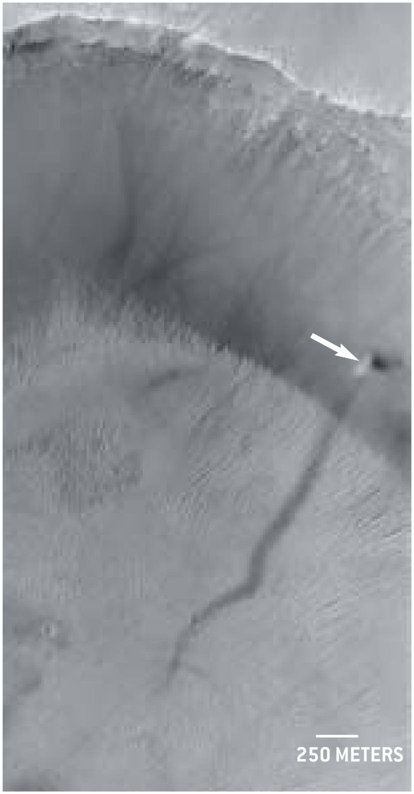


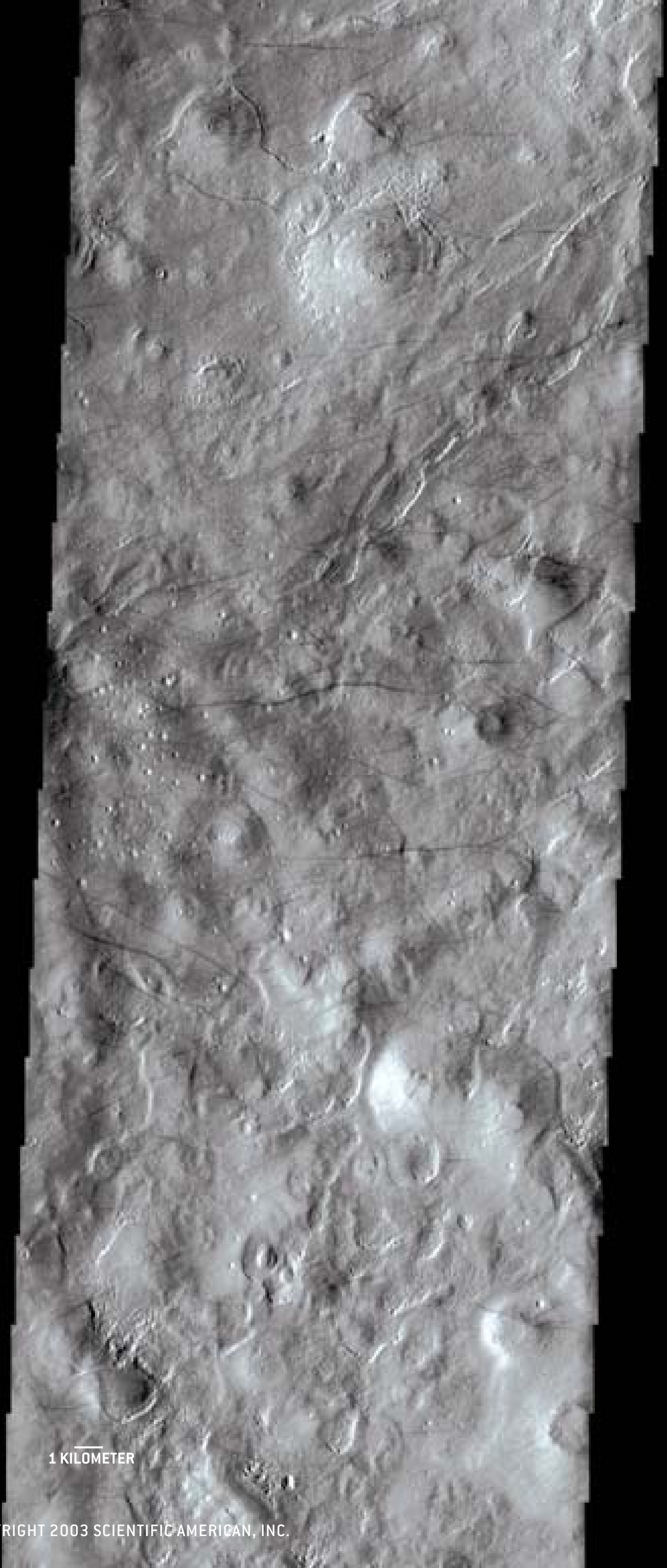
NASA's Mars Exploration Rover

NASA ARTIST CONCEPT/MARS EXPLORATION PROGRAM

IN THE PLAINS northwest of Olympus Mons, dust devils are sweeping over the land and leaving streaks in their wake (*right*). A similar scene has unfolded in the Argyre Basin (*below*) and east of Valles Marineris (*bottom*), where a devil was caught in the act. These tornadolike vortices—thought to be created as warm air rises off the surface—clear away light-colored dust and expose comparatively dark soil. They are one of the many wind-related processes that are continuously reshaping the Martian surface. The picture at the right is from Mars Odyssey; the two below are from Mars Global Surveyor.







NASA/ARIZONA STATE UNIVERSITY THEMIS SCIENCE TEAM, FEBRUARY 6, 2003 (*right*); NASA/JPL/MALIN SPACE SCIENCE SYSTEMS MOC 2-220 (*left top*) AND MOC 2-318 (*left bottom*)

edge of the highlands show that immense amounts of material eroded during—rather than after—the intense meteor bombardment that took place early in Martian history. These analyses imply that the distribution of water kept changing as impacts reworked the landscape. Craters filled up with water and debris, and channels began to link them together into a network, but impacts continually disrupted this process. For instance, the Argyre Basin, 1,000 kilometers in diameter, may once have been filled to its brim with water. It is part of a valley system that brought water from near the south pole, through the basin, into channels that crossed the equator. The roles of water and ice in these systems, both aboveground and underground, remain unclear. In any case, these networks are very different from hydrologic systems on Earth.

A final clue to Martian history comes from one of the biggest surprises delivered by Mars Global Surveyor, the extent to which the uppermost crust consists of layered deposits. Almost everywhere that the subsurface is exposed—on walls of canyons, craters, mesas and valleys—it is layered. The layers differ from one another in thickness, color and strength. They show that the Martian surface has undergone complex sequences of deposition, crater formation and erosion. The oldest layers are the most extensive. The higher layers have been partially stripped away, apparently blown by the wind.

Where did the layers come from? The lack of boulderlike blocks argues against their being volcanic flows, although they could be volcanic ash. Ultimately, however, most of the layers probably originated in impact debris. On the moon, scientists observe overlapping rings of impact debris, which mark craters of differing ages. Similarly, Mars is so heavily cratered that the upper crust has been stirred up like soil tilled by a gardener. Water and wind then scattered this material.

Blue Mars?

IN A SENSE, scientists' ideas about early Mars are more uncertain than they have ever been. This doubt comes to the fore when researchers address the question of liquid water. The presence or absence of liquid water is fundamental to geologic processes, climate change and the origin of life. The early valley networks and the later flood channels attest to an abundance of water. The evidence for early rainfall suggests that the atmosphere was once much denser. But spacecraft have found no evidence for deposits of carbonate minerals, which would be the vestiges expected from an early dense carbon dioxide atmosphere [see "The Climate of Mars," by Robert M. Haberle; SCIENTIFIC AMERICAN, May 1986].

At this point, scientists have three main hypotheses. Perhaps the early atmosphere was indeed thick. The planet might have had lakes, even oceans, free of ice. Robert A. Craddock of the National Air and Space Museum and Alan

Monitoring Mars, 24/7/687

BOTH MARS GLOBAL SURVEYOR and Mars Odyssey circle the planet in paths that take them over both poles. Their orbits remain fixed as Mars spins below them, allowing the instruments to observe day and night swaths over the entire planet. This continuous coverage can track changes in the surface, atmosphere, gravity and magnetic field.

Global Surveyor has five main instruments. Its laser altimeter has measured the overall shape and the topography of Mars with an altitude precision of about five meters, which means that Mars is now better mapped than most of Earth. The camera takes red and blue medium-resolution images of the entire surface, as well as high-resolution images—1.4 meters per pixel, as good as the pictures taken by the spy satellites of the 1960s—of limited areas. A Michelson interferometer measures the emitted thermal infrared spectrum with high spectral resolution but low spatial resolution, suitable for mapping the mineral composition and thermal properties of the surface. A magnetometer determines the magnetic field. Finally, the spacecraft itself counts as an instrument, because its motion is sensitive to variations in Martian gravity. The gravitational field reveals the thickness of the crust and changes in the size of the polar ice caps.

Odyssey complements Global Surveyor. Its camera lacks a high-resolution mode but takes images in five selected color bands. Its infrared imager has low spectral resolution but high spatial resolution. Another instrument measures gamma-ray and neutron fluxes, which are sensitive to hydrogen just below the planet's surface; Odyssey is therefore the first spacecraft capable of peeking under the surface of Mars, to a depth of about one meter.

These spacecraft also monitor the atmosphere. Cameras scan the entire planet daily, much like Earth-orbiting weather satellites. Twelve times a day the thermal-emission spectrometer takes readings of temperature, pressure, cloud cover and dust abundance. Additionally, as radio transmissions pass through the Martian atmosphere, they are diffracted; signal processing can infer the variation of temperature and pressure with altitude.

—*A.L.A.*

MARS GLOBAL SURVEYOR (artist's conception)

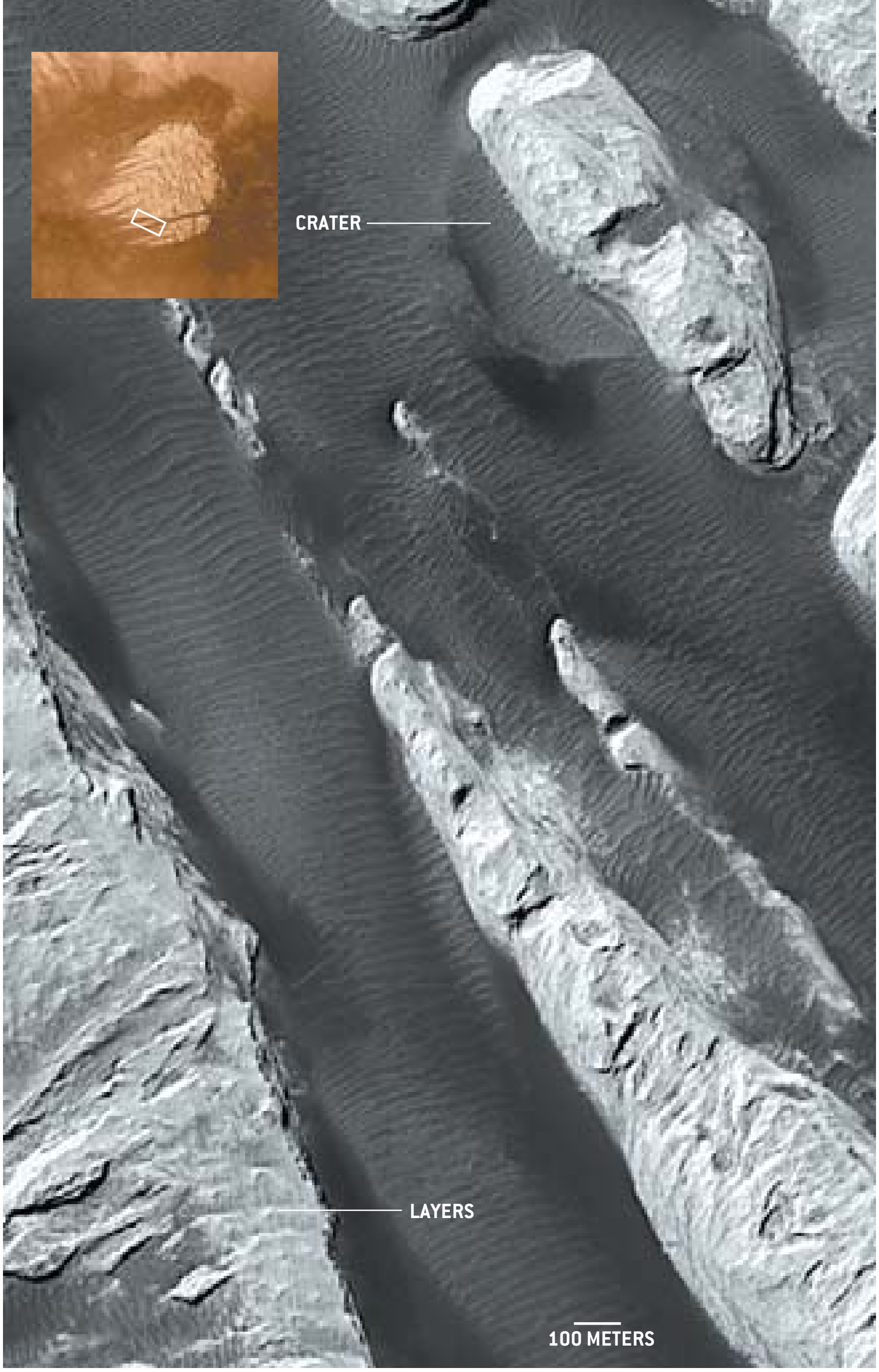
NASA ARTIST CONCEPT/GLOBAL SURVEYOR PROGRAM

D. Howard of the University of Virginia recently suggested that the carbon dioxide was lost to space or locked up in carbonate minerals that have so far escaped detection. Intriguingly, Mars Odyssey spectra have revealed small amounts of carbonate in the dust.

Alternatively, perhaps Mars had a fairly thin atmosphere. It was a wintry world. Any standing bodies of water were covered in ice. Snow might have fallen, recharging the groundwater and leading to temporary trickles of water across the surface. Steven M. Clifford of the Lunar and Planetary Science Institute in Houston, among others, has conjectured that melting under a glacier or a thick layer of permafrost could also have recharged subterranean water sources. Although Mars was bitterly cold, periodic bursts of relatively warmer temperatures could have reinvigorated the planet. Orbital shifts, similar to those that trigger ice ages on Earth, drove these climate cycles. Head, John F. Mustard of Brown and others have pointed to the latitude dependence of the ice and dust cover as evidence for climate change.

Finally, perhaps the climate cycles were insufficient to make Mars warm enough to sustain liquid water. The planet had clement conditions for only brief periods after major impacts. Each such impact deposited water-rich material and pumped enough heat and water into the atmosphere to permit rain. Soon, though, the planet returned to its usual frozen state. Victor Baker of the University of Arizona has argued that the intensive volcanism in the Tharsis region periodically made early Mars quite a temperate place.

It is also very possible that none of these options is correct. We simply do not yet know enough about early Mars to have any real understanding of its climate. We must wait for future exploration. Unlike Earth, Mars has preserved much of its ancient landscape, which may yield clues to the conditions under which it formed. Indeed, understanding how Mars became so different from Earth will help geologists grapple with Earth's own history. The new lander missions will soon provide some of these clues.



"WHITE ROCK," seen by the Viking spacecraft in the 1970s (*inset*), is a prime example of how Mars's Earth-like appearance can deceive. The feature looks tantalizingly like a pile of salt deposited by liquid water. But spectral measurements now show it to be generic dust that somehow got compacted or cemented together. The reddish dust buried existing features, such as the crater at the top right, and was in turn buried by black sand. The image, taken by Mars Global Surveyor, attests to a bafflingly complex sequence of geologic events.

NASA/JPL/MALIN SPACE SCIENCE SYSTEMS MOC 2–264; CALVIN J. HAMILTON (*inset*)

MORE TO EXPLORE

Mars 2000. Arden L. Albee in *Annual Review of Earth and Planetary Sciences,* Vol. 28, pages 281–304; 2000.

Mars: The Lure of the Red Planet. William Sheehan and Stephen James O'Meara. Prometheus Books, 2001.

The Mars Global Surveyor Mars Orbiter Camera: Interplanetary Cruise through Primary Mission. Michael C. Malin and Kenneth S. Edgett in *Journal of Geophysical Research,* Vol. 106, No. E10, pages 23429–23570; October 25, 2001.

Nature Insight: Mars. Special section in *Nature,* Vol. 412, pages 207–253; July 12, 2001.

For information on NASA's Mars missions, visit **www.jpl.nasa.gov**

For information on the European Space Agency's Mars Express, visit **sci.esa.int/marsexpress**

For information on the Japanese Nozomi mission, visit **www.isas.ac.jp/e/enterp/missions/index.html**

ASTEROIDS HAVE BECOME NOTORIOUS AS CELESTIAL MENACES
BUT ARE BEST APPRECIATED IN A POSITIVE LIGHT, AS SURREAL WORLDS
BEARING TESTIMONY TO THE ORIGIN OF THE PLANETS

By Erik Asphaug

The Small Planets

Growing up in the Space Age, my friends and I would sometimes play the gravity game. One of us would shout, "Pretend you're on the moon!" and we'd all take the exaggerated slow strides we'd seen on television. "Pretend you're on Jupiter!" another would say, and we'd crawl on our hands and knees. But no one ever shouted, "Pretend you're on an asteroid!" In that pre-*Armageddon* era, who knew what "asteroid" meant? Now a grown-up who studies asteroids for a living, I still don't know how to respond.

Although we haven't seen any of the largest asteroids up close, they probably resemble shrunken, battered versions of the moon. In their weaker gravity, visiting astronauts would simply take longer strides. But below a few dozen kilometers in diameter, gravity is too feeble to press these so-called minor planets into even an approximately round shape. The smallest worlds instead take on a carnival of forms, resembling lizard heads, kidney beans, molars, peanuts and skulls. Because of their irregularity, gravity often tugs away from the center of mass; when added to the centrifugal forces induced by rotation, the result can seem absurd. Down might not be down. You could fall up a mountain. You could jump too high, never to return, or launch yourself into a chaotic (though majestically slow) orbit for days before landing at an unpredictable location. A pebble thrown forward might strike you on the head. A gentle vertical hop might land you 100 meters to your left or even shift the structure of the asteroid underfoot. Even the most catlike visitor would leave dust floating everywhere, a debris "atmosphere" remaining aloft for days or weeks.

These aspects of asteroid physics are no longer only theoretical curiosities or a game for children. Space missions such as the Near Earth Asteroid Rendezvous (NEAR), the first probe to go into orbit around a minor planet, are dramatically modernizing our perception of these baffling objects. But in spite of careful observations and the occasional proximity of these bodies to Earth, we know less about asteroids (and their relatives, the comets) than we knew about the moon at the dawn of space exploration. Minor planets exhibit a delicate interplay of minor forces, none of which can be readily ignored and none of which can be easily simulated in a laboratory on Earth. Are they solid inside, or aggregate assemblages? What minerals are they composed of? How do they survive collisions with other small bodies? Could a lander or astronaut negotiate an asteroid's weird surface?

Half-Baked Planets

MY GRADUATE STUDIES began during the first Bush administration, when asteroids were mere dots—a thousand points of light known to orbit primarily in a belt between Mars and Jupiter. A few lesser populations were known to swoop closer to Earth, and then there were comets in the Great Beyond. From periodic variations in color and brightness, asteroids were inferred to be irregular bodies ranging in size from a house to a country, rotating every several hours or days. More detailed properties were largely the stuff of scientific imagination.

Asteroids residing closer to Mars and Earth commonly have the spectra of rocky minerals mixed with iron, whereas asteroids on the Jupiter side are generally dark and red, suggesting a primitive composition only coarsely differentiated from that of the primordial nebula out of which the planets began to coalesce 4.56 billion years ago [*see illustration on page 46*]. This timing is precisely determined from analysis of lead isotopes—the products of the radioactive decay of uranium—in the oldest grains of the most primitive meteorites. In fact, meteorites have

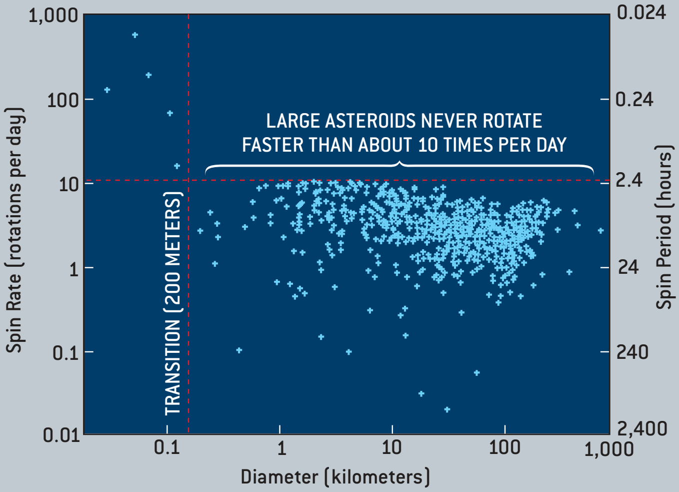
JOHNS HOPKINS UNIVERSITY APPLIED PHYSICS LABORATORY AND NASA

Updated from the May 2000 issue

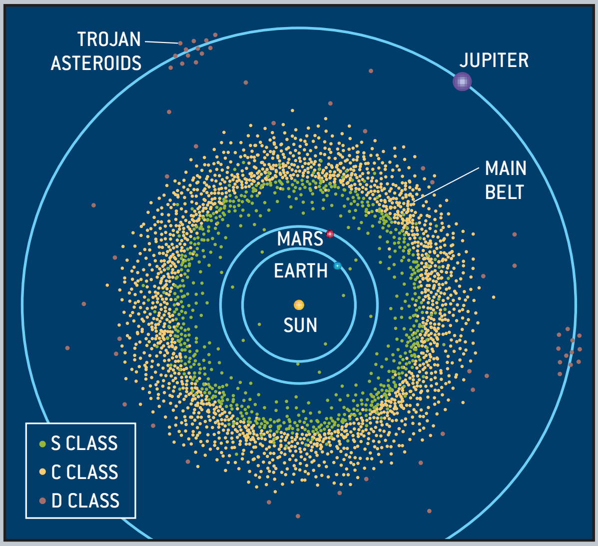
GIANT PAW PRINT is a strange crater (*center of lower image*) on the asteroid Eros, so dubbed by scientists now studying this 33-kilometer-long space rock with data from the NEAR space probe. On the other side of the body is a youthful, saddle-shaped gouge (*left of upper image*) full of unexplained markings. Through images such as these, asteroids are now turning from astronomical objects—mere points of light—into geologic objects—whole worlds whose exploration has only begun.

Types of Asteroids

Where They Roam



TWO GROUPS OF ASTEROIDS emerge on a plot of their rotation rates (*vertical axis*) versus size (*horizontal axis*). No known asteroid larger than 200 meters across rotates faster than once every 2.2 hours. The cutoff is easy to explain if these asteroids are piles of rubble that fly apart if spun too fast. Smaller asteroids, which can turn once every few minutes, must be solid rocks. The transition probably arose because of collisions.



MAIN ASTEROID BELT lies between the orbits of Mars and Jupiter, but stragglers cross Earth's orbit (and sometimes collide with Earth) or revolve in sync with Jupiter (in two groups known as the Trojan asteroids). The inner main belt consists mainly of stony or stony-iron asteroids (S class); farther out the asteroids are darker, redder and richer in carbon (C class and D class).

LAURIE GRACE; SOURCE: PETR PRAVEC Academy of Sciences of the Czech Republic (from sunkl.asu.cas.cz/~asteroid/3FRAs)

long been suspected to derive from asteroids. The spectra of certain meteorites nearly match the spectra of certain classes of asteroids. We therefore have pieces of asteroids in our possession.

Many astronomers used to think that telescope observations, combined with meteorite analysis, could substitute for spacecraft exploration of asteroids. Although the puzzles proved more stubborn than expected, researchers have been able to piece together a tentative outline of solar system history. For the planets to have accreted from a nebula of dust and gas, there had to be an initial stage in which the first tiny grains coagulated into growing bodies known as planetesimals. These became the building blocks of planets. But in the zone beyond Mars, gravitational resonances with massive Jupiter stirred the cauldron and prevented any body from growing larger than 1,000 kilometers across—leaving unaccreted remnants to become the present asteroids.

The largest of these would-be planets nonetheless accumulated enough internal heat to differentiate: their dense metals percolated inward, pooling and perhaps forming cores, leaving behind lighter rocky residues in their outer layers. Igneous activity further metamorphosed their rock types, and volcanoes erupted on some. Although none grew large enough to hold on to an atmosphere, hydrated minerals found in some meteorites reveal that liquid water was often present.

Encounters among the planetesimals became increasingly violent as Jupiter randomized the orientation and ellipticity of their orbits. Instead of continuing to grow, the would-be planets were chiseled or blasted apart by mutual collisions. Their pieces often continued to orbit the sun in families with common orbital characteristics and related spectra. Many current-day asteroids and meteorites are the rock- or metal-rich debris of these differentiated protoplanets. Other asteroids (and most comets) are more primitive bodies that for various reasons never differentiated. They are relics from the distant time before planets existed.

The Sky Is Falling

A DECADE AGO no asteroid had been imaged in any useful detail, and many astronomers had trouble taking them seriously. The first asteroids, discovered in the early 1800s, were named in the grand mythological manner. But with the tenth, the hundredth and the thousandth, asteroids began taking on the names of their discoverers and then of discoverers' spouses, benefactors, colleagues and dogs. Now, after a century of near-neglect, serious interest in asteroids is waxing as new observations transform them from dim twinkles in the sky into mind-boggling landforms. For this, asteroid scientists can thank former NASA administrator Daniel S. Goldin and the dinosaurs.

Goldin's "faster, better, cheaper" mantra has been a boon to asteroid science, because a visit to a tiny neighbor is both faster and cheaper than a mission to a major planet. The specter of fiery death from above has also focused minds. The

THE AUTHOR

ERIK ASPHAUG recalls his quarter-dropping days playing the video game Asteroids: "You get two big chunks and maybe two smaller chunks whenever you kill an asteroid. In truth you'll get hundreds of pieces filling your screen." Now an earth science professor at the University of California, Santa Cruz, Asphaug enjoys the ocean, his two sons, and simulating asteroid and planet collisions on supercomputers. Asphaug was awarded the 1998 Urey Prize of the American Astronomical Society.

discovery of the Chicxulub crater in the Yucatán vindicated the idea that the impact of an asteroid or comet 65 million years ago extinguished well over half the species on Earth.

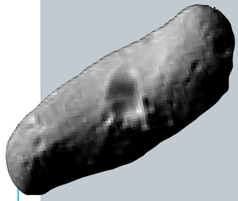
A repeat is only a matter of time, but when? Until we completely catalogue all significant near-Earth asteroids—a job we have just begun—we cannot say, but the chance in a given year is incredibly small. (We will never completely catalogue the comet hazard, because each comet visits the inner solar system so rarely.) None of us is remotely likely to die by asteroid impact, yet even scientists are drawn to the excitement of apocalypse, perhaps too often characterizing asteroids by their potential explosive yield in megatons instead of by diameter. Our professional dilemma is akin to notoriety in art: we want asteroids to be appreciated for higher reasons, but notoriety pays the bills.

Egged on by this nervous curiosity, we are entering the golden age of comet and asteroid exploration. Over a dozen have been imaged, and each new member of the menagerie is welcomed with delight and perplexity. They are not what we expected. Small asteroids were predicted to be hard and rocky, as any loose surface material (called regolith) generated by impacts was expected to escape their weak gravity. Aggregate small bodies were not thought to exist, because the slightest sustained relative motion would cause them to separate.

Reduced to Rubble

BUT OBSERVATIONS and modeling are proving otherwise. Most asteroids larger than a kilometer are now believed to be composites of smaller pieces. Those imaged at high resolution show evidence for copious regolith despite the weak gravity. Most of them have one or more extraordinarily large craters, some wider than the mean radius of the whole body. Such colossal impacts would not just gouge out a crater—they would break any monolithic body into pieces. Evidence of fragmentation also comes from measurements of asteroid bulk density. The values are improbably low, indicat-

Continued on page 51



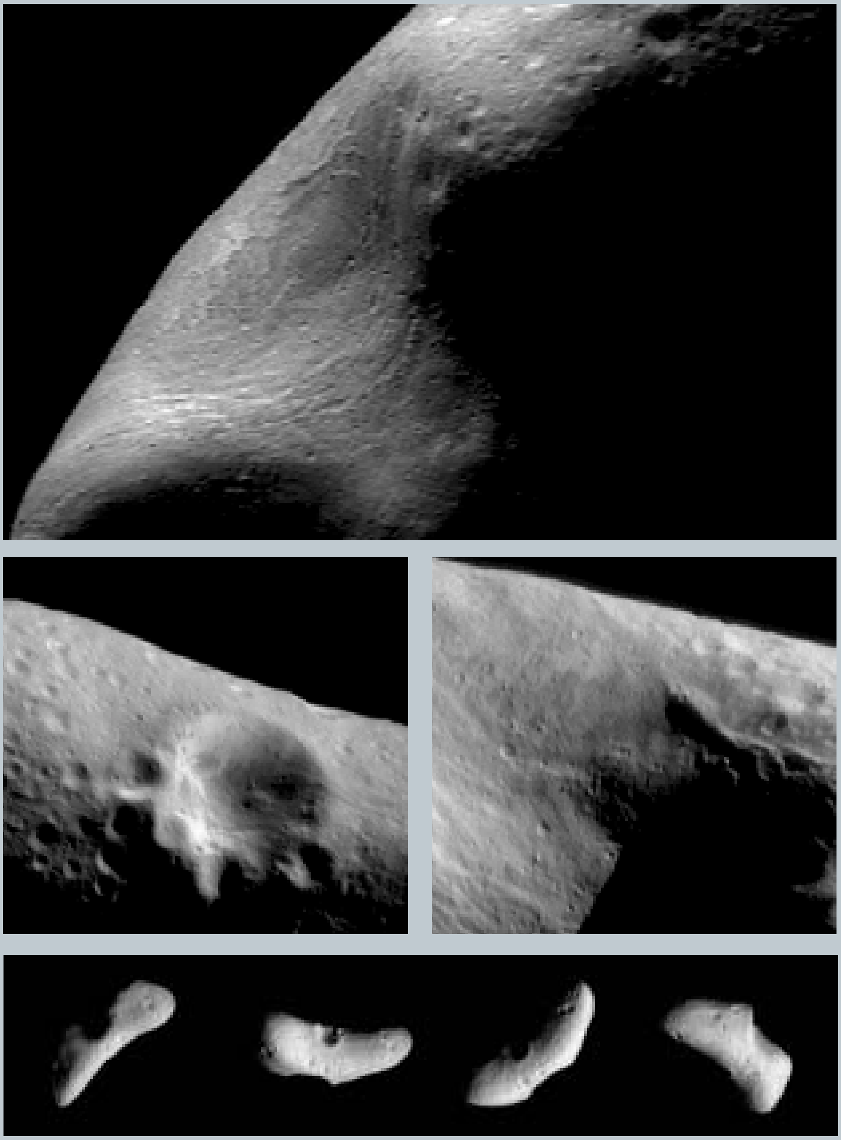
NEAR's Courtship with Eros

A Lovely Rock

EROS—orbited, and then visited, by the NEAR spacecraft—resembles a boat with a narrow bow, a wide stern and a prominent crater on the concave deck. Copious mounded and blocky debris around this crater show the influence of gravity during its formation. A boulder is inside, stopped halfway; it can't seem to figure out which way is down. Another prominent divot, on the opposite side, is so big that it is part of Eros' overall shape. If it is of impact origin, as is probable, its formation must have cracked Eros into a few great pieces mantled in lesser fragments and debris.

The name "Eros" befits a coy flirtation with Earth. Unfortunately, this love affair may end in sorrow. Paolo Farinella of the University of Trieste and Patrick Michel of Nice Observatory have calculated that Eros has a 5 percent chance of colliding with Earth in the next one billion years, with an intensity exceeding that which extinguished the dinosaurs.

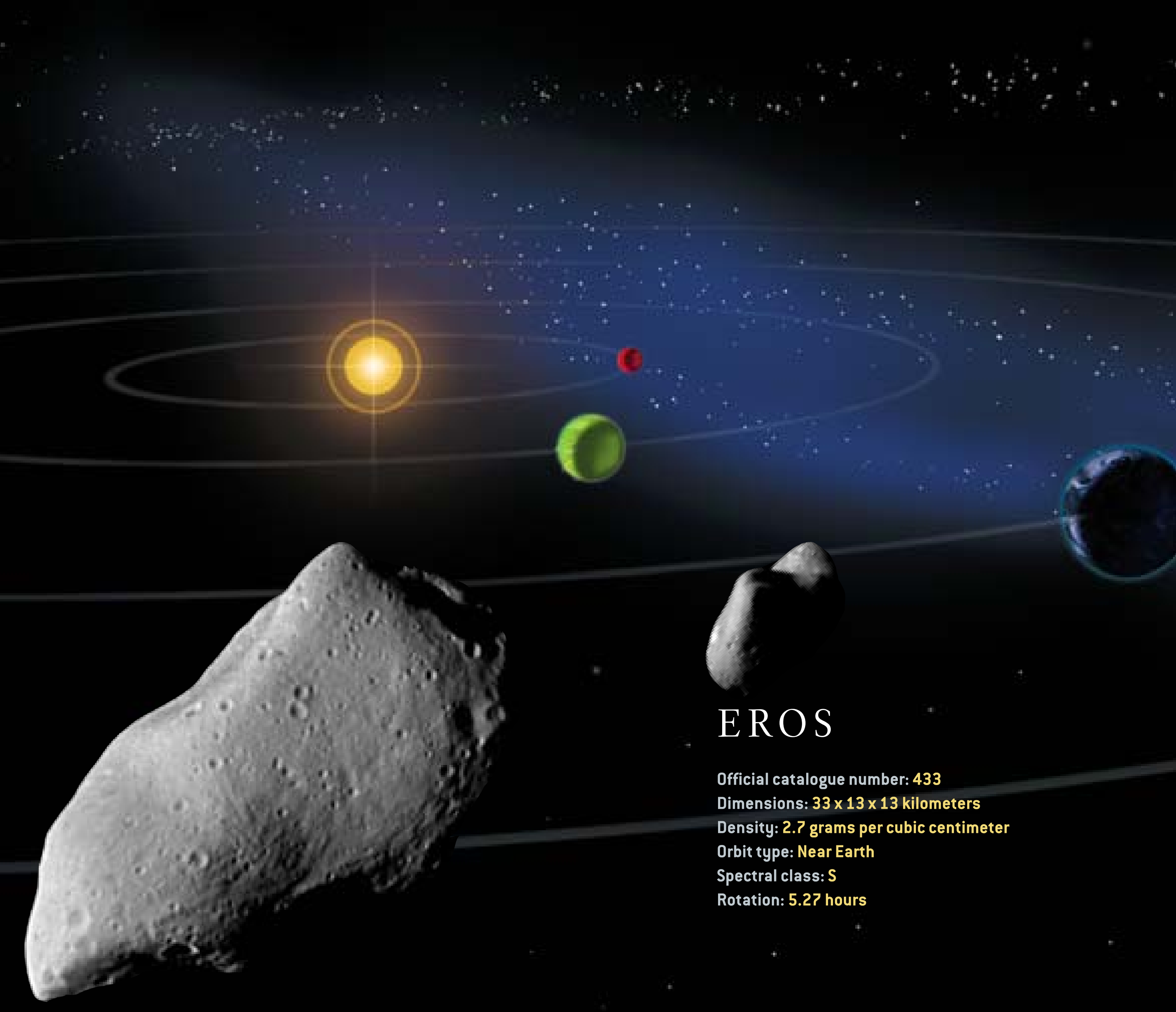
NEITHER SOLID ROCK NOR DUST BUNNY, Eros is a conglomerate of several major pieces crosscut by faults, scarps and ridges. The largest structure is a smooth, striated gouge that is nearly devoid of craters (*below*). The most prominent crater—the "paw print" six kilometers across—has massive deposits on its rim, which indicate that gravity dictated its formation (*center left*). A steep ridge, which runs parallel to the linear markings, suggests faulting in a coherent material (*center right*). The asteroid rotates once every five and a half hours (*bottom*).



JOHNS HOPKINS UNIVERSITY APL AND NASA

Asteroid Visitations

The Rock Concert



EROS

Official catalogue number: 433
Dimensions: 33 x 13 x 13 kilometers
Density: 2.7 grams per cubic centimeter
Orbit type: Near Earth
Spectral class: S
Rotation: 5.27 hours

GASPRA

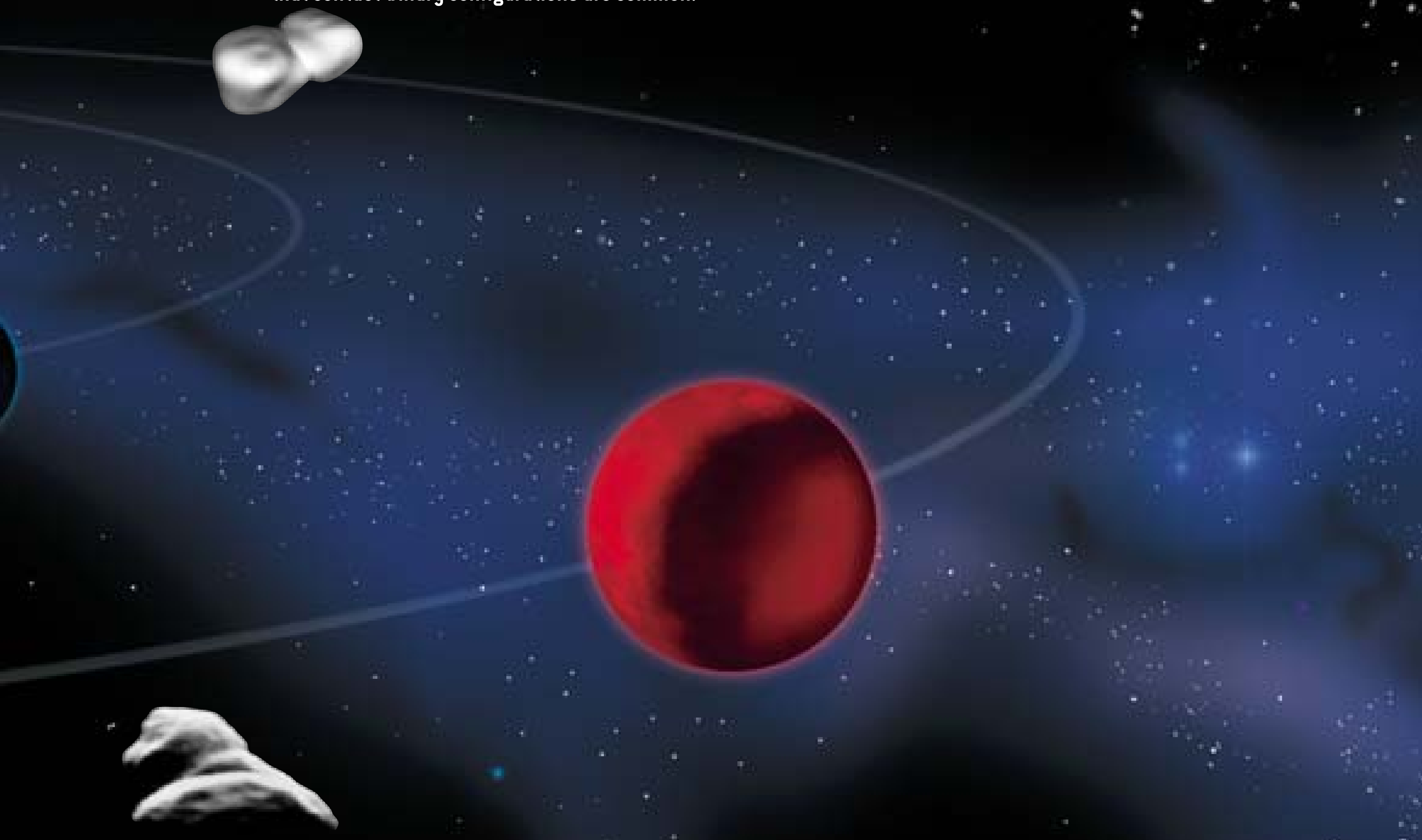
Official catalogue number: 951 **Dimensions:** 19 x 12 x 11 kilometers
Density: Not known **Orbit type:** Main belt (Flora family)
Spectral class: S **Rotation:** 7.04 hours

Gaspra was the first asteroid visited by a spacecraft: Galileo flew by in 1991 on its way to Jupiter. Some have argued that its six large concavities are not craters but facets formed when Gaspra broke off from its parent asteroid. On the other hand, in the weak, irregular gravity of Gaspra, the largest impact craters would naturally take on such a flat, lopsided shape.

CASTALIA

Official catalogue number: 4769 Dimensions: 1.8 x 0.8 kilometers
Density: 2.1 grams per cubic centimeter (surface) Orbit type: Earth crossing
Spectral class: S Rotation: 4 hours

Castalia was the first asteroid ever to be imaged. In August 1989 it flew within 11 lunar distances of Earth—still too far for optical telescopes but close enough for radar. Steven J. Ostro and his group at the Jet Propulsion Laboratory targeted the body with powerful and precise beams from the world's largest radio telescope, in Arecibo, Puerto Rico. Castalia's peanut shape—derived from radar echoes—suggests two 800-meter pieces resting together despite the very weak gravity. Radar echoes from other Earth-crossing asteroids now tell us that contact-binary configurations are common.



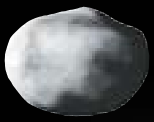
TOUTATIS

Official catalogue number: 4179 Dimensions: 4.5 by 2.4 by 1.9 kilometers
Density: 2.1 grams per cubic centimeter (surface) Orbit type: Earth crossing
Spectral class: S Rotation: Two separate periods (5.41 and 7.35 days)

Since the first observations of Castalia, better opportunities for radar detection of asteroids have presented themselves, most notably for asteroid Toutatis. Strongly influenced by Earth's gravity, its orbit is chaotic. Also, it wobbles with two types of motion that combine to create nonperiodic rotation. A visitor would never see the same horizon twice. On September 29, 2004, Toutatis will come within four lunar distances, whereupon it will be visible with binoculars.

VESTA

Official catalogue number: 4 Dimensions: 525 kilometers in diameter
Density: 3.3 grams per cubic centimeter Orbit type: Main belt
Spectral class: V Rotation: 5.34 hours



Only Vesta among the large asteroids has a surface of basaltic rock from ancient lava flows. In the distant past, it evidently differentiated into layers and underwent many of the same geologic processes as early Mars or Earth. Dozens of Vesta-like bodies presumably existed at one time but were broken apart into families of smaller asteroids. Iron meteorites are thought to come from the cores of these shattered worlds and igneous meteorites from their crusts and mantles. In 1996 the Hubble Space Telescope obtained this image of Vesta, showing a huge crater 430 kilometers across. Perhaps a billion years old, this crater might be the source of the small V-type asteroids observed today.

MATHILDE

Official catalogue number: 253 Dimensions: 66 x 48 x 46 kilometers
Density: 1.3 grams per cubic centimeter Orbit type: Main belt
Spectral class: C Rotation: 17.4 days

On its way to Eros, NEAR made the first spacecraft encounter with a primitive C-type asteroid. This blacker-than-coal spheroid, the largest asteroid yet visited and one of the slowest rotators, follows an eccentric orbit extending to the outer reaches of the main belt. The spacecraft's trajectory was slightly deflected by Mathilde, telling us its mass. The implied density is less than half that of the closest matching meteorites, carbonaceous chondrites, so if Mathilde is made of the same material, it must be very loosely packed. The same is true of another C-type asteroid, Eugenia, recently studied using a ground-based telescope equipped with sophisticated adaptive optics.

The giant craters are amazing. Several are wider than Mathilde's average radius, yet none have rims or ejecta deposits, which are associated with large craters on other worlds. Also, none of the craters have been degraded by subsequent cratering; we can't even tell which impact happened first or last. It is as though some deity has come and taken tidy bites from a cosmic apple.

IDA & DACTYL

Official catalogue number: 243
Dimensions: 56 x 24 x 21 kilometers Density: About 2.5 grams per cubic centimeter
Orbit type: Main belt (Koronis family) Spectral class: S Rotation: 4.63 hours

Two years after visiting Gaspra, Galileo flew by this main-belt asteroid. An unexpected gem from this encounter was the discovery of Dactyl—the first known asteroid satellite, a mere 1.4 kilometers in diameter. The Galileo team used Dactyl's orbit to calculate Ida's mass. The implied density is much lower than that of the most closely related type of meteorite, ordinary chondrites, so Ida must be of different composition or else porous. Some believe that Dactyl  agglomerated out of slow ejecta from one of Ida's largest craters, although this would have been very difficult to achieve dynamically. Daniel Durda of the Southwest Research Institute in Boulder, Colo., showed that Dactyl and Ida could have formed as a pair a billion or more years ago, when Ida's parent body was disrupted. But it is hard to explain how Dactyl could have survived for so long without being destroyed.

Continued from page 47

ing that these bodies are threaded with voids of unknown size.

In short, asteroids larger than a kilometer across may look like nuggets of hard rock but are more likely to be aggregate assemblages—or even piles of loose rubble so fragmented that no bedrock is left. This rubble-pile hypothesis was first proposed two decades ago by Don Davis and Clark Chapman, both then at the Planetary Science Institute in Tucson, but they did not suspect that it would apply to such small diameters.

Shortly after the NEAR spacecraft flew by asteroid Mathilde on its way to Eros, the late planetologist Eugene M. Shoemaker (for whom NEAR has been renamed) realized that the huge craters on this asteroid and its very low density could only make sense together: a porous body such as a rubble pile can withstand a battering much better than an integral object. It will absorb and dissipate a large fraction of the energy of an impact; the far side might hardly feel a thing. A fair analogy is a bullet hitting a sandbag, as opposed to a crystal vase.

What about the jagged shapes of most asteroids? No regional slope on any imaged asteroid or comet exceeds a typical angle of repose (about 45 degrees), the incline at which loose debris tumbles down. In the steepest regions, we do see debris slides. So, small bodies could as well be made of boulders or even sand and still hold their shape. Dunes, after all, have distinct ridges yet are hardly monolithic. Rapid rotation would contribute to an elongated, lumpy appearance for a rubble pile.

Direct support for the rubble-pile hypothesis emerged in 1992, when Comet Shoemaker-Levy 9 strayed too close to Jupiter and was torn into two dozen pieces. Two years later this "string of pearls" collided with the giant planet. According to a model I developed with Willy Benz of the University of Bern, the comet could have disassembled as it did only if it consisted of hundreds of loose grains in a slow cosmic landslide. As the comet was stretched by Jupiter's tides, the grains gravitated into clumps much like water beading in a fountain. From this breakup we proposed that comets are likely to be granular structures with a density around two thirds that of water ice. What applies to comets might apply to asteroids as well.

When Everything Matters

YET THE RUBBLE-PILE hypothesis is conceptually troublesome. The material strength of an asteroid is nearly zero, and gravity is so low we are tempted to neglect that, too. What's left? The truth is that neither strength nor gravity can be ignored. Paltry though it may be, gravity binds a rubble pile together. And anyone who builds sand castles knows that even loose debris can cohere. Oft-ignored details of motion begin to matter: sliding friction, chemical bonding, damping of kinetic energy, electrostatic attraction and so on. We are just beginning to fathom the subtle interplay of these minuscule forces.

The size of an asteroid should determine which force dominates. One indication is rotation rate. Some collisions cause an asteroid to spin faster; others slow it down. If asteroids are monolithic rocks undergoing random collisions, a graph of their rotation rates should show a bell-shaped distribution with a statistical "tail" of very fast rotators. If nearly all asteroids are rubble piles, however, this tail would be missing, because any rubble pile spinning faster than once every two or three hours (depending on bulk density) would fly apart. Alan Harris of the Jet Propulsion Laboratory in Pasadena, Calif., Petr Pravec of the Academy of Sciences of the Czech Republic in Prague and their colleagues have discovered that all but five observed asteroids obey a strict rotation limit [*see illustration on page* 46]. The exceptions are all smaller than about 150 meters in diameter, with an abrupt cutoff for asteroids larger than about 200 meters.

The evident conclusion—that asteroids larger than 200 meters across are multicomponent structures or rubble piles—agrees with recent computer modeling of collisions, which also finds a transition at that diameter. A collision can blast a large asteroid to bits, but those bits will usually be moving slower than their mutual escape velocity (which, as a rule of thumb, is about one meter per second, per kilometer of radius). Over several hours, gravity will reassemble all but the fastest pieces into a newer rubble pile [*see illustration on next page*]. Because collisions among asteroids are relatively frequent, most large bodies have already suffered this fate. Conversely, most small asteroids should be monolithic, because impact fragments easily escape their feeble gravity.

Qualitatively, a "small" asteroid sustains dramatic topography, and its impact craters do not retain the debris they eject. A "large" asteroid is an assemblage of smaller pieces that gravity and random collisions might nudge into a rounded or, if spinning, an elongated shape. Its craters will have raised rims and ejecta deposits, and its surface will be covered in regolith. But this distinction is not straightforward. Asteroid Mathilde could be considered small, as it has no visible rims or ejecta around its enormous craters, or large, as it is approximately spheroidal. The ambiguity is a sign that the underlying science is uncertain.

Shock Value

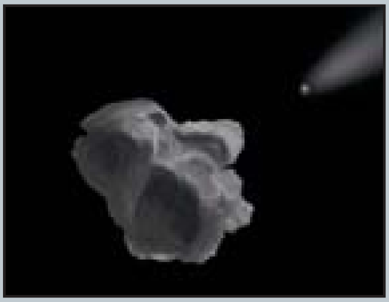
GIVEN THAT GEOPHYSICISTS are still figuring out how sand behaves on Earth and how landslides flow, we must be humble in trying to understand conglomerate asteroids. Two approaches are making inroads into one of their key attributes: how they respond to collisions.

Derek Richardson and his colleagues at the University of Washington simulate asteroids as piles of discrete spheres. Like cosmic billiards on a warped pool table—the warp being gravity—these spheres can hit one another, rebound and slow down because of friction and other forms of energy dissipation. If balls have enough collisional energy, they disperse; more commonly, some or all pile back together. Richardson's model is useful for studying the gentle accretionary encounters in the early solar system, before relative velocities started to increase under the gravitational influence of nascent Jupiter. It turns out to be difficult for planetesimals to accrete mass during even the most gentle collisions.

SCOTT HUDSON *Washington State University* [*Castalia*]; BEN H. ZELLNER *Georgia Southern University* AND NASA [*Vesta*];
JOHNS HOPKINS UNIVERSITY APL AND NASA [*Eros and Mathilde*]; NASA/JET PROPULSION LABORATORY [*Gaspra*]; NASA/JPL [*Ida*]

Flintstone or Rubble?

Really Deep Impacts



RUBBLE-PILE ASTEROID, whose fragmented structure bears the scars of collisions, is struck again by a smaller asteroid at high speed. Such bang-ups are common.

In the aggregate body the blast remains confined to the local area. Within a few minutes, the smallest, fastest debris has escaped. Larger fragments drift outward.

Some large pieces escape; some return. A few days later things settle. Over time the wound will be covered in debris from bombardment and other processes.

SOLID ASTEROID responds very differently to a collision than the rubble pile does—just as a log responds differently to the blow of an ax than a mound of wood chips does.

The shock wave propagates deep, blasting apart the body. The fastest ejecta are soon gone, leaving larger fragments to undergo a gentle gravitational dance for hours.

Many of these pieces come to rest in a pile of rubble. Because it is so easy to turn a rock into a rubble pile, few asteroids larger than a few hundred meters across are still solid.

BRYAN CHRISTIE DESIGN

High-speed collisions, more typical of the past four billion years, are complicated because they involve the minutiae of material characteristics such as strength, brittle fracture, phase transformations, and the generation and propagation of shock waves. Benz and I have developed new computational techniques to deal with this case. Rather than divide an asteroid into discrete spheres, we treat it as a continuous body, albeit with layers, cracks, or networks of voids.

In one sample simulation, we watch a 6,000-ton impactor hit billion-ton Castalia at five kilometers per second. This collision releases 17 kilotons of energy, the equivalent of the Hiroshima explosion—and enough to break up Castalia. We simulate Castalia as a two-piece object held together by gravity. The projectile and an equal mass of Castalia are vaporized in milliseconds, and a powerful shock wave is spawned. Because the wave cannot propagate through vacuum, it rebounds off surfaces, including the fracture between the two pieces of the asteroid. Consequently, the far piece avoids damage. The near piece cracks into dozens of major fragments, which take hours to disperse; the largest ones eventually reassemble. This outcome is very sensitive to what we start with. Other initial configurations and material parameters lead to vastly different outcomes. Asteroids that start off as rubble piles, for example, are hard to blast apart.

Another surprise is the asteroid pair, which theory suggested was needed to explain why 15 percent of the impact craters on Venus, Earth, the moon and Mars are doublets. The Galileo spacecraft's discovery of Dactyl orbiting Ida indicated that the idea had merit. Advanced imaging of the solar system in recent years has proved the notion correct. The menagerie of asteroids is sublime with dancing pairs, such as 1999 KW4 (the pair gets a name; the individuals do not). I delight in thinking about floating between the duo, the larger a kilometer across, separated by a few hundred meters. A gentle hop would do it (escape velocity is one meter per second, so watch out!) and would occupy a few tranquil minutes. One might then enjoy a relaxing view from the smaller body before hopping back home for a glass of Chablis. There could be trouble, though: 1999 KW4 is a Potentially Hazardous Asteroid—its most likely fate is to collide with us at an unknown date at least a few thousand years from now. I anticipate that humans, if we stick around, will have figured out how to give it a gentle nudge by then.

Rendezvous with Eros

THE EXISTENCE of abundant asteroid satellites supports the rubble-pile hypothesis, because we can only conceive of these systems forming in the inner solar system as disruption remnants from near-catastrophic collisions or in the aftermath of a tidal encounter, where loosely bound asteroids get ripped apart by tidal forces. Such was the fate that befell Comet Shoemaker-Levy 9. Still, this view puts a lot of weight in a theory, and theories are a dime a dozen in this business. How else can we know the structures of asteroids?

We can infer the rock properties of an asteroid by trying out different initial guesses and comparing simulations with observations. As an example, I have worked with Peter Thomas of Cornell University to re-create the largest crater on Mathilde: its diameter and shape (easy enough), its lack of fracture grooves or damage to existing craters (somewhat harder), and the absence of crater ejecta deposits (very hard).

If we assume that Mathilde was originally solid and monolithic, our model can reproduce the crater but predicts that the asteroid would have cracked into dozens of pieces, contrary to observations. If we assume that Mathilde was originally a rubble pile, then our impact model easily matches the observations. Kevin Housen of the Boeing Shock Physics Lab and his colleagues have also argued that Mathilde is a rubble pile, although they regard the craters as compaction pits—like dents in a beanbag—rather than excavated features.

Understanding asteroid structure will be crucial for future missions. A rubble pile will not respond like a chunk of rock if we hope to gather material for a sample return to Earth or, in the more distant future, construct remote telescopes, conduct mining operations or attempt to divert a doomsday asteroid headed for Earth. The irregular gravity is also a problem; spacecraft orbits around comets and asteroids can be chaotic, making it difficult to avoid crashing into the surface, let alone point cameras and instruments. NEAR therefore conducted most of its science 100 kilometers or more away from Eros. At this distance the irregular, rapidly rotating potato exerts almost the same gravity that a sphere would.

Orbiting Eros at the speed of a casual bicyclist (corresponding to the low gravity), NEAR beamed a torrent of data toward Earth. The primary objective was to clarify the link between asteroids and meteorites. Cameras mapped the body to a few meters' resolution, spectrometers analyzed the mineral composition, and a magnetometer searched for a native magnetic field and for interactions with the solar field. Upcoming missions will probe asteroids and comets in ever greater detail, using a broader range of instruments such as landers, penetrators and sample returns.

These discoveries will help plug a vast conceptual hole in astronomy. We simply don't understand small planetary bodies, where gravity and strength compete on sometimes equal footing. Asteroids are a balancing act, as serene as the moon yet of cataclysmic potential, large enough to hang onto their pieces yet too small to lose their exotic shape. Neither rocks nor planets, they are something of Earth and Heaven.

Upcoming Missions

The NEAR Future

BEFORE ANY OF US can set foot on an asteroid, minor planets must be poked and prodded just as the moon was before Apollo 11 could land. To this end, several new missions will follow up the concluded Near Earth Asteroid Rendezvous.

Two will collect samples and bring them back to Earth. NASA's Stardust spacecraft was launched in February 1999 toward Comet Wild 2 and is expected to return in 2006 with a piece of the tail (some grains of precious dust). In May 2003 the Japanese space agency launched its Hayabusa space probe to collect material from the small asteroid 1998 SF36. In part a technology demonstration mission, Hayabusa will collect several grams of surface rock by shooting bullets into the body and capturing the ejecta.

JAPAN'S HAYABUSA SPACECRAFT will return samples from the small Earth-approaching asteroid 1998 SF36.

A. IKESHITA/MEF/ISAS

In early 2004 the European Space Agency will launch its Rosetta spacecraft for a 2014 rendezvous with Comet Churyumov-Gerasimenko. This billion-dollar mission is designed to determine the comet's interior structure by performing radio transmission experiments between a lander and an orbiter. Under NASA's Discovery program, an orbiter dubbed Dawn will launch in 2006 and spend the years 2010 and 2011 mapping the asteroid Vesta. It will then head for Ceres in 2014–2015. These large asteroids are believed to be relics from the time when planets were accreting. Their geologies will also reveal how tectonics, volcanism and impact mechanics operate in low-gravitational environments.

The goal of NASA's Deep Impact mission is to blow a crater in Comet Tempel 1 on July 4, 2005, using a 370-kilogram copper projectile. Possibilities range from a lunarlike crater the size of a sports arena to a carrot-shaped hole puncturing deep into the comet; the results should elucidate the unknown properties of comets.

Human visitation to an asteroid could even be in the offing. Veteran NASA astronaut Tom Jones and his colleagues have designed a tour that begins with a modified capsule docked to the International Space Station. Several astronauts would take a 60-day "vacation," spending a few weeks getting to a near-Earth asteroid, a few weeks studying the body from a safe distance and perhaps taking a few exploratory space walks. The journey would be a small step (and maybe a few new bootprints) toward Mars. —*E.A.*

MORE TO EXPLORE

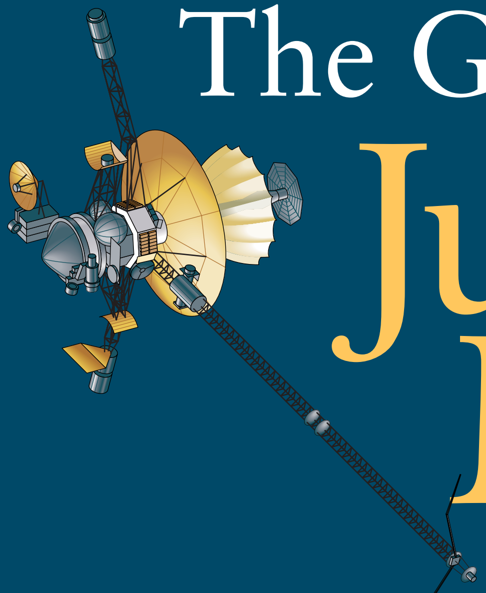
Disruption of Kilometre-Sized Asteroids by Energetic Collisions. Erik Asphaug, Steven J. Ostro, R. S. Hudson, D. J. Scheeres and Willy Benz in *Nature,* Vol. 393, pages 437–440; June 4, 1998.

Asteroids III. Edited by W. F. Bottke, Jr., et al. University of Arizona Press, 2002.

For updates on the Near Earth Asteroid Rendezvous mission, visit **http://near.jhuapl.edu**

For general information on near-Earth objects, go to **http://neo.jpl.nasa.gov**

The author's Web site is at **http://es.ucsc.edu/~asphaug**



The Galileo Mission to Jupiter and Its Moons

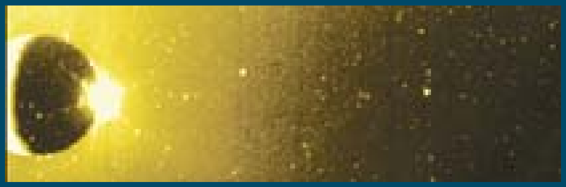
GALILEO SPACECRAFT, beset by technical troubles, still conducted a comprehensive study of the **JOVIAN SYSTEM.** Few predicted that the innards of these worlds would prove so varied

By Torrence V. Johnson

To conserve power, the probe was traveling in radio silence, with only a small clock counting down the seconds. Racing 215,000 kilometers overhead, its companion spacecraft was ready to receive its transmissions. Back on Earth, engineers and scientists, many of whom had spent most of two decades involved in the project, awaited two key signals. The first was a single data bit, a simple yes or no indicating whether the little probe had survived its fiery plunge into Jupiter's massive atmosphere.

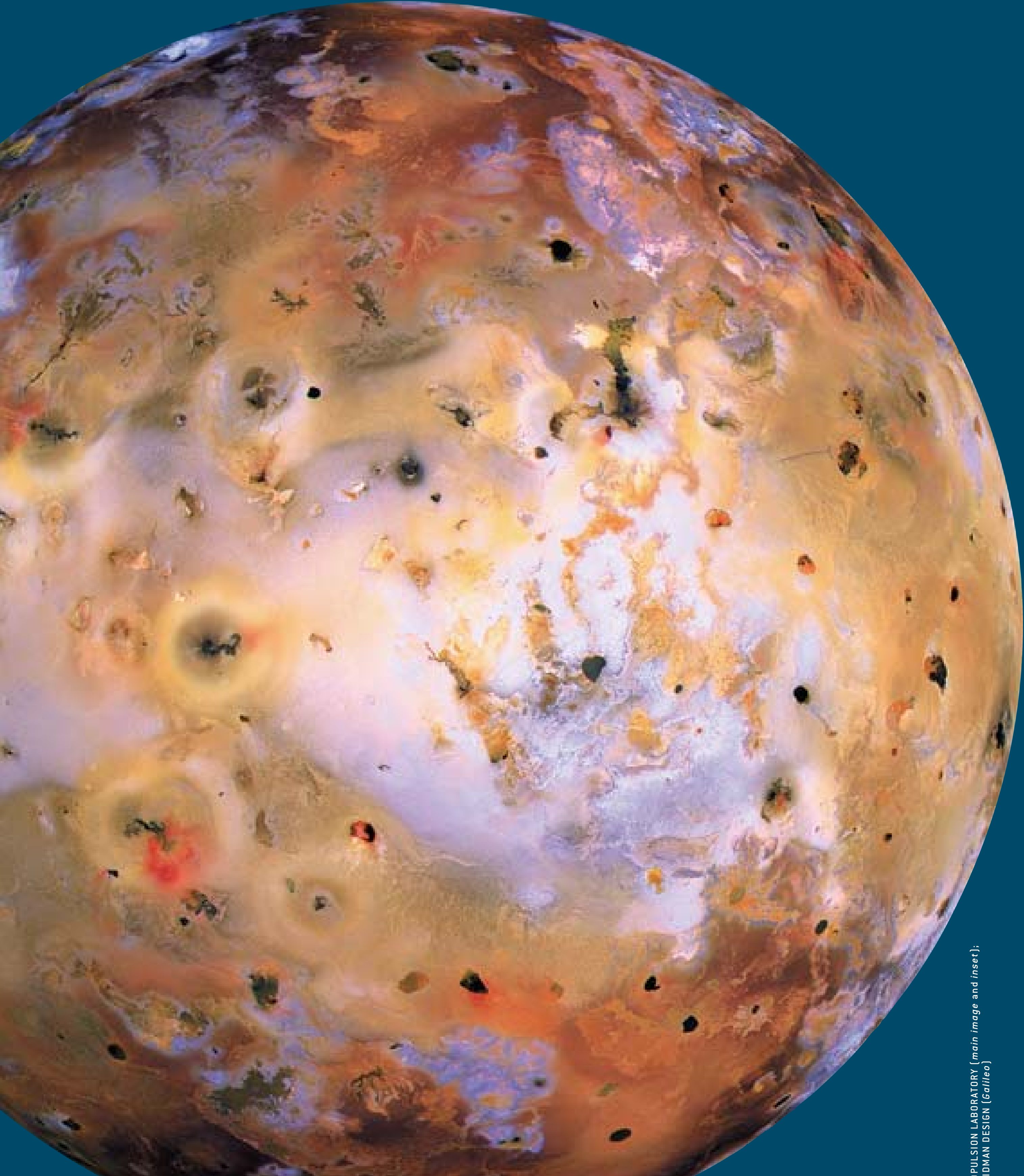
Getting this far had not been easy for the Galileo mission. When conceived in the mid-1970s, the two-part unmanned spacecraft was supposed to set forth in 1982, carried into Earth orbit on board the space shuttle and sent onward to Jupiter by a special upper rocket stage. But slips in the first shuttle launches and problems with upper-stage development kept pushing the schedule back. Then came the *Challenger* tragedy in 1986, which occurred just as Galileo was being readied for launch. Forced by the circumstances to switch to a safer but weaker upper stage, engineers had to plot a harrowing gravity-assist trajectory, using close flybys of Venus and Earth to provide the boost the new rocket could not. From launch in October 1989, the journey took six years. Two years into the flight, disaster struck again when the umbrellalike main communications antenna refused to unfurl, leaving the spacecraft with only its low-capacity backup antenna. Later, the tape recorder—vital for storing data—got stuck.

When engineers finally received the "golden bit" confirming that the probe was still alive,



WRACKED BY MORE THAN 100 VOLCANOES, the surface of Io makes Earth look geologically inert by comparison. The yellow, brown and red patches on this false-color mosaic (*main image*) represent different sulfur-based minerals—in other words, brimstone. A sulfur dioxide frost coats the white areas. Gas and dust have been swept into orbit, as is evident when Io is backlit by the sun (*inset at right*). Much of the yellowish glow comes from sodium gas. The burst of white light is sunlight scattered by the plume of the volcano Prometheus.

Updated from the February 2000 issue



NASA/JET PROPULSION LABORATORY (*main image* and *inset*);
JARED SCHNEIDMAN DESIGN (*Galileo*)

The Gas Giant
Jupiter

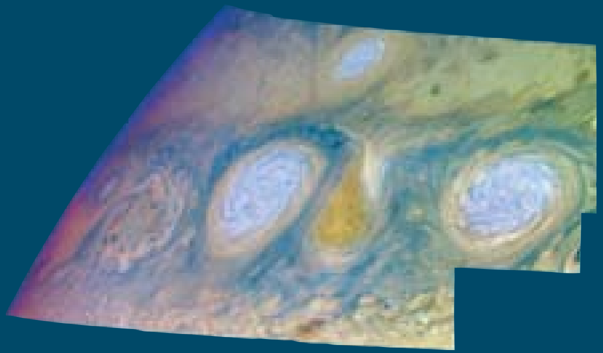

RETA BEEBE *New Mexico State University* AND NASA

Until the Galileo mission, no man-made object had ever made contact with a gas giant planet. The spacecraft dropped a probe into the atmosphere just north of the equator, a location shown (*arrow*) on this Hubble Space Telescope image taken after the probe had been targeted (*left*). The probe descended for more than an hour, measuring the composition (*table below*) before succumbing to the rising temperature and pressure (*sequence at right*). The primordial solar composition is assumed to be the same as that of the outer layers of the sun.

CHEMICAL COMPOSITION OF UPPER ATMOSPHERE
(Number of atoms per atom of hydrogen)

ELEMENT	CHEMICAL FORM	JUPITER	SATURN	SUN
HELIUM	HELIUM	0.078	0.070 ± 0.015	0.097
CARBON	METHANE	1.0×10^{-3}	2×10^{-3}	3.6×10^{-4}
NITROGEN	AMMONIA	4.0×10^{-4}	$3 \pm 1 \times 10^{-4}$	1.1×10^{-4}
OXYGEN	WATER	3.0×10^{-4}	unmeasured	8.5×10^{-4}
SULFUR	HYDROGEN SULFIDE	4.0×10^{-5}	unmeasured	1.6×10^{-5}
DEUTERIUM	DEUTERIUM	3×10^{-5}	3×10^{-5}	3.0×10^{-5}
NEON	NEON	1.1×10^{-5}	unmeasured	1.1×10^{-4}
ARGON	ARGON	7.5×10^{-6}	unmeasured	3.0×10^{-6}
KRYPTON	KRYPTON	2.5×10^{-9}	unmeasured	9.2×10^{-10}
XENON	XENON	1.1×10^{-10}	unmeasured	4.4×10^{-11}

SOURCES: SUSHIL K. ATREYA *University of Michigan*; HASSO B. NIEMANN *NASA Goddard Space Flight Center*

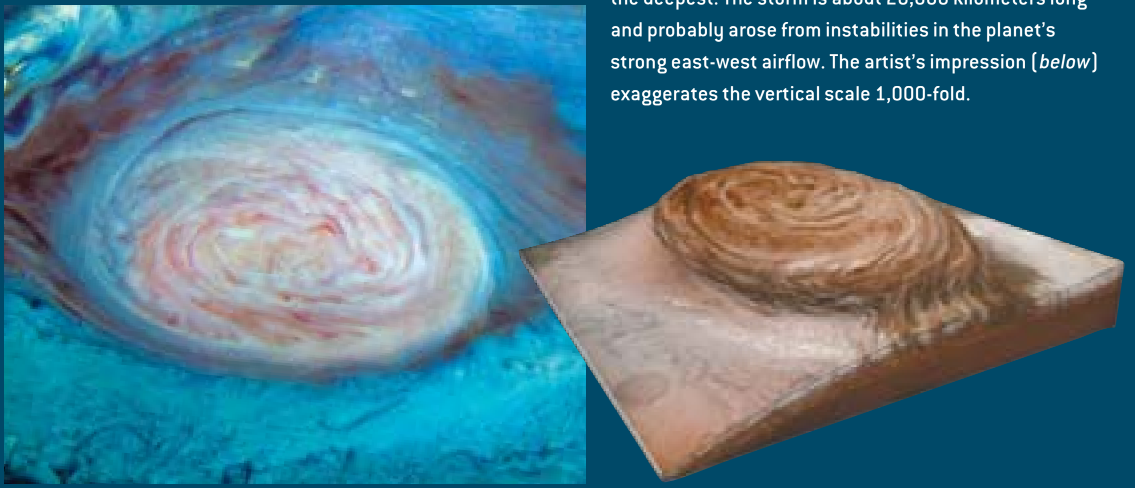

NASA/JPL

OVAL CLOUDS were seen by Galileo in early 1997. They have trapped a pear-shaped region between them. The ovals rotate counterclockwise; the pear-shaped region, clockwise. On this false-color mosaic of three near-infrared images, bluish clouds are thin, white ones are thick, and reddish ones are deep. A year later the ovals merged together—a vivid example of Jupiter's dynamic weather. Each oval is about 9,000 kilometers across.


NASA/JPL

HOLE IN THE UPPER CLOUD DECK reveals the comparatively warm regions deeper down. As on other near-infrared images, bluish clouds are thin, white ones are thick, and reddish ones are deep (*diagram at right*). The Galileo probe entered just such an area, known as a hot spot. This image depicts an area 34,000 kilometers across.

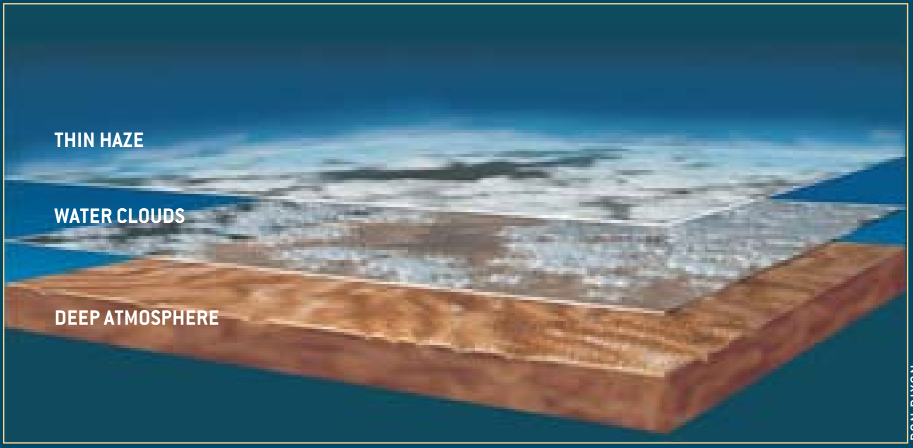
GREAT RED SPOT is a vast storm system that towers some 30 kilometers above the surrounding clouds (*left*). From red to green to blue, the color coding is decreasingly sensitive to the amount of methane along the line of sight. Consequently, the pink and white areas are highest, and bluish and black areas the deepest. The storm is about 26,000 kilometers long and probably arose from instabilities in the planet's strong east-west airflow. The artist's impression (*below*) exaggerates the vertical scale 1,000-fold.


NASA/JPL

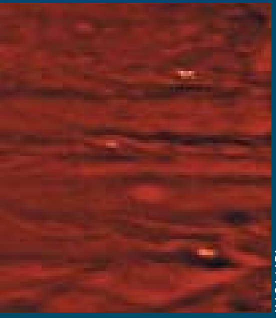
DON DIXON

LIGHTNING FLASHES appear in these orbiter images of the night side of Jupiter. Moonlight from Io dimly illuminates the ammonia cloud deck. The flashes probably originate from water clouds 100 kilometers deeper. Lightning strikes at about the same rate as in thunderstorms on Earth but 1,000 times more brightly. Each image shows an area roughly 60,000 kilometers square.

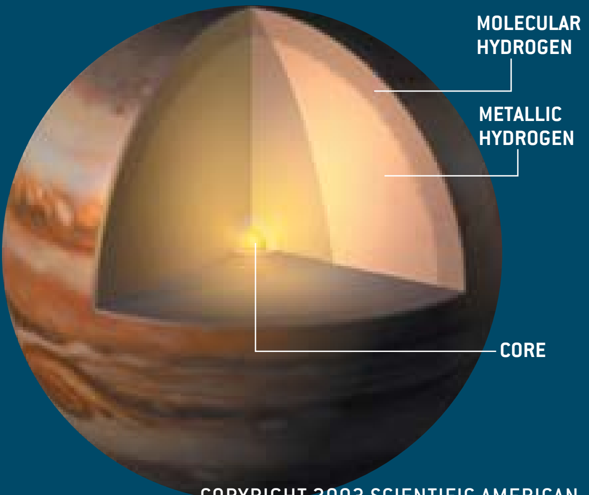
DON DAVIS (*capsule*); DON DIXON



DON DIXON



NASA/JPL



INTERIOR OF JUPITER shows that the term "gas giant" is something of a misnomer. The bulk of the planet consists of hydrogen under such immense pressures that it has become liquid and metallic. Underneath it all is a core of rock around which the hydrogen accumulated.

ALFRED T. KAMAJIAN

cheers went up in the control room and the tension began to ease. But the team still had to wait out the next two hours for the second critical event: insertion of the companion spacecraft into orbit. To slow it from interplanetary cruise enough for Jupiter's gravity to capture it, engineers instructed the German-built main engine to fire for 45 minutes. Ultimately, word came through that this maneuver had succeeded. The orbiter had become the first known artificial satellite of the giant planet.

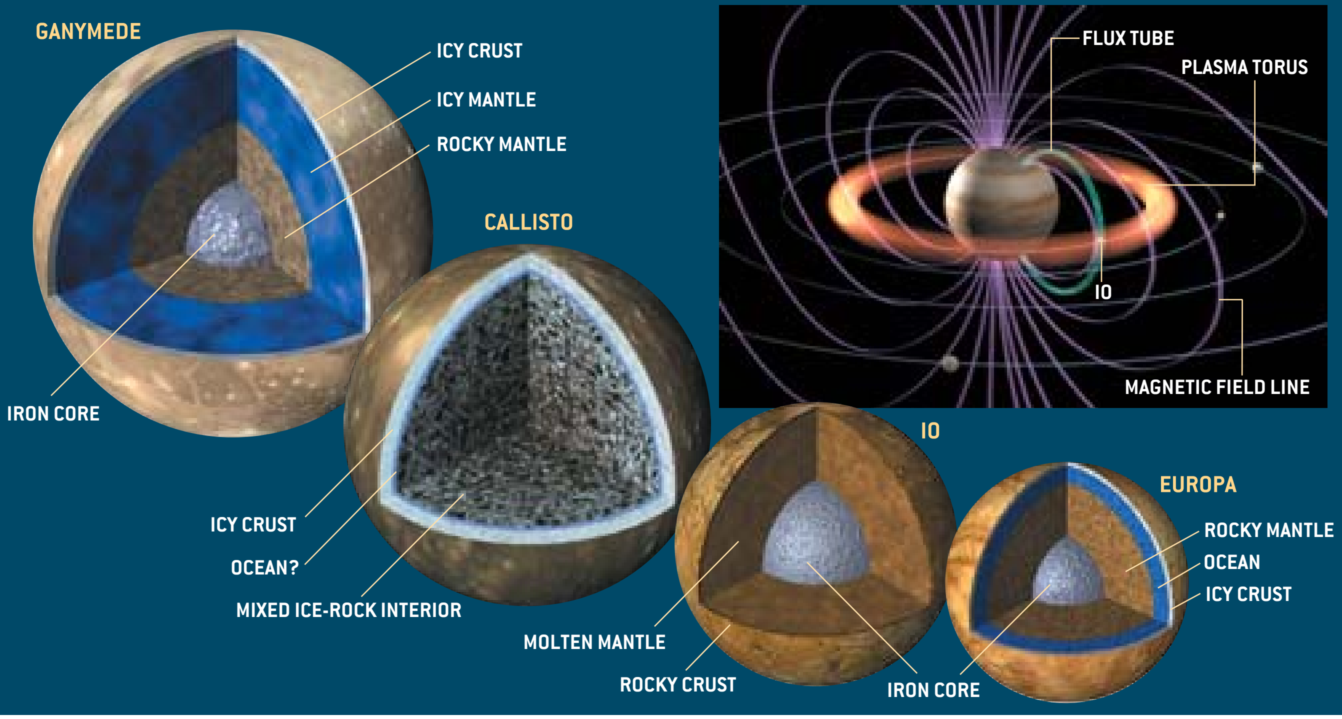
Since that day in December 1995, a mission that once seemed doomed has given researchers their first detailed view of the Jovian system, revealed only fleetingly in the Pioneer and Voyager flybys of the 1970s. The atmospheric probe penetrated the kaleidoscopic clouds and conducted the first in situ sampling of an outer planet's atmosphere, transmitting data for an hour before it was lost in the gaseous depths. The orbiter is still going strong. It has photographed and analyzed the planet, its rings and its diverse moons. Most famously, it has bolstered the case that an ocean of liquid water lurks inside Europa, one of the four natural satellites discovered by Galileo Galilei in 1610 [see "The Hidden Ocean of Europa," on page 64]. But the other large moons have revealed surprises of their own: beams of electrons that connect Io, the most volcanically tormented body in the solar system, to Jupiter; a magnetic field generated within Ganymede, the first such field ever discovered on a moon; and the subtle mysteries of Callisto, including signs that it, too, has an ocean.

Mother of All Downdrafts

ACCORDING TO modern theories of planet formation, Jupiter and the other giant planets emerged from the primordial solar nebula in two stages. First, icy planetesimals—essentially large comets that had condensed out of the cloud of gas and dust—clumped together. Then, as the protoplanet grew to a critical size, it swept up gas directly from the nebula. Jupiter thus started off with a sample of the raw material of the solar system, which had roughly the same composition as the early sun. Since then, the

The Interiors and Magnetic Fields of
Galilean Satellites

The four Galilean satellites of Jupiter do not really deserve to be called "moons." In many ways, they are planets in their own right. The inner two, Io and Europa, are about the size and density of Earth's moon. The outer two, Ganymede and Callisto, are about the size of Mercury but much less dense. Although Galileo did not land on or dig into them, it inferred their interior structure from their gravitational forces and magnetic fields. Of the four, only Callisto does not seem to have differentiated into distinct layers of metal, rock and water ice. Jupiter's electromagnetic fields interact with all four, but especially with Io (*diagram below*). The fields scoop up ionized gases from Io's volcanic eruptions, creating a torus of plasma. A flux tube between the planet and moons carries an electric current of five million amperes. (In this diagram, the planet and moons are not to scale.)



planet has been shaped by processes such as internal differentiation and the continuing infall of comets. Disentangling these processes was the main goal of the atmospheric probe.

Perhaps the most mysterious discovery by the probe involved the so-called condensable species, including elements such as nitrogen, sulfur, oxygen and carbon. Scientists have long known that Jupiter has about three times as much carbon (in the form of methane gas) as the sun. The other species (in the form of ammonia, ammonia sulfides and water) are thought to condense and form cloud layers at various depths. Impurities in the cloud droplets, possibly sulfur or phosphorus, give each layer a distinctive color. The probe was designed to descend below the lowest expected cloud deck, believed to be a water cloud at about 5 to 10 atmospheres of pressure—some 100 kilometers below the upper ammonia ice clouds. The expected weather report was windy, cloudy, hot and humid.

Yet the instruments saw almost no evidence for clouds, detecting only light hazes at a pressure level of 1.6 atmospheres. The water and sulfur abundances were low. The lightning detector—basically an AM radio that listened for bursts of static—registered only faint discharges. In short, the weather was clear and dry. So what had gone wrong with the prediction? One piece of the answer came quickly. Infrared images from Earth-based telescopes discovered that the probe had unwittingly hit a special type of atmospheric region known as a hot spot—a clearing where infrared radiation from lower, hotter levels leaks out. Jupiter has many such regions, and they continually change, so the probe could not be targeted to either hit or avoid them.

The luck (both good and bad) of descending in a hot spot did not entirely solve the mystery, however. Scientists had expected that even in these regions the gases at the depths the probe reached would match the average composition of the atmosphere. If so, Jupiter has an anomalously low amount of such elements as oxygen and sulfur. But no one has proposed a process that would eliminate them so efficiently. The other possibility is that the composition of the hot spot differs from the average, perhaps because of a massive downdraft of cold, dry gas from the upper atmosphere.

The latter theory has its own difficulties but currently seems the more likely interpretation. Just before the probe ceased transmitting, concentrations of water, ammonia and hydrogen sulfide were beginning to rise rapidly—just as if the probe was approaching the base of a downdraft. Orbiter images of another prominent hot spot show that winds converge on the center of the hot spot from all directions [*see illustration on page 56*]. The only place the gas can go is down. Orbiter spectra showed that the abundance of water and ammonia varies by a factor of 100 among different hot spots, supporting the idea that local me-

NASA/JPL; ALFRED T. KAMAJIAN (*diagram*)

teorological conditions dictate the detailed composition of the atmosphere.

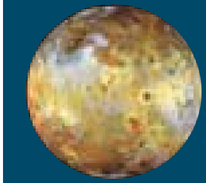
The one part of the weather prediction that proved correct was "windy." Jupiter's cloud bands are associated with high-velocity jet streams: westerlies and easterlies that blow steadily at several hundred kilometers an hour. On Earth the analogous winds die off near the surface. On Jupiter there is no surface; the wind profile depends on which energy source dominates the atmosphere. If a source of internal energy (such as slow contraction under the force of gravity) dominates, the winds should stay strong or increase with depth. The opposite is true if external energy (such as sunlight) is the main contributor. By tracking the probe's radio signal, scientists ascertained that winds at first increase rapidly with depth and then remain constant—indicating that Jupiter's atmosphere is driven by internal energy.

Onto Each Planet Some Rain Must Fall

ALTHOUGH THE PROBE detected only distant radio noise from lightning, the orbiter saw bright flashes illuminating the clouds in what are obviously massive thunderstorms [*see box on pages 56 and 57*]. Like Voyager, Galileo found that lightning was concentrated in just a few zones of latitude. These zones are regions of anticyclonic shear: the winds change speed abruptly going from north to south, creating turbulent, stormy conditions. As on Earth, lightning may occur in water clouds where partially frozen ice granules rise and fall in the turbulence, causing positive and negative charges to separate. How deep the lightning occurs can be estimated from the size of the illuminated spot on the clouds; the bigger the spot, the deeper the discharge. Galileo deduced that the lightning is indeed originating from layers in the atmosphere where water clouds are expected to form.

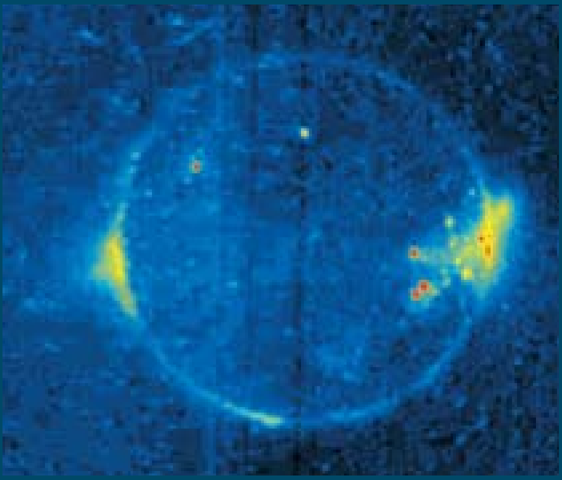
For all its pains, the probe descended less than 0.1 percent of the way to the center of the planet before succumbing to the high pressures and temperatures. Nevertheless, some of its measurements hint at what happens deeper down. The concentrations of noble gases—helium

NASA/JPL

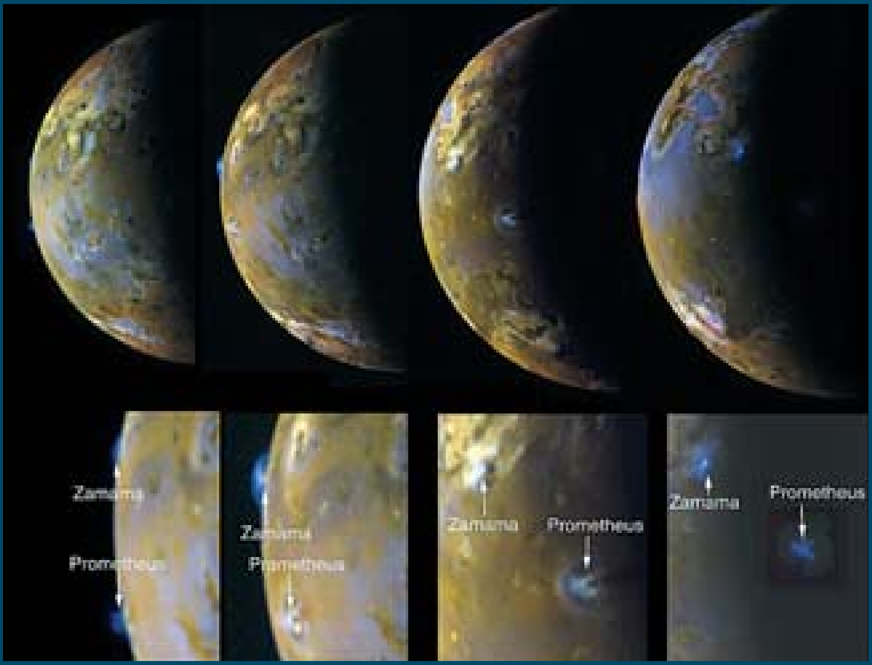


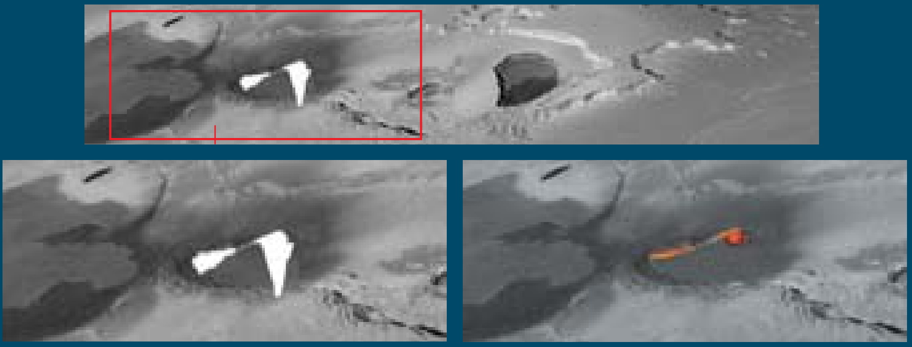
The Infernal Moon

Io

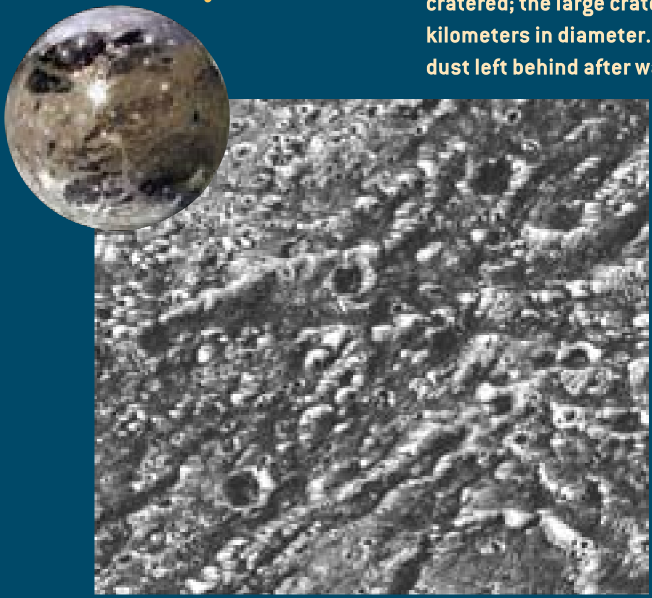


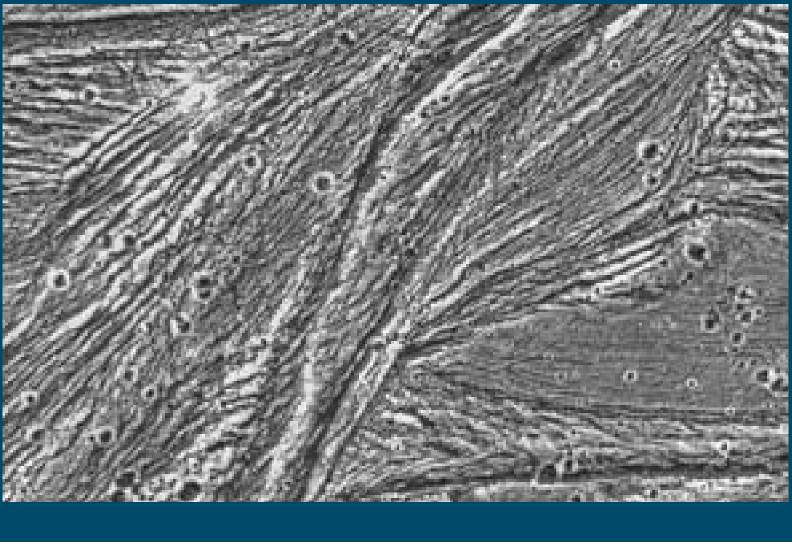
A pizza color distinguishes Io in Voyager images from two decades ago (*above*). Galileo's greater wavelength range permits more spectacular false-color views. When Io is in Jupiter's shadow (*right*), light becomes evident from hot lava flows (*red and yellow spots*) and electron-stimulated emissions in volcano plumes (*blue and yellow glows at edges*). A sequence of enhanced images (*below, top*) shows Prometheus and Zamama coming into view—first the smoky plumes, then the volcanoes themselves surrounded by rings of debris 100 kilometers in diameter. In November 1999 Galileo captured a huge volcanic complex in Io's northern climes (*below, bottom*). Craters and a massive curtain of fire are visible. Fresh lava glows so brightly it overexposed the CCD camera (*white*).





The Ice-Laced Moon
Ganymede





The largest satellite in the solar system is a strange quilt of dark and bright terrains. The dark regions, like Galileo Regio (*left*), are heavily cratered; the large crater in the foreground is 19 kilometers in diameter. Deep furrows may contain dust left behind after water ice sublimated away.

The bright regions, like Uruk Sulcus (*below*), have fewer craters and more tectonic features such as grooves. This image depicts an area roughly 400 kilometers square. Some regions, like Tiamat Sulcus (*right*), shown here just after sunrise, contain both types of terrain.

NASA/JPL

(the second most abundant element in Jupiter, after hydrogen), neon, argon, krypton and xenon—are particularly instructive. Because these gases do not react, they are comparatively unambiguous tracers of physical conditions within the planet. So informative is the concentration of helium that the Galileo atmospheric probe carried an instrument dedicated solely to its measurement.

Infrared spectra obtained by Voyager suggested that Jupiter contains proportionately much less helium than the sun does, an indication that something must have drained this element from the upper atmosphere. Galileo, however, found that Jupiter has nearly the same helium content as the outer layers of the sun. This result still requires that some process remove helium from the atmosphere, because the outer layers of the sun have themselves lost helium. But that process must have started later in the planet's history than researchers had thought. Galileo also discovered that the concentration of neon is a tenth of its solar value.

Both these results support the once controversial hypothesis that the interior of Jupiter is deluged with helium rain. There helium becomes immiscible in the hydrogen-rich atmosphere, which at high pressures—millions of times sea-level pressure on Earth—is perhaps better thought of as an ocean. Being heavier, the helium settles toward the center of the planet. Under certain conditions, neon dissolves in the helium raindrops. Helium may also precipitate out on Saturn, where its depletion may be even more extreme.

After several years of analysis, researchers announced the abundance of the other noble gases. Argon, krypton and xenon are enriched compared with the solar composition by about the same factor as carbon and sulfur. That, too, is a mystery. The only way to trap the inferred quantities of these gases is to freeze them—which is not possible at Jupiter's current distance from the sun. Therefore, much of the material that makes up the planet must have come from colder, more distant regions. Jupiter itself may even have formed farther from the sun, then drifted inward.

A final clue to Jovian history came from the measurement of deuterium, one of the heavy isotopes of hydrogen. The concentration is similar to that on the sun and different from that of comets or of Earth's oceans. This suggests that comets have not had a major effect on the composition of Jupiter's atmosphere, despite the spectacular effects when they hit, as demonstrated during the Shoemaker-Levy 9 collisions in 1994.

World of Fire

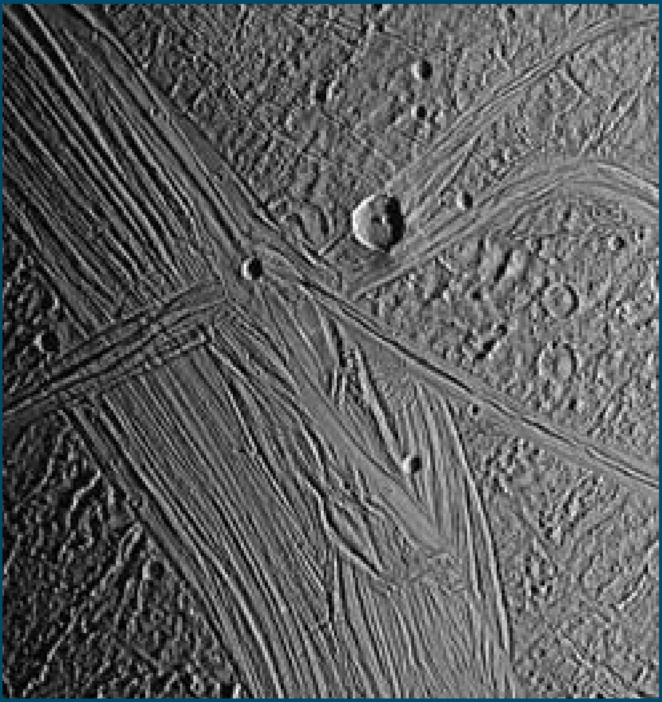
AFTER THE ORBITER relayed the probe information to Earth, it commenced its tour of the Jovian system—more than 30 orbits of the planet, with multiple flybys of each of the four Galilean satellites. The limelight has been placed on Europa, whose surface geology and other features point to the existence of a liquid ocean beneath the ice sometime in Europa's history, probably in the geologically recent past.

The innermost Galilean satellite, Io, stole the show during the two Voyager encounters. The initial pictures showed a remarkably young surface, the only one in the solar system with essentially no impact craters. Later, images taken for navigation purposes serendipitously caught immense eruptive plumes. Subsequent observations confirmed that Io is wracked by volcanic activity. The size of Earth's moon, it spews 100 times more lava than Earth does.

Galileo has spent less time looking at Io than at the other moons, because of danger to the craft: Io lies deep in Jupi-

THE AUTHOR

TORRENCE V. JOHNSON has an asteroid named after him: 2614 Torrence, a body about one kilometer in diameter. Working at the Jet Propulsion Laboratory in Pasadena, Calif., he has been the project scientist for Galileo since 1977—some three quarters of his career as a planetary scientist. He was a member of the imaging team for Voyager and is now on the imaging team for the Cassini mission to Saturn.



Artist's impression of the surface

NASA/JPL (*left*); DON DAVIS (*right*)

ter's intense radiation belts. Galileo flew within 900 kilometers of Io's surface just before the orbit insertion in 1995 but did not revisit until October 1999, when the bulk of its mission had been completed and scientists felt free to take more risks. Although concerns about the jam-prone tape recorder forced cancellation of imaging and spectroscopy during the 1995 flyby, the particle detector and magnetometer remained active.

They found that the empty space around Io is anything but. It seethes with subatomic particles blasted out by volcanic eruptions and stirred up by Jupiter's magnetic field. Electron beams course down the field lines that connect Io to Jupiter's atmosphere; dense, cold plasmas permeate the wake left behind Io by the magnetic field sweeping by. Whenever Io passed through Jupiter's shadow, Galileo saw the moon outlined by a thin ring of glowing gas, lit up by the impact of electrons from the Jovian magnetosphere. In short, Io is tightly linked to the giant planet by what amounts to the largest electric circuit in the solar system [*see illustration on page 58*].

For most of its mission Galileo studied the tortured surface of Io from a safe distance. Based on how brightly the volcanoes glow at different visible and near-infrared wavelengths, it inferred their temperature, a measurement critical to determining the composition of the lavas. Most volcanoes on Earth disgorge lava of basaltic composition—iron, magnesium and calcium silicates rich in the minerals olivine and pyroxene. Basaltic melts typically have temperatures ranging from 1,300 to 1,450 kelvins (1,050 to 1,200 degrees Celsius). In contrast, telescopic observations of Io half a dozen years ago suggested temperatures of 1,500 to 1,800 kelvins. These temperatures ruled out substances that melt at lower temperatures, such as liquid sulfur, which had been suggested previously as a dominant volcanic fluid on Io.

When Galileo's measurements came down, the enigma intensified. Lavas are actually 1,700 to 2,000 kelvins. Magma this hot has not been common on Earth for more than three billion years. Io may thus be giving scientists an unexpected glimpse into Earth's geologic youth, a time when its interior temperatures were higher and the composition of the upper mantle differed.

When Galileo finally returned to Io in autumn 1999, the mission team was uncertain whether the spacecraft would survive the radiation. On one of its passes, it autonomously aborted the data-taking sequence just four hours before reaching Io, and the team rebooted with only minutes to spare. Several instruments also suffered damage, but all continued to work and in the end returned spectacular data. Io's active volcanoes were finally captured up close and personal [*see illustration on page 59*].

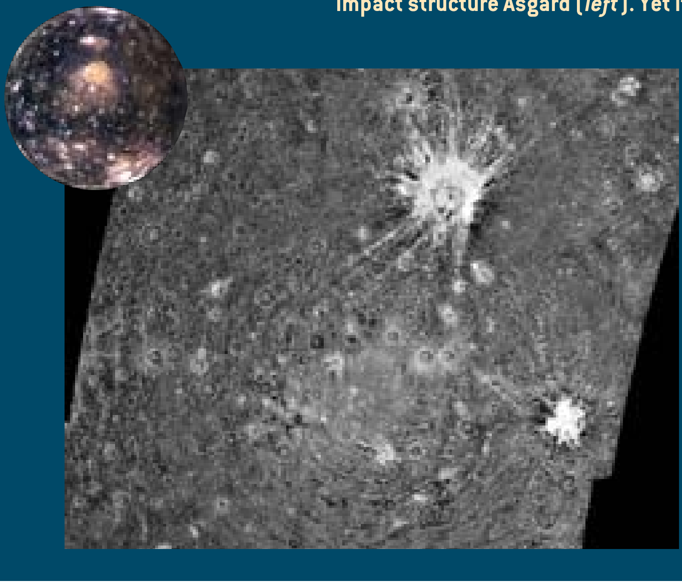
In a Field of Its Own

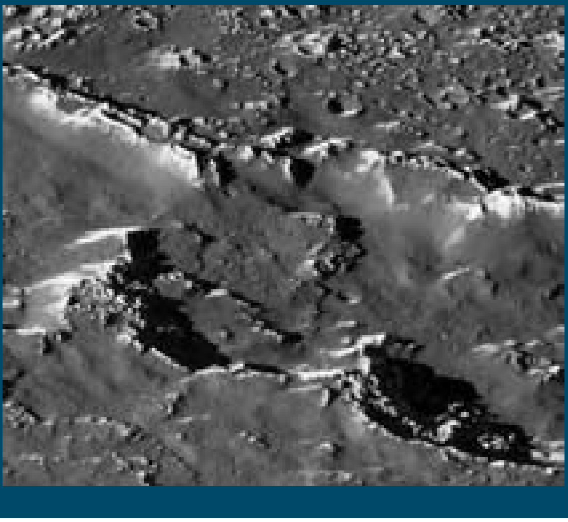
ONE OF GALILEO'S major discoveries was made during its very first orbital encounter—with Ganymede, Jupiter's largest moon. About half an hour before the spacecraft reached its closest approach, the radio-noise instrument, designed to record ambient electrical fields, began to go haywire. The relatively quiet background radio signals seen in most of the Jovian system changed abruptly to a complex, active radio spectrum. For 45 minutes the activity remained intense, and then it ceased as suddenly as it had begun. When the radio noise commenced, the magnetometer readings shot up fivefold.

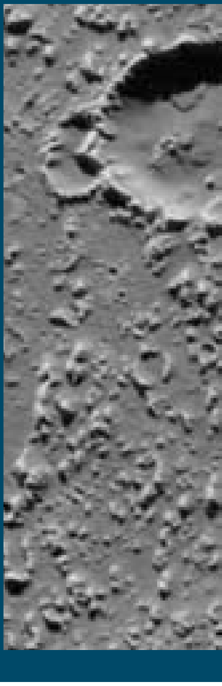
Plasma researchers had seen signatures of this sort before, when spacecraft carrying similar instruments entered and exited magnetospheres at Earth, Jupiter, Saturn, Uranus and Neptune. Detailed magnetic field measurements on two subsequent Ganymede flybys confirmed their suspicions: the moon is magnetized, generating a dipole field similar to those of these planets. No other satellite explored to date has such a field. Earth's moon and Mars may have had fields in the past, but currently they exhibit only limited patches of magnetic variation that represent magnetized rocks on the surface. Like a

Pockmarked Callisto

This most baffling of the Galilean satellites is densely packed with large craters, such as the massive, multiringed impact structure Asgard (*left*). Yet it is comparatively free of small craters, and those that do exist are fuzzy (*below* and *right*)—suggesting that dusty material has somehow flowed across the surface.







NASA/JPL

set of nested Russian dolls, Ganymede has a magnetosphere contained within Jupiter's huge magnetic domain, which in turn is embedded in the sun's.

Tracking of the spacecraft signal allowed researchers to probe Ganymede's gravity field and therefore its internal structure. They concluded that it probably has a dense core about 1,500 kilometers in radius with a surrounding icy mantle 700 kilometers deep. Geochemical models suggest that the core consists of a sphere of iron or iron sulfide enveloped in rock. The inner metallic core could produce the dipolar magnetic field.

Yet theorists are not sure quite how. Solid iron at the center would be too hot to retain a permanent magnetic field. Instead a magnetic field is thought to involve a convecting, conductive liquid. Models of Ganymede indicate that its interior can easily become hot enough to melt iron or iron sulfide. But the same models show that convection will cease as the core gradually cools; the conditions required for convection should last only a billion years or so.

The answer may lie in the orbital resonance of the inner three Galilean satellites. Io goes around Jupiter precisely four times for each time Europa completes two circuits and Ganymede one. Like pushing a child's swing in time with its natural pendulum period, this congruence allows small forces to accumulate into large outcomes—distorting the orbits from their default circular shape into more oblong ellipses. The effect on the moons is profound. Because the distance between them and Jupiter is continuously changing, the influence of Jupiter's gravity waxes and wanes, stretching the moons by varying amounts. The process, known as tidal heating, drives the volcanism on Io and keeps Europa's putative ocean from freezing.

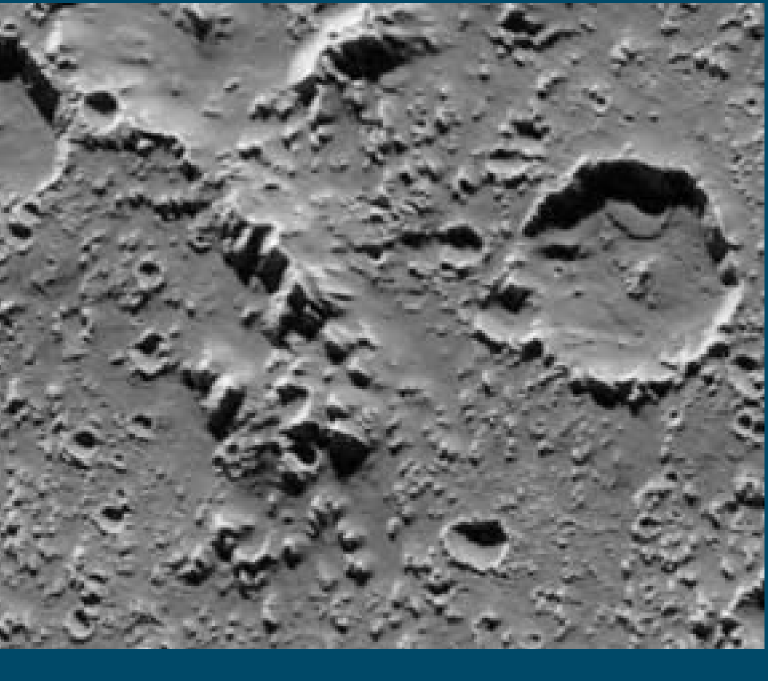
Researchers used to think that tidal heating was of little consequence for Ganymede, the outermost of these three moons. But now they realize that the orbits may have shifted over time. The resonances may once have been stronger and Ganymede's orbit more perturbed. The immense fault systems across the surface may record this earlier period of intense heating. If so, the moon is still cooling, and its core can continue to generate a magnetic field.

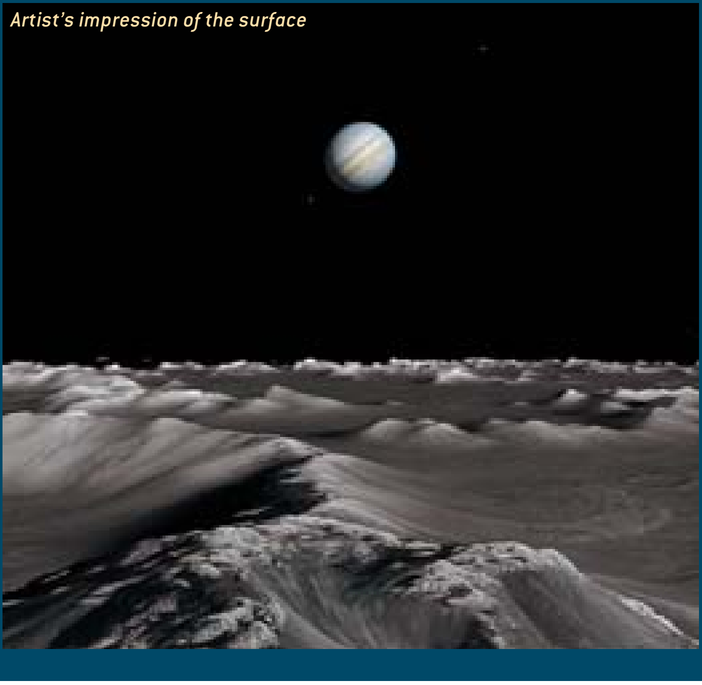
Compared with flamboyant Europa, Io and Ganymede, Callisto was thought rather drab. In Voyager images it epitomized the traditional stereotype for icy satellites: an old, frozen, pockmarked mudball. But Galileo observations tell a different story.

Old but Hardly Dull

CALLISTO IS COVERED with large impact scars, ranging from craters kilometers in diameter to the so-called palimpsest named Valhalla, some 1,500 kilometers across. The surface is believed to date back more than four billion years to the rain of meteoritic and cometary debris left after the formation of the planets and satellites. In this sense, Callisto is indeed old. Seen close-up, however, Callisto's surface is blanketed by fine, dark debris. Small craters, which on most other bodies are produced in abundance, are largely absent. Surface features appear softened and eroded. Clearly, some young processes have been at work. Among the ideas proposed have been electrostatic levitation of fine dust, which would allow it to "flow" across the surface, and evaporation of ices from the surface, which would leave behind deposits of darker, less volatile material. But none of these explanations is satisfying.

Intriguingly, near-infrared spectra show not only water ice and hydrated minerals, as expected, but also four unusual absorption features near a wavelength of four microns. One appears to be carbon dioxide trapped in the surface, perhaps as inclusions in icy particles or bubbles produced by radiation damage to the surface. Two other spectral features probably represent sulfur in the surface, which may originate in Io's volcanic eruptions. The fourth spectral feature is the strangest. Its wavelength corresponds to that absorbed by carbon-nitrogen bonds. In fact, laboratory spectra of complex organic molecules called tholins by the late Carl Sagan are similar. Tholins are thought to resemble organic material in





Artist's impression of the surface

NASA/JPL (*left*); DON DAVIS (*right*)

the solar nebula; clouds of interstellar ice grains have comparable spectra. Taken together, the data provide the first direct evidence that icy satellites contain the carbon, nitrogen and sulfur compounds common in primitive meteorites and comets. These materials are also some of the most important for life.

Similar spectral signatures have now been seen on Europa and Ganymede. Unlike the other Galilean satellites, Callisto seems more like a uniformly dense sphere, indicating that most of its rock and ice are mixed together. A core is ruled out. Therefore, the interior has never been heated strongly. The moon does not participate in the orbital resonance that kneads the other Galilean moons.

On the other hand, the orb is far from dead. As the Galileo magnetometer found, Callisto seems to perturb the surrounding Jovian magnetic field in a peculiar pattern. This disturbance resembles what is seen in physics experiments in which a hollow copper sphere is subjected to a changing magnetic field: electric currents are set up in the shell, which produces a magnetic field that counters the imposed field.

But what could form the electrically conducting layer? Rock, ice and ionospheric particles are poor conductors. Researchers are left with a possibility that not long ago seemed outrageous: salty ocean water. Seawater is a weak conductor, and a global liquid layer some tens of kilometers thick could produce the observed signature. The combination of evidence for a comparatively undifferentiated interior and for a global ocean presents a severe challenge for theorists. Somehow Callisto must be hot enough to support an ocean but not so hot that light and heavy materials separate. The water layer might be sandwiched between a radioactively heated interior, where convection keeps the material mixed, and a thin icy shell, where a different convection cycle cools the ocean. These results were so exciting they prompted the Galileo team to design, while the mission was under way, special measurements that were taken during flybys that provided key evidence for oceans. So much for dull, old Callisto.

Galileo's camera also captured each of the four inner, small moons—Metis, Adrastea, Amalthea and Thebe. A major finding was that these bodies are responsible for Jupiter's rings. A special series of pictures was taken while the spacecraft was within Jupiter's shadow, allowing the sun to backlight the tiny dust particles that make up the rings. These images reveal the complex structure of the tenuous gossamer ring. It consists of multiple layers directly related to the orbits of Amalthea and Thebe. Thus, the rings are probably microscopic debris kicked off the moons by tiny meteoroids.

The data gathered by Galileo have revolutionized scientists' view of Jupiter and its moons, which we have come to recognize as a kind of planetary system comparable in complexity to the solar system itself. The Voyager flybys provided the adrenaline rush of seeing worlds for the first time, but only an intensive investigation such as Galileo's could have revealed the nuances and the limitations of deceptively straightforward characterizations such as "thundercloud" and "icy satellite." Soon it will be Saturn's turn to enter this new phase of exploration. Another two-in-one spacecraft—Cassini-Huygens—arrives there in 2004. It, too, will probably raise more questions than it answers.

MORE TO EXPLORE

Galileo Probes Jupiter's Atmosphere. Special section in *Science,* Vol. 272, pages 837–860; May 10, 1996.

Galileo Orbiter. Special section in *Science,* Vol. 274, pages 377–413; October 18, 1996.

Remote Sensing of the Galileo Orbiter Mission. Special issue of *Icarus,* Vol. 135, No. 1; September 1998.

Encyclopedia of the Solar System. Edited by Paul R. Weissman, Lucy-Ann McFadden and Torrence V. Johnson. Academic Press, 1999.

The New Solar System. Fourth edition. Edited by J. Kelly Beatty, Carolyn Collins Peterson and Andrew Chaikin. Sky Publishing, 1999.

By Robert T. Pappalardo,
James W. Head and Ronald Greeley

The Hidden Ocean of EUROPA

Doodles and freckles, creamy plains and crypto-icebergs—the amazing surface of Jupiter's brightest icy moon hints at a global sea underneath

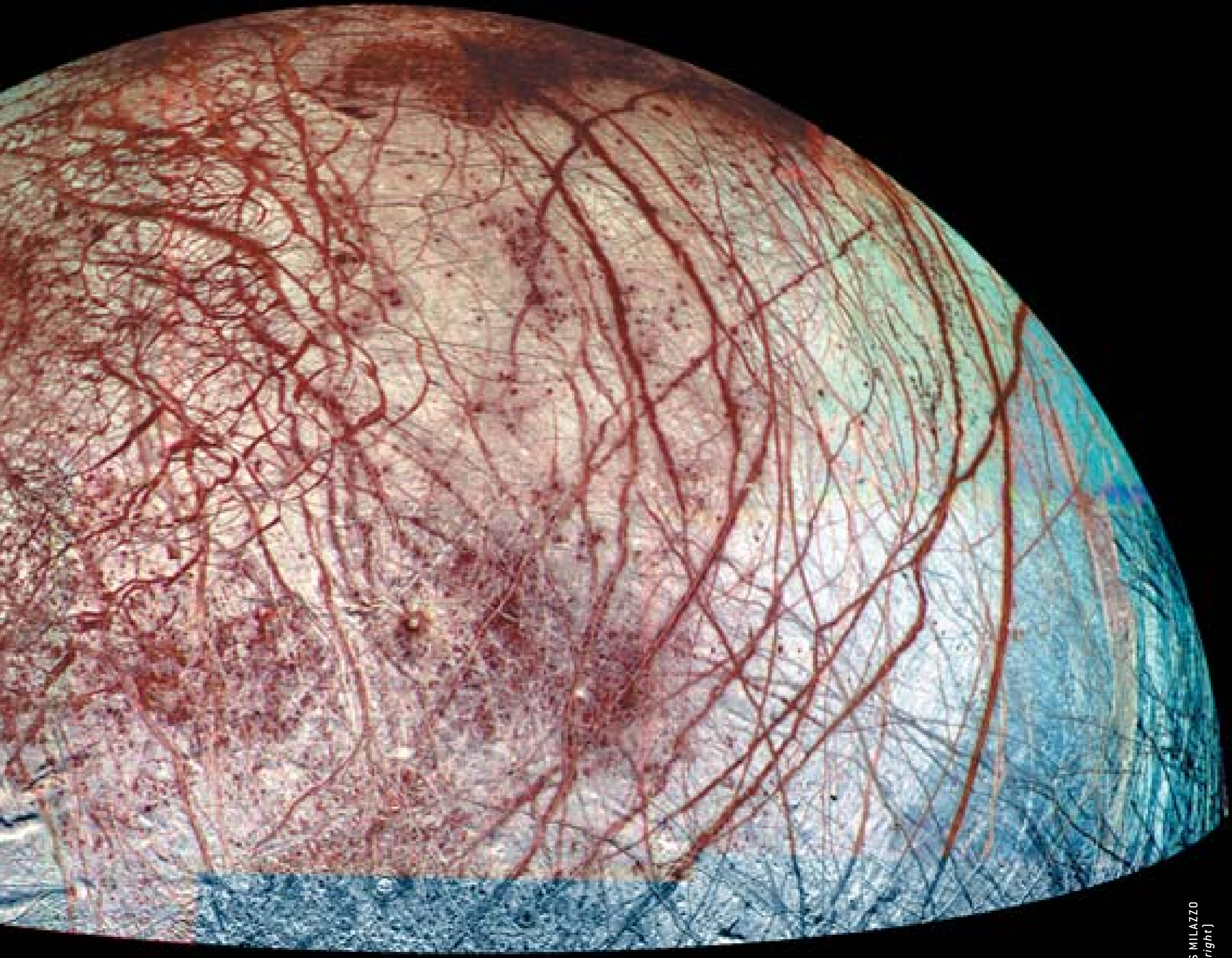
Do living things flourish elsewhere within our solar system, or is Earth's environment uniquely nurturing? This question is central to planetary exploration today. Three decades into humankind's reconnaissance of the planets and their natural satellites, only a short list of possible abodes remains. Perhaps the most intriguing is Jupiter's ice-rich moon Europa.

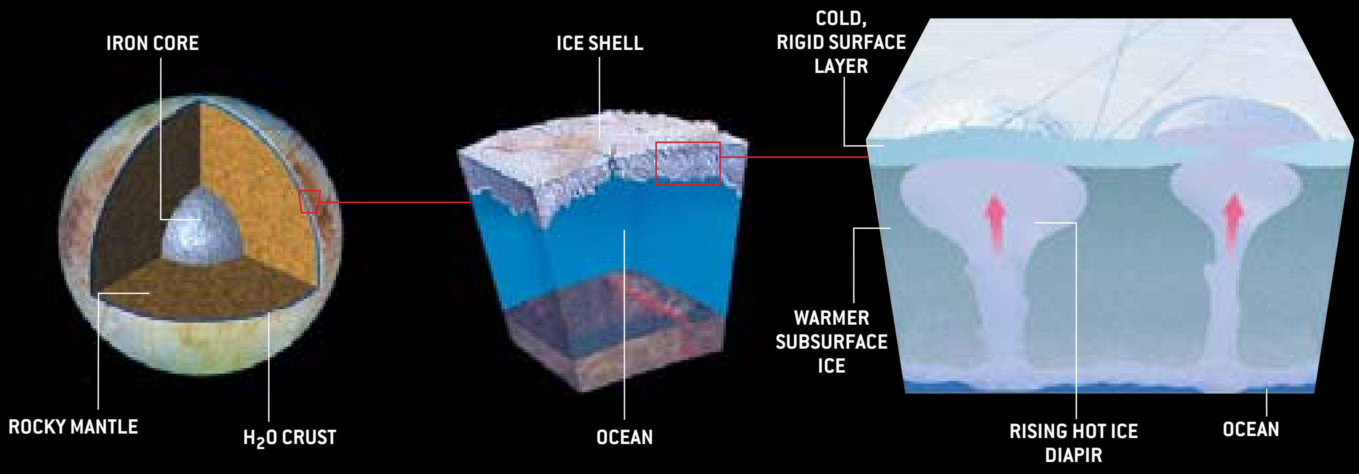
For centuries, astronomers knew Europa only as a pinprick of light in even the most powerful telescopes. In the 1960s spectroscopy showed that the satellite, like many others in the cold reaches of the outer solar system, is covered with ice. With surface temperatures of 110 kelvins (–260 degrees Fahrenheit) near the equator and 50 kelvins near the poles, that ice must form a rock-hard skin. Researchers had no way to probe deeper and little reason to expect anything special. But in the past two decades and especially in the 1990s, spectacular images radioed from visiting spacecraft have revealed a young and tremendously deformed surface. Somewhere under the icy shell, it seems, must be a warm, mobile interior. Are Europa's innards warm enough to sustain an ocean of liquid water? If so, we can stretch our imaginations and ask whether life might have arisen within the lightless depths.

Planetary scientists have been trying to infer what lies inside Europa ever since the two Voyager spacecraft flew by Jupiter and its companions in 1979. Celestial mechanics dictated that these spacecraft could pass by Europa only distantly. The photographs they did obtain were nonetheless tantalizing. Europa looked like a ball of string, its bright plains crisscrossed with bands and ridges. Researchers noticed that some dark wedge-shaped bands have opposing sides that match each other perfectly. Somehow the bright icy surface has been wrenched apart, exposing dark material that was fluid enough to permeate the ensuing void. These features resemble liquid-filled openings between floating plates of sea ice on Earth.

Unexpectedly, the Voyagers found very few large impact craters on Europa. A planetary surface slowly accumulates impact craters as it is occasionally hit by cometary and asteroidal debris. If Europa all but lacks visible craters, it must have been repaved by volcanic or tectonic events in the relatively recent past. Based on the number of comets with Jupiter-crossing orbits, the late cratering expert Eugene Shoemaker deduced that a crater larger than 10 kilometers (six miles) in diameter should form once every 1.5 million years on average. Extrapolation from the few known Europan craters

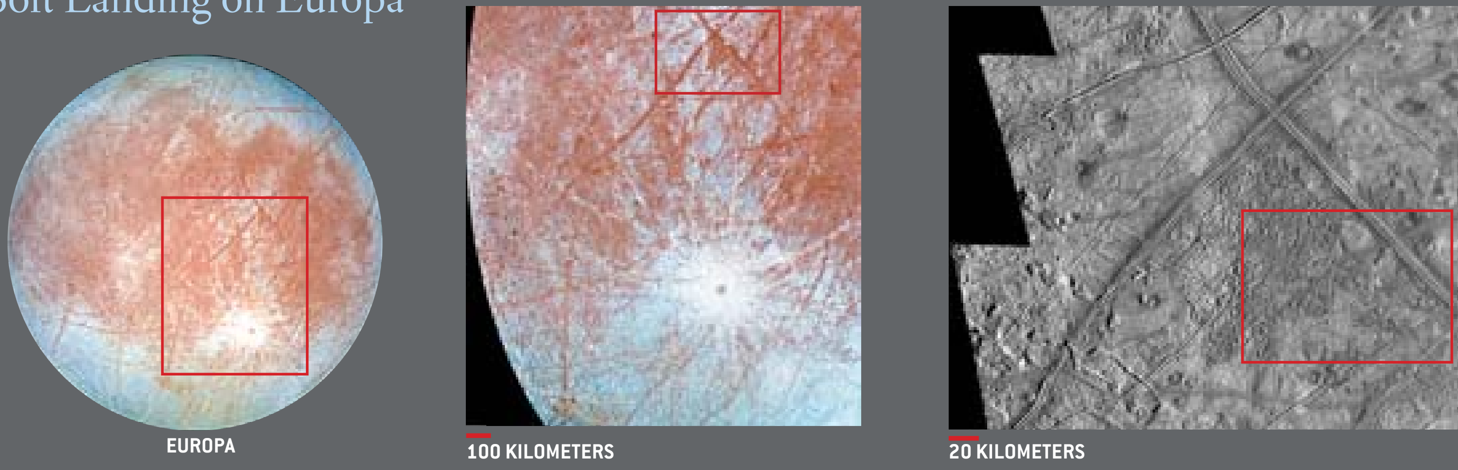
EUROPA'S ICY COUNTENANCE resembles a cracked eggshell. Reddish material has oozed out of fractures opened up by Jupiter's gravitational forces. Very few craters are present, indicating that the surface is geologically young. On this Galileo spacecraft image, the colors are exaggerated but real. Other spacecraft instruments have found that Europa's interior is mainly rock, with an outer layer of water (in either liquid or solid form) about 100 kilometers thick (*bottom right*). Most of that water must be fluid or semifluid to account for surface features, such as the circular mounds pushed up by rising blobs of relatively hot ice (*far right*), which occasionally puncture the surface.





NASA/JET PROPULSION LABORATORY; CYNTHIA PHILLIPS AND MOSES MILAZZO *University of Arizona* (*top*); NASA (*left and center*); TOM MOORE (*right*)

Soft Landing on Europa

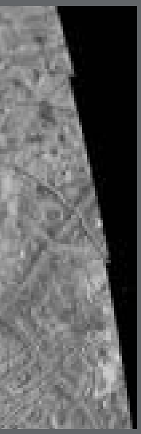


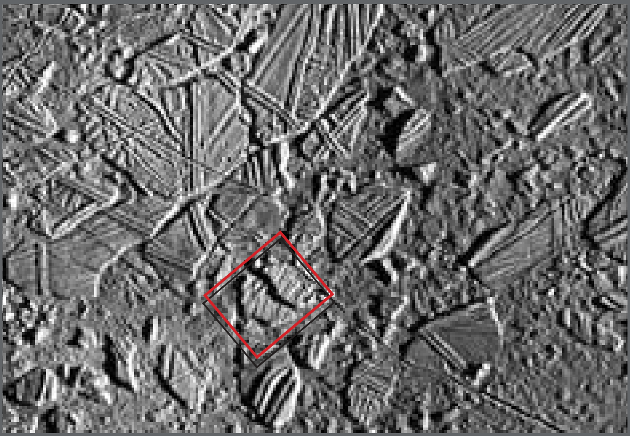
suggested a surface age of just 30 million years—a geologic eyeblink. Shoemaker added that Europa's large craters might have flattened out over time if the interior were warm. The satellite could be active even today.

But this hypothesis remained uncertain. The Voyager images were too coarse to pick out smaller craters. Indeed, intermingled with the bright plains is mottled terrain filled with dark spots, mounds and pits. Some researchers pointed out that craters could be hiding in these odd regions, in which case the satellite's surface would be ancient. Besides, how could a moon so small possibly be active? Similarly sized bodies, such as Earth's moon, are inert balls of rock, having lost most of their internal radioactively generated heat long ago. By all rights, Europa should now be cold and dead.

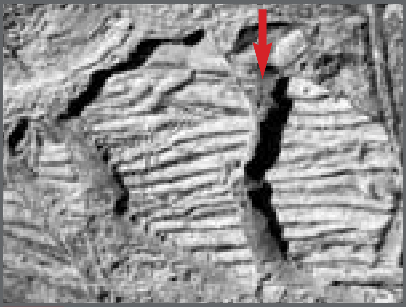
Then researchers came to appreciate the power of an exotic heat source: tidal kneading, the process that drives volcanism on Europa's pizza-colored neighbor Io. Of the four large moons of Jupiter—Io, Europa, Ganymede and Callisto, collectively known as the Galilean satellites in honor of their discoverer—the first three are engaged in an elegant orbital dance called the Laplace resonance. With clockwork precision, each time Ganymede orbits Jupiter once (with a period of 7.2 Earth days), Europa orbits twice (3.6 days) and Io four times (1.8 days). The consequent gravitational push and pull distorts

NASA/JPL; ALFRED T. KAMAJIAN (*bottom left*)





Simulated view of icy blocks, roughly three kilometers from side to side



GIANT BLOCKS OF ICE the size of a small city are progressively revealed in this sequence of Galileo images. A colossal X formed by two ridges (*first three images in top row*) helpfully marks the spot as the pictures zoom in on a dark splotch known as Conamara Chaos (*next two images*) and thence to individual iceberglike blocks (*above*). Warm ice, slush or liquid water once filled the low-lying "sea" in which the blocks sit, now frozen in place. An artist's impression (*left*), simulating a vantage point a few hundred meters above the surface looking south (*arrow in above image*), shows dirt tumbling downslope as fine ice particles slowly evaporate away.

their orbits into oblong ellipses. They move nearer to, then farther from, their parent planet during each orbital revolution. In response, tides are raised and lowered in the body of each satellite. Like bending a paper clip rapidly back and forth, this tidal flexing generates heat [*see bottom illustration on page 71*].

The effects are felt most profoundly on Io, which is the closest to Jupiter. The interior temperature rises to the melting point of rock, powering continual volcanic eruptions. Europa, farther away, is heated less intensely. But calculations indicate that its interior might be kept warm enough to melt ice below a depth of 20 to 30 kilometers, maintaining a global subsurface ocean.

After Voyager, observational tests of the ocean hypothesis had to wait nearly two decades, until the worlds of Galileo could be visited by the spacecraft named for him. That spacecraft swung into orbit around Jupiter in December 1995. Every few months, through 2002, its trajectory brought it speeding closely past one of the Galilean satellites—including, a dozen times, Europa.

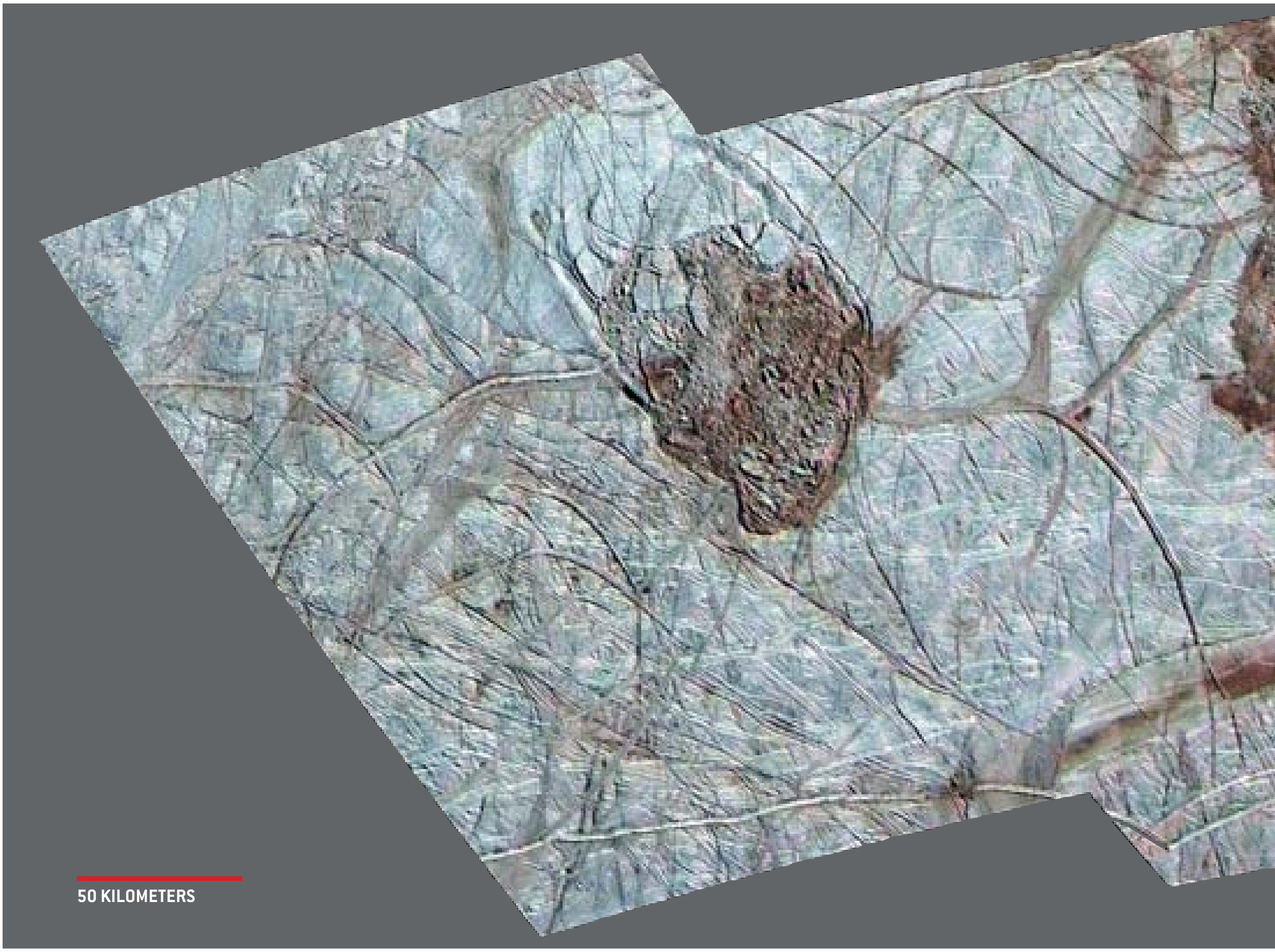
Even if Galileo had not sent back a single picture, it would have provided a vital insight. On each flyby, engineers and scientists have carefully tracked the spacecraft's radio signal in order to measure Europa's gravitational field. Any rotating and tidally distorted moon is slightly flattened, or oblate, so its gravitational field is also nonspherical. The irregular force causes slight shifts in the frequency of Galileo's signal, from which researchers have quantified the satellite's oblateness and, in turn, its internal mass distribution. (For a given rotation rate, a satellite with a more centrally concentrated mass will be less oblate than a homogeneous satellite.)

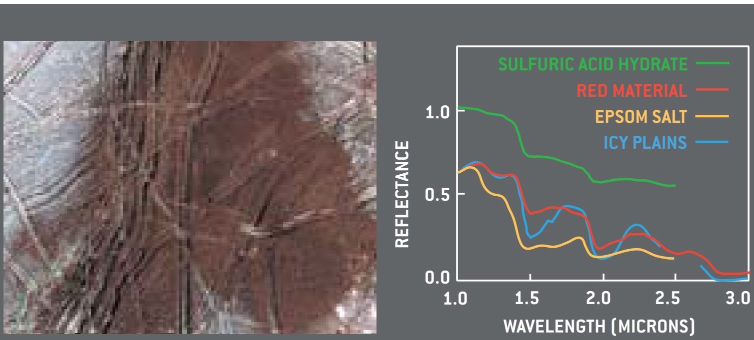
Tapestries and Hydraulic Bearings

JUDGING FROM its average density of 3.04 grams per cubic centimeter, Europa is predominantly a rocky object. The gravity data indicate that the rock is sandwiched between a central iron core and an outer crust of H_2O. Considering the likely range in density values for the iron core and rocky mantle, the water crust is between 80 and 170 kilometers thick, most likely about 100 kilometers. If a significant portion of it is liquid, its volume exceeds that of all the oceans of Earth combined. But Galileo's gravity data cannot tell whether this water layer is completely solid or partially liquid.

To address that question, one must look to the other data, beginning with the pictures. The Galileo imaging team found a world like no other. Its surface is an elaborate weave of fractures, ridges, bands and spots. The fractures presumably formed as tidal forces distorted the icy surface until it cracked. Ridges are similarly ubiquitous. They slice across the surface in pairs, each with a narrow valley down the center. Plausible models for their formation invoke the rise of liquid water or warm glacial ice along fractures. A watery or icy "magma" might have forced the rigid near-surface ice upward, warping it into a double ridge. Or an icy slurry might have erupted onto the surface to build each ridge. Multiple parallel ridges also occur, indicating that the process can repeat to create ridges side by side. The widest ones are commonly flanked by dark, reddish, diffuse-edged stripes. Perhaps the heat pulse associated with ridge formation created these darkened margins through icy volcanism or sublimation of a dirty ice surface. Whatever the formation mechanism, ridges point to a dynamic geologic history and warm subsurface.

From the seemingly random doodlings of fractures and ridges, scientists have attempted to understand the manner in which Europa has been stretched and distorted. Tidal knead-





RUDDY SPOT marks where briny liquid poured out onto Europa's surface (*left*). Spectral measurements (*right*) found that surrounding bright plains (*blue*) consist mainly of water ice. The ruddy material (*red*) more nearly matches the laboratory spectrum of magnesium sulfate (Epsom salt, *yellow*) and sulfuric acid (battery acid, *green*).

NASA/JPL (*left*); SARAH DONELSON; SOURCE: THOMAS B. MCCORD *University of Hawaii* (*right*)

ing produces a distinctive pattern, and some of Europa's freshest cracks and ridges fit that pattern. But something else must also have been going on. Strangely, it appears that the stress pattern has swept across the surface over time.

In fact, the pattern can be explained if Europa's surface has rotated faster than its interior. Most of the solar system's natural satellites are in synchronous rotation: torqued by tidal forces, they come to rotate exactly once for each orbital revolution, always showing the same face to their parent planet. But if Europa's icy surface were decoupled—mechanically separated—from its rocky mantle, Jupiter's gravity would cause the surface to spin slightly faster than the synchronous rate. A subsurface ocean could easily act as such a bearing, allowing the floating ice shell to rotate nonsynchronously.

It cannot be said whether nonsynchronous rotation is going on today or whether the surface instead records an ancient pattern of now inactive lineaments. Scientists have compared the locations of features in Galileo images with their locations in Voyager images and found no measurable change over that 20-year period. Relative to the interior, the surface today cannot be rotating faster than once every 10,000 years.

LEATHERY EXOSKELETON represented by chaos terrain on Europa is one of the strangest landscapes in the solar system. The Thera region (*left half of mosaic*) contains bright, dislodged plates of ice. The Thrace region (*right half*) is elongated, hummocky and higher in elevation. It spills into a gray band, Libya Linea, to the south. Such chaos regions might have formed when an underground ocean melted through the moon's icy shell or when upwelling blobs of warm ice disrupted the surface.

NASA/JPL; CYNTHIA PHILLIPS *University of Arizona*

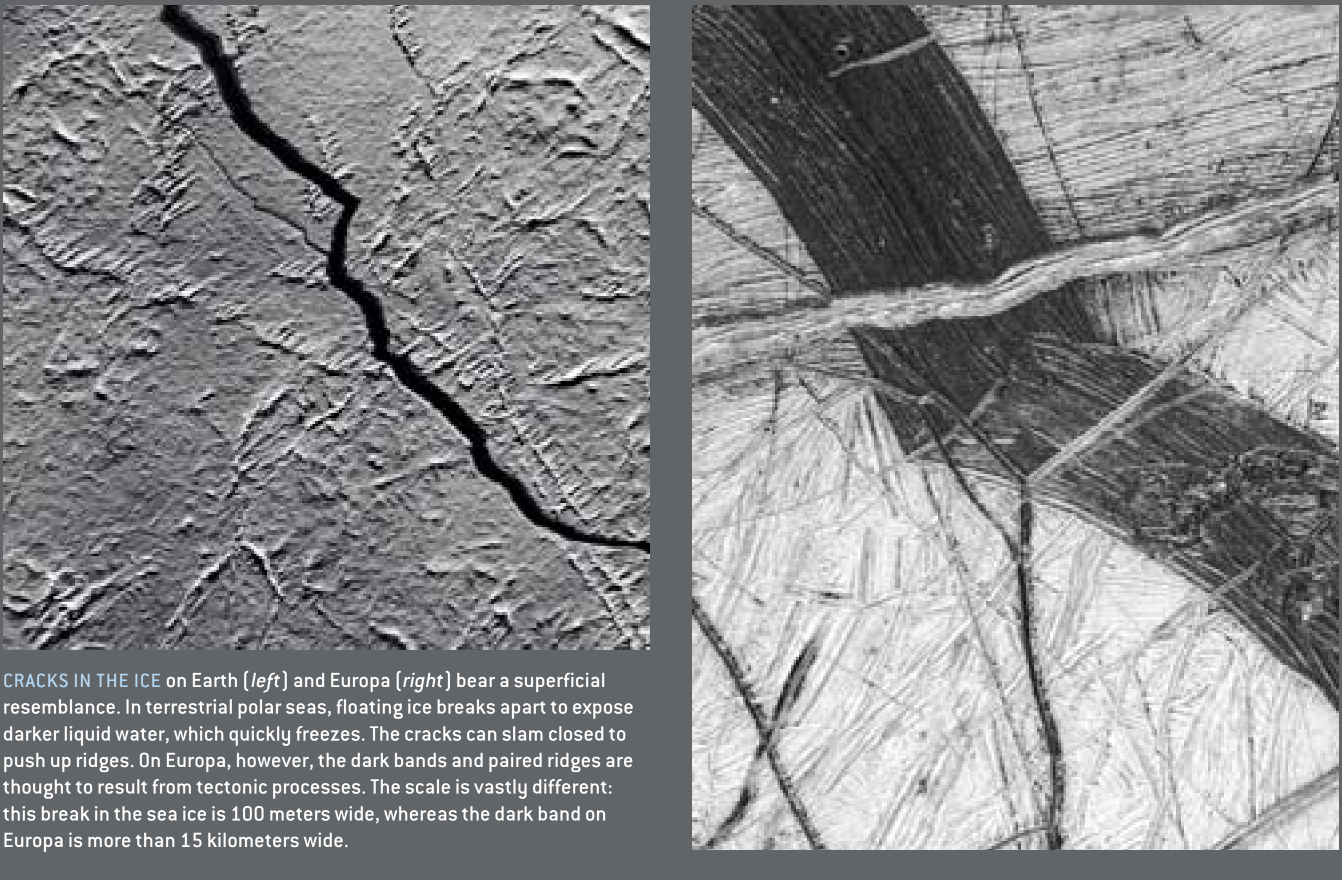
Galileo's camera has also homed in on the dark wedge-shaped bands. Recent analyses have confirmed that the opposing sides of these bands are perfectly matched. The dark material in between is finely striated, commonly having a prominent central groove and some degree of symmetry [*see illustration on next page*]. These bands may be the icy equivalents of spreading centers—locations on Earth's ocean floors where tectonic plates move apart and new rock surges up. If so, the subsurface ice must have been mobile and warm when the features formed. But plate tectonics is a zero-sum game: if some material emerges from the interior, other material must descend. On Earth this descent occurs at subduction zones. No such zones have yet been identified on Europa.

Blame the Blobs

THE MYSTERIOUS MOTTLED TERRAIN provides further clues to Europa's interior. Galileo's images show it to be peppered with circular and elliptical features that the imaging team named lenticulae, Latin for "freckles." Many are domes, some are pits and some are smooth dark spots; others have a jumbled and rough texture. The dome tops look like pieces of the older ridged plains, intimating that the domes formed when the plains were pushed upward from below.

The variety of lenticulae can be explained if Europa's icy shell has behaved like a planetary lava lamp, with blobs of warm ice rising up through the colder near-surface ice. In that case, domes formed when the blobs pressed against the underside of the surface. Rough textures may be places where blobs disrupted and destroyed the plains. Smooth dark patches may be meltwater unleashed by blobs and quickly refrozen.

Blobs—technically, diapirs—would develop if Europa's icy shell floated above liquid water. Tidal flexing pumps heat into the base of the shell, where ice is near its melting temperature and most easily deformed. The warm ice is less dense than the cold ice above, so it attempts to rise. If the ice shell is thick enough, buoyancy forces can overcome the viscous resistance to flow (which lessens with depth). Like wax rising in a lava



CRACKS IN THE ICE on Earth (*left*) and Europa (*right*) bear a superficial resemblance. In terrestrial polar seas, floating ice breaks apart to expose darker liquid water, which quickly freezes. The cracks can slam closed to push up ridges. On Europa, however, the dark bands and paired ridges are thought to result from tectonic processes. The scale is vastly different: this break in the sea ice is 100 meters wide, whereas the dark band on Europa is more than 15 kilometers wide.

MAX COON NORTHWEST RESEARCH ASSOCIATES (*left*); NASA/JPL; ROLAND WAGNER *German Aerospace Research Agency* (*right*)

lamp, warm-ice diapirs will rise toward the surface, where they could create the visible lenticulae. Models suggest that the shell would have to be at least 20 kilometers thick.

As well as the lenticulae, mottled terrain contains the most spectacular of Europa's features: regions of "chaos." In these jumbled areas, small icy remnants of preexisting ridged plains appear to have jostled in a hummocky matrix—like icebergs calved into a slushy sea. The original arrangement of the iceberglike blocks can be reconstructed like a jigsaw puzzle, and researchers have done so for one of these areas, Conamara Chaos [*see illustration on page 73*]. If the regions formed when subsurface water melted through Europa's icy shell and then refroze, the iceberg analogy may be right on target. Another possibility is that one or more diapirs welled up and heated the near-surface ice, creating a slushy bed of ice and liquid on which the cracked and dislodged blocks of ice could slide freely. Either way, the chaos regions tell of a warm subsurface and at least partial melting.

The one type of feature the mottled terrain conspicuously lacks is small impact craters. So the surface of Europa must indeed be young. Following on Shoemaker's pioneering age estimates, researchers have modeled the solar system's comets and asteroids to understand the rate at which they strike Europa. They agree with Shoemaker's suggestion that it is primarily comets that slam into the Galilean satellites; asteroids are simply too few in number. From the presumed and observed numbers of comets in the vicinity of Jupiter—including Comet Shoemaker-Levy 9, which plunged into the gas giant in July 1994—scientists calculate that the sparsely cratered landscape of Europa is about 60 million years old. By geologic standards, that is a short amount of time. Therefore, it seems likely that Europa is still active today, although no volcanic smoking guns have been found, as on Io.

The few craters that do exist on Europa's surface are themselves a probe of the thickness of the icy shell. Unlike the bowl-shaped or flat-floored impact craters on other worlds, Europa's two largest impact features have a central smooth patch surrounded by concentric rings [*see top illustration on opposite page*]. The blasts that created these features must have penetrated the rigid near-surface ice to a weak layer below. Because the weak layer was unable to maintain a crater shape, melt and slush quickly filled in, dragging the near-surface ice inward and

THE AUTHORS

ROBERT T. PAPPALARDO, JAMES W. HEAD and *RONALD GREELEY* have worked together on the Galileo imaging team for several years. Pappalardo learned to appreciate Jupiter's satellites during the 1979 Voyager encounters, when he was in high school. Now a professor of planetary science at the University of Colorado, he researches processes that have shaped the surfaces of icy moons. Head began his career helping to choose Apollo landing sites and train astronauts. Since that time, he has been a geology professor at Brown University and a participant in nearly every major planetary mission. He has collaborated with Russian scientists for several decades. Greeley began to work for NASA while on military assignment in the pre-Apollo days. Seeing how geologic principles could be applied to non-Earth objects—still a new idea at the time—he stayed on at the space agency. He is now a professor at Arizona State University.

fracturing the surface in concentric rings. In essence, the rings are the frozen record of a rock thrown into a pond—a very big rock and a very big pond. Scientists have estimated the dimensions of the impact from the visible scars; in turn, the depth to the weak layer is 20 or more kilometers, in agreement with values from the tidal-heating theory and the blob models. Models are necessarily uncertain, however, and researchers continue to debate the thickness of Europa's ice shell.

The Bands of NIMS

IN ADDITION TO ITS CAMERA, the Galileo spacecraft carried a near-infrared mapping spectrometer (the NIMS instrument), which has analyzed the light reflected by Europa's surface. As expected, NIMS found the characteristic spectral bands of water ice. Yet the bands are skewed and asymmetric in shape, a sign that some impurity is mixed into the ice, especially in areas that appear dark and reddish at visible wavelengths. A prime suspect is a salt—specifically, magnesium sulfate [*see top illustration on page 69*]. If so, sitting on Europa are the biggest deposits of Epsom salt in the solar system. Another possibility is sulfuric acid, like that found in car batteries.

Because salts are generally colorless or white, some other material must be present as well to account for the reddish color. The identity of that contaminant so far eludes scientists, but sulfur compounds are suspected. Before the Galileo mission, some investigators had predicted that an internal ocean on Europa would probably be quite briny, given that many meteorites contain salts. Europa's surface materials may be revealing the chemistry of a hidden brackish ocean.

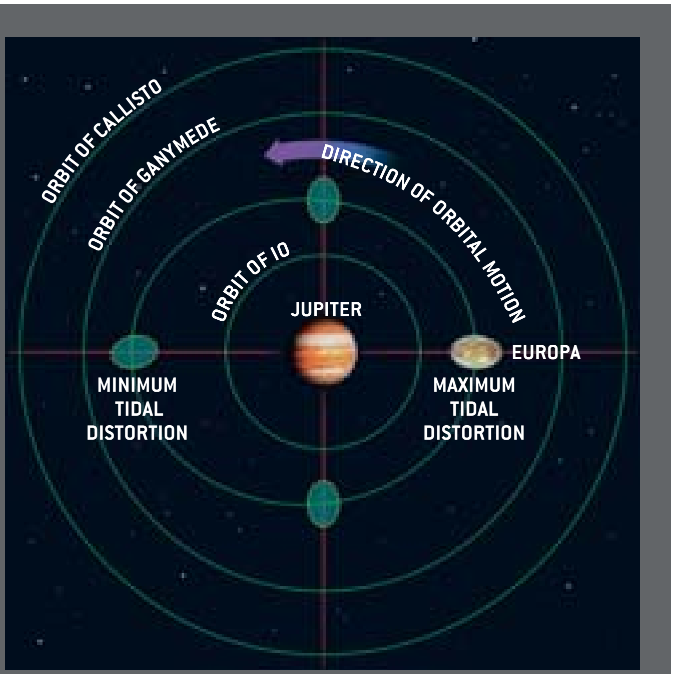
Perhaps the most fascinating indication of Europa's interior state—and probably the best indication of a current-day ocean—has come from a seemingly unlikely source: the Galileo magnetometer. The Galilean satellites are immersed within the powerful magnetic field of Jupiter. Measurements of the ambient field in the vicinity of Europa show deviations associated with the satellite. These rapid and systematic variations cannot be explained by an internal dynamo field of the type that Earth possesses. Instead Europa's subsurface must be behaving as an electrical conductor, responding to the time-varying Jovian magnetic field with an induced field of its own. In this scenario the internal conductor must be as conductive as salty seawater.

Surprisingly, the magnetometer also detected a similar field near Callisto, a satellite with a heavily cratered surface that provides no hint of a subsurface ocean. Signs of such an induced field are detected at Ganymede, too, a moon that also has an intrinsic dynamo field. An exciting possibility is that all the solar system's large icy satellites possess salty oceans within, vestiges of their warmer pasts.

Theory and observation have combined to provide a strong self-consistent case for a global ocean within Europa today. But its existence is not unequivocally proved. Warm subsurface ice could mimic many of the effects of an internal ocean. Although the satellite's surface is sparsely cratered and probably geologically young, searches for definitive evidence of ongoing geologic activity have been fruitless. Important issues such as the thickness of the ice crust are unresolved. Clearly, the next step is to return a spacecraft to Europa and this time go into orbit.

That is just what NASA is planning to do. The Jupiter Icy Moons Orbiter mission could be launched as early as 2011 and would enter Jupiter's orbit seven years later. The spacecraft would orbit Callisto, then Ganymede, then Europa. Precise

BULL'S-EYE SHAPE of the impact site known as Tyre—one of the few large craters on Europa—hints at the presence of liquid water. The central depression is about 40 kilometers across. The surrounding rings are fractures that were created as the crater collapsed inward. Small secondary craters, gouged by debris from the impact, pock the surface. The black lines mark gaps in the data.



ORBITAL CHOREOGRAPHY of Jupiter's largest satellites coerces Europa into an elliptical orbit. As a result, Jupiter's tidal forces, which stretch Europa into an oblong shape, are stronger at some points in the moon's orbit than at others. (The orbits, moon, planet and tidal effects are not to scale.)

NASA/JPL; PAUL GEISSLER AND MOSES MILAZZO *University of Arizona* (*top*); TOM MOORE (*bottom*)

THE LAKE THAT TIME FORGOT

By Frank D. Carsey and Joan C. Horvath

If ever there were a middle of nowhere, Lake Vostok in Antarctica would be it. To get there, one would first have to go to the eponymous Russian scientific base, a place famed for its climate—widely regarded as the world's worst. Then one would have to drill four kilometers straight down. There, cut off from the outside world for the past several million years, is a body of fresh water roughly the extent of Lake Ontario and twice as deep. It may be the closest thing on Earth to the putative ocean of Europa.

Researchers only fully comprehended the true size of the lake in 1996, after the smooth expanse of its icy roof had been probed by the European Remote Sensing Satellite. So far no one has drilled into it, although plans are afoot. The Lamont-Doherty Earth Observatory group performed a thorough radar study in 2000–2001 and tracked the source and fate of lake water that froze to the bottom of the ice sheet and was removed by ice motion.

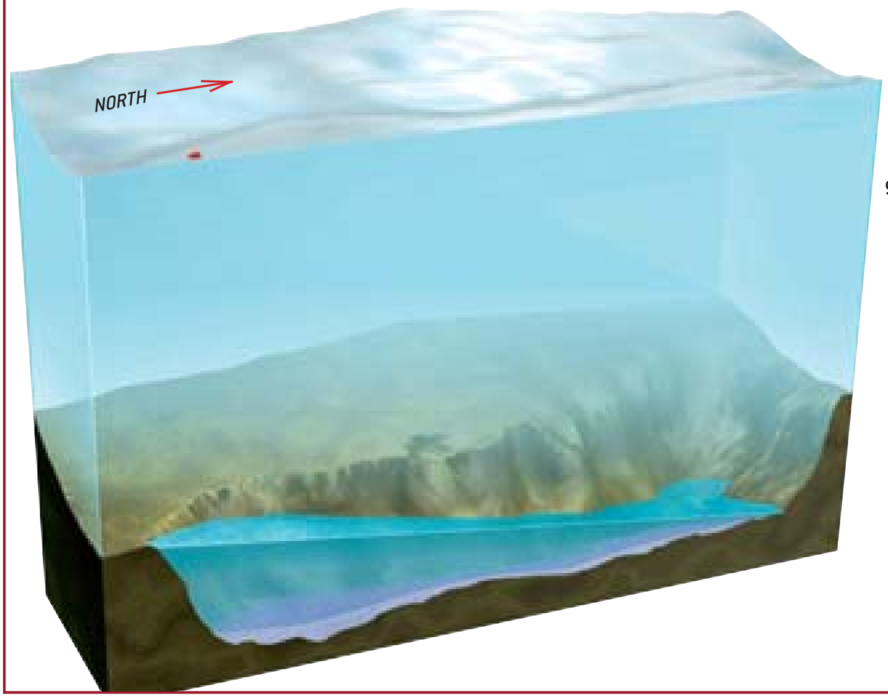
The top of the ice is at about 3,700 meters (12,000 feet) altitude, and the lake surface itself is just below sea level. Judging from the contours of the surrounding bedrock, the lake basin may be a tectonic rift—a ruptured area of Earth like those filled by Lake Baikal and the Red Sea. And why is water there, rather than simply more ice? Some geologic evidence suggests the presence of a hot spot similar to (but smaller than) that responsible for building the Hawaiian Islands. But even without a hot spot, the trickle of heat from Earth's interior is sufficient to reach the local melting point because of the insulating effect of the ice. In fact, under-ice lakes are common in Antarctica; Vostok is simply the largest.

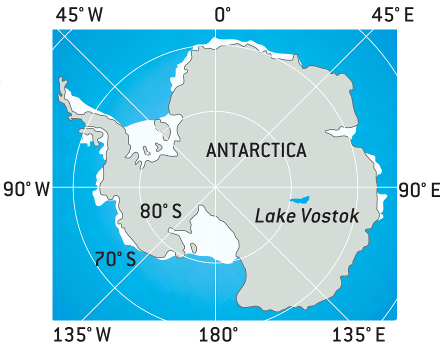
Other evidence also revealed how precious Vostok's pristine waters could be for science. Microbes were turning up in harsh environments—around deep-sea volcanic vents, in shallower ice-covered Antarctic lakes, in alkaline lakes—that had only one thing in common: liquid water. Meanwhile the Galileo spacecraft began finding that Europa might have its own ocean under the ice. The depth of the ice cover on Vostok and Europa is similar; except for the lower pressures on Europa (its gravity is one seventh as strong as Earth's), conditions could be comparable. If life could colonize Lake Vostok, then maybe it could find a niche in Europa.

In 1996 we and others at the Jet Propulsion Laboratory proposed exploring both Lake Vostok and Europa using the same basic approach. Vostok would benefit from technology developed for Europa, whereas a Europan explorer could go through its paces near to home. We, as well as international experts, have examined the possibility of a pair of devices: a "cryobot," which melts its way through the ice, and a small submarine, or "hydrobot," which searches for life and makes other measurements.

Needless to say, the design will be a challenge. The high pressures in the subsurface seas seem to demand a large and heavily armored hydrobot, but a large hydrobot would be difficult to send to Europa. The hydrobot must be autonomous and able to respond to a complex environment with cracks, rocks and so on. Its tiny onboard chemistry laboratories must survey the environment and search out microbes, even if they are utterly unlike those seen elsewhere. And both devices must be fully sterilized so they do not contaminate the water with commonplace microbes. Meeting all these demands is beyond the state of the art in ice coring and miniature submersibles. But engineers are optimistic about this remarkable lake.

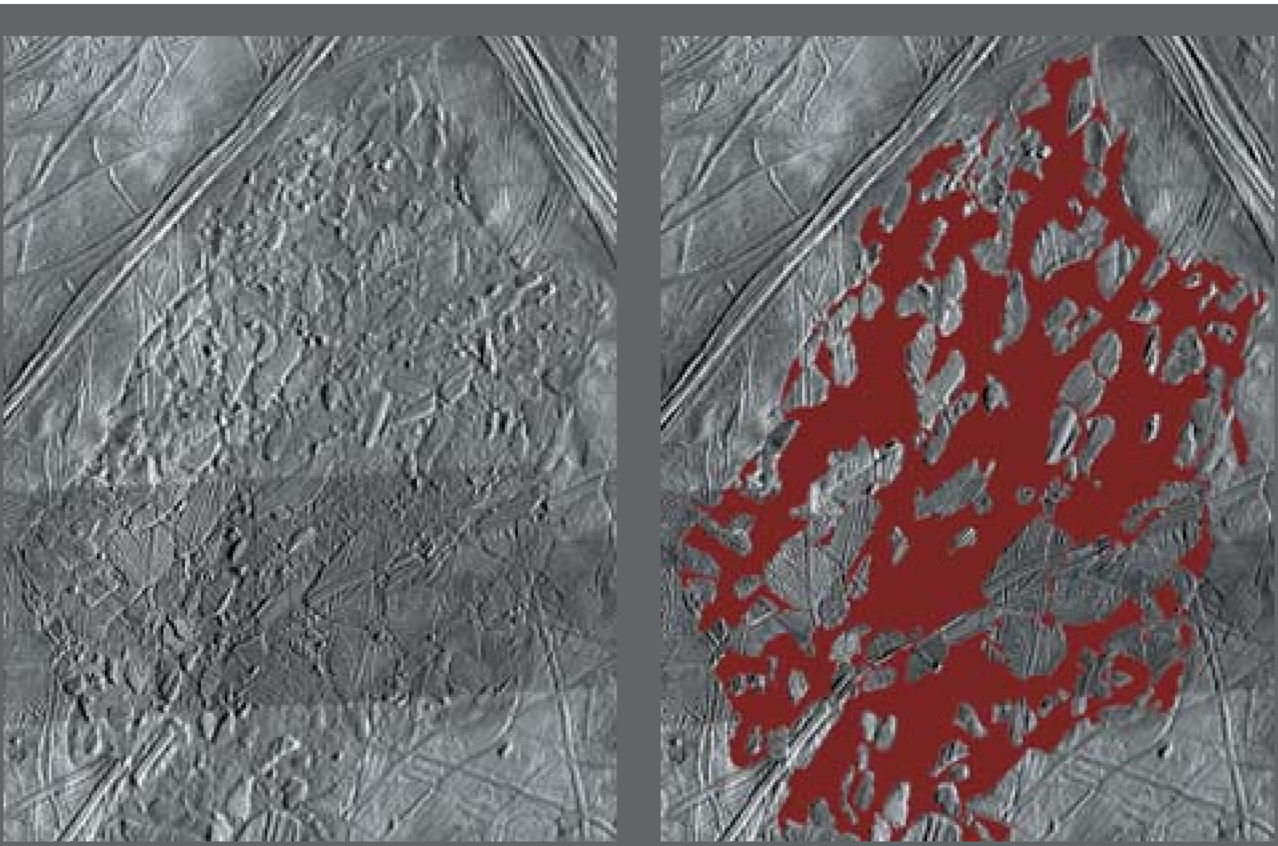
FRANK D. CARSEY and JOAN C. HORVATH lead the Europa/Lake Vostok Initiative at the Jet Propulsion Laboratory in Pasadena, Calif.





LAKE VOSTOK is tucked away in an East Antarctic rift valley and covered by about four kilometers of slowly moving ice. Above the lake, the ice floats much as an iceberg does (*left*); to support a slight rise of the ice surface, the lake surface slopes down by about 400 meters from south to north. On the bottom may be sediments. The Russian Vostok station is directly above the south end (*red dot*). (The vertical scale is distorted.)

TOM MOORE; SOURCE: FRANK D. CARSEY AND MARTIN J. SIEGERT *University of Bristol*



LIKE A JIGSAW PUZZLE, Europa's Conamara Chaos can be pieced back together. The Galileo spacecraft saw a jumble of ice blocks that had been jostled and twisted within a frozen icy matrix (*left*). In their reconstruction (*right*), scientists have identified the matrix (*red*) and restored the ridge-topped ice blocks as best as possible to their original positions. But still, more than half of the pieces have gone missing, converted into matrix. The gnarled region testifies to the compelling geologic vivacity of Europa.

NASA/JPL; CYNTHIA PHILLIPS *University of Arizona* (*left*); NICOLE SPAUN *Brown University* (*right*)

tracking of its position and altitude would map the gravitational field and shape of Europa and other icy moons in enough detail to track the ebb and flow of tides as the moons trundle around Jupiter. If Europa does have a subsurface sea, the moon's surface should rise and fall 30 meters every 3.6-day orbit; otherwise the tidal bulge will change by just one meter. In this way, the Europa orbiting spacecraft would provide the definitive test for an ocean.

Meanwhile the spacecraft's camera would photograph each satellite in turn, and its radar would probe the subsurface for any shallow melt zones. Depending on the ice temperature and purity, the radar signal might even be able to penetrate Europa's ice shell to detect an ocean beneath, in the way Antarctica's Lake Vostok was recently mapped by radar below four kilometers of cold glacial ice [*see box on opposite page*].

Life as we know and understand it requires three basic ingredients: energy, carbon and liquid water. More than any other moon in our solar system, Europa could have all three. Tidal flexing may heat the rocky mantle and lead to volcanism on Europa's ocean floor. At volcanic regions on Earth's ocean floors, water circulates through hot rock and emerges rich in chemical nutrients. Biological communities thrive at these warm oases. Another possible supply of chemical energy comes from above. Fast-moving particles trapped in Jupiter's magnetic field constantly slam into Europa, converting ice and other compounds into free oxygen and other oxidants. If delivered to the ocean by geologic processes, these chemicals could possibly power life. The available chemical-energy resources would be very limited. Although microbial life might make do, biologically complex and diverse organisms of the type that inhabit Earth's oceans probably could not.

If a future orbiter mission confirms the existence of a subsurface ocean, the next logical step would be to examine the surface in situ. A small robotic lander could analyze a scoopful of ice for organic compounds. Ultimately, it may be possible for a robotic submarine to melt a path through the ice shell. Europa's briny waters, now surmised only by indirect means, would then be known firsthand. It might turn out that we are not alone in the solar system after all.

MORE TO EXPLORE

Geology of Europa. Baerbel K. Lucchitta and Laurence A. Soderblom in *Satellites of Jupiter.* Edited by David Morrison. University of Arizona Press, 1982.

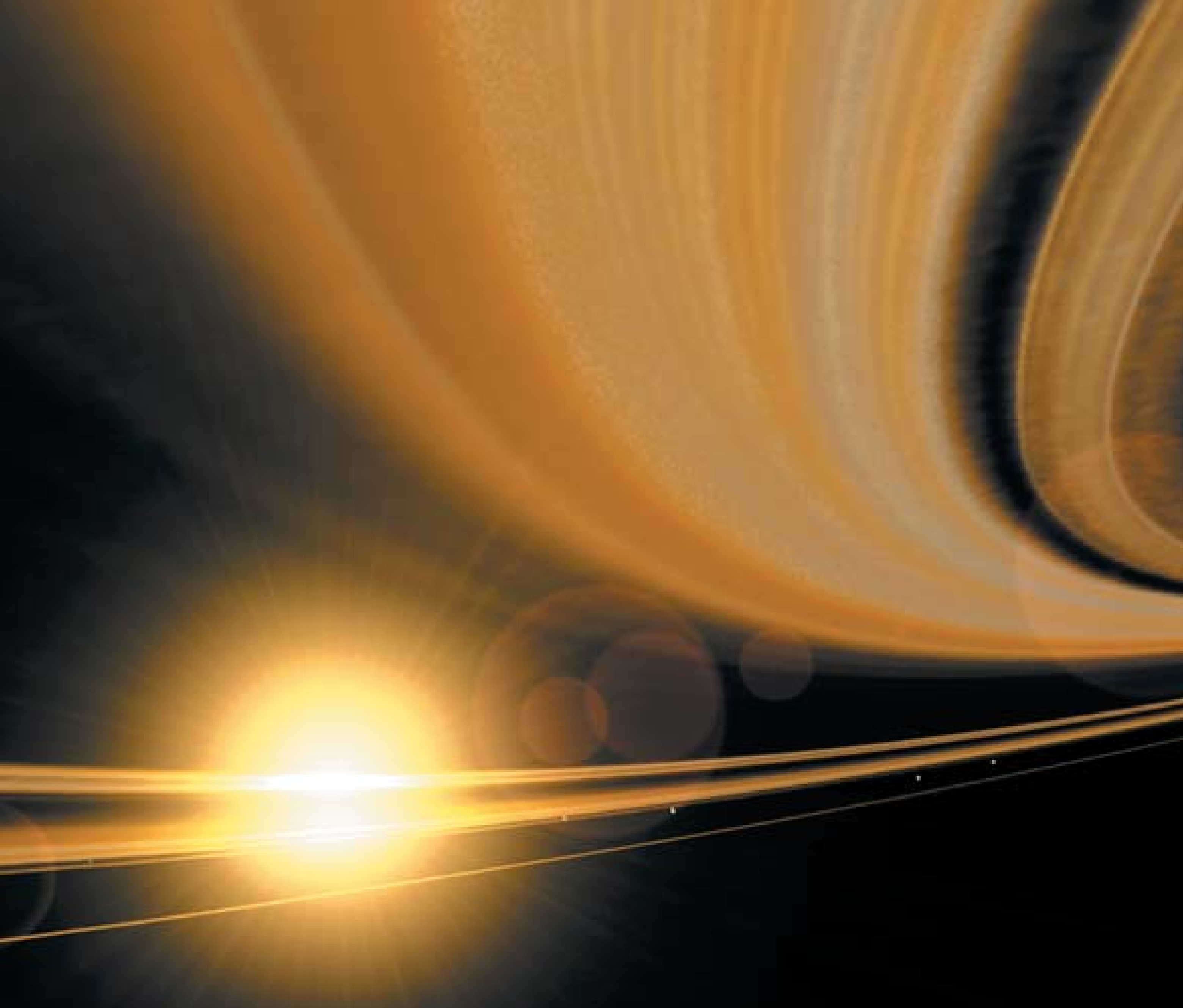
Does Europa Have an Ocean? Evaluation of the Geological Evidence. Robert T. Pappalardo et al. in *Journal of Geophysical Research—Planets,* Vol. 104, pages 24015–24055; October 1999.

Europa: Initial Galileo Geological Observations. Ronald Greeley et al. in *Icarus,* Vol. 135, pages 4–24; September 1998.

The New Solar System. Fourth edition. Edited by J. Kelly Beatty, Carolyn Collins Petersen and Andrew Chaikin. Cambridge University Press, 1998.

The Galileo project site is at **www.jpl.nasa.gov/galileo**

The Jupiter Icy Moons Orbiter site is **www.jpl.nasa.gov/jimo**



DANCING ALONG THE RINGS OF SATURN, small moons (*dots*) are responsible for much of the system's intricacy. The sun shines through the Cassini Division—a prominent gap visible from amateur telescopes on Earth—and glints off the faint outer rings.

Bejeweled

By Joseph A. Burns, Douglas P. Hamilton and Mark R. Showalter

Updated from the February 2002 issue

Worlds

What an impoverished universe it would be if Saturn and the other giant planets lacked rings. Planetary scientists are finally working out how gravity has sculpted these elegant ornaments

Much of the modern world's economy is based on inventions made possible by 19th-century physicist James Clerk Maxwell, father of electromagnetism and pioneer of thermodynamics. In terms of raw economic benefit, though, not much can be said for another of Maxwell's favorite subjects: the rings of Saturn. Apart from inspiring the sales of executive desk toys, planetary rings do not contribute conspicuously to the material wealth of nations. And yet that does not blunt their appeal. In his 1857 Adams Prize essay, Maxwell wrote:

> There are some questions in Astronomy to which we are attracted ... on account of their peculiarity ... [rather] than from any direct advantage which their solution would afford to mankind.... I am not aware that any practical use has been made of Saturn's Rings ... [b]ut when we contemplate the Rings from a purely scientific point of view, they become the most remarkable bodies in the heavens, except, perhaps, those still less *useful* bodies—the spiral [galaxies].... When we have actually seen that great arch swung over the equator of the planet without any visible connection, we cannot bring our minds to rest.

A century and a half later Saturn's rings remain a symbol of all that is exotic and wondrous about the universe. Better observations have only heightened their allure. The findings of the past two decades have so overturned previous knowledge that essentially a new ring system—one much more complex and interesting than theory, observation or imagination had suggested—has been revealed.

Other giant planets besides Saturn have rings, and no two systems look alike. Rings are strange, even by the standards of astronomy. They are sculpted by processes that can be feeble and counterintuitive. For example, in rings, gravity can effectively repel material. We now appreciate that rings, once thought to be static, are continually evolving. We have seen the vital symbiosis between satellites and rings. Most important, we have recognized that planetary rings are more than just exquisite phenomena. Like Maxwell, modern scientists see analogies between rings and galaxies; in a very fundamental way, rings may also afford a glimpse into the solar system's ancient beginnings.

Saturn's rings, initially spied in 1610 by Galileo Galilei and interpreted as a planet-encircling hoop five decades later by Christiaan Huygens, stood alone for more than three and a half centuries. Then, in a span of just seven years, rings were discovered around the other three giant planets. Uranus's were detected first, in 1977. James L. Elliot, then at Cornell University, monitoring a star's brightness as Uranus crossed in front of it, noticed the signal blinking on and off. He inferred that a series of narrow bands, slightly elliptical or inclined, circumscribe the planet [see "The Rings of Uranus," by Jeffrey N. Cuzzi and Larry W. Esposito; SCIENTIFIC AMERICAN, July 1987]. In 1979 the Voyager 1 spacecraft sighted Jupiter's diaphanous rings. Finally, in 1984, a technique like Elliot's detected pieces of rings—but not full rings—around Neptune.

Those heady days passed, and ring research stagnated until the mid-1990s. Since then, a new era of ring exploration has begun. Observations have poured in from the Hubble Space Telescope, ground-based telescopes and the Galileo probe in orbit around Jupiter. Saturn's faintest rings and satellites became visible in 1995 and 1996, when the positions of Earth and Saturn made the system appear edge-on, thereby reducing the glare from the main rings. And in July 2004 the Cassini spacecraft will begin its highly anticipated four-year tour of the Saturnian system.

Four-Ring Circus

ALTHOUGH THE FOUR KNOWN ring systems differ in detail, they share many general attributes. They are all richly textured, made up of multiple concentric rings often separated by gaps of various widths. Each ring is composed of innumerable

THE AUTHORS

JOSEPH A. BURNS, DOUGLAS P. HAMILTON* and *MARK R. SHOWALTER started working together at Cornell University, where Burns is a professor and Hamilton and Showalter were graduate students. Burns studied naval architecture in college but then got caught up in the excitement of the space age and changed fields. He is now I. P. Church Professor of Engineering and Astronomy. Hamilton, a professor at the University of Maryland, received the 1999 Urey Prize of the American Astronomical Society for his studies of the celestial mechanics of dust. Showalter is a researcher at Stanford University, where he oversees NASA's archive of planetary ring data. All three authors are deeply involved in space missions to the outer planets.

DON DIXON (*preceding pages*); DON DIXON (*drawing*); NASA/JPL (*spacecraft images*) (*opposite page*)

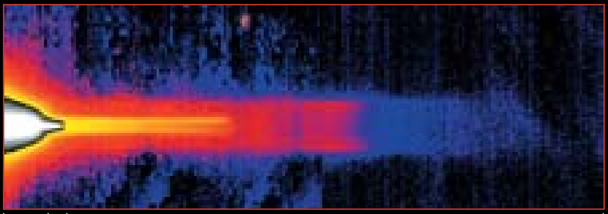
Jupiter

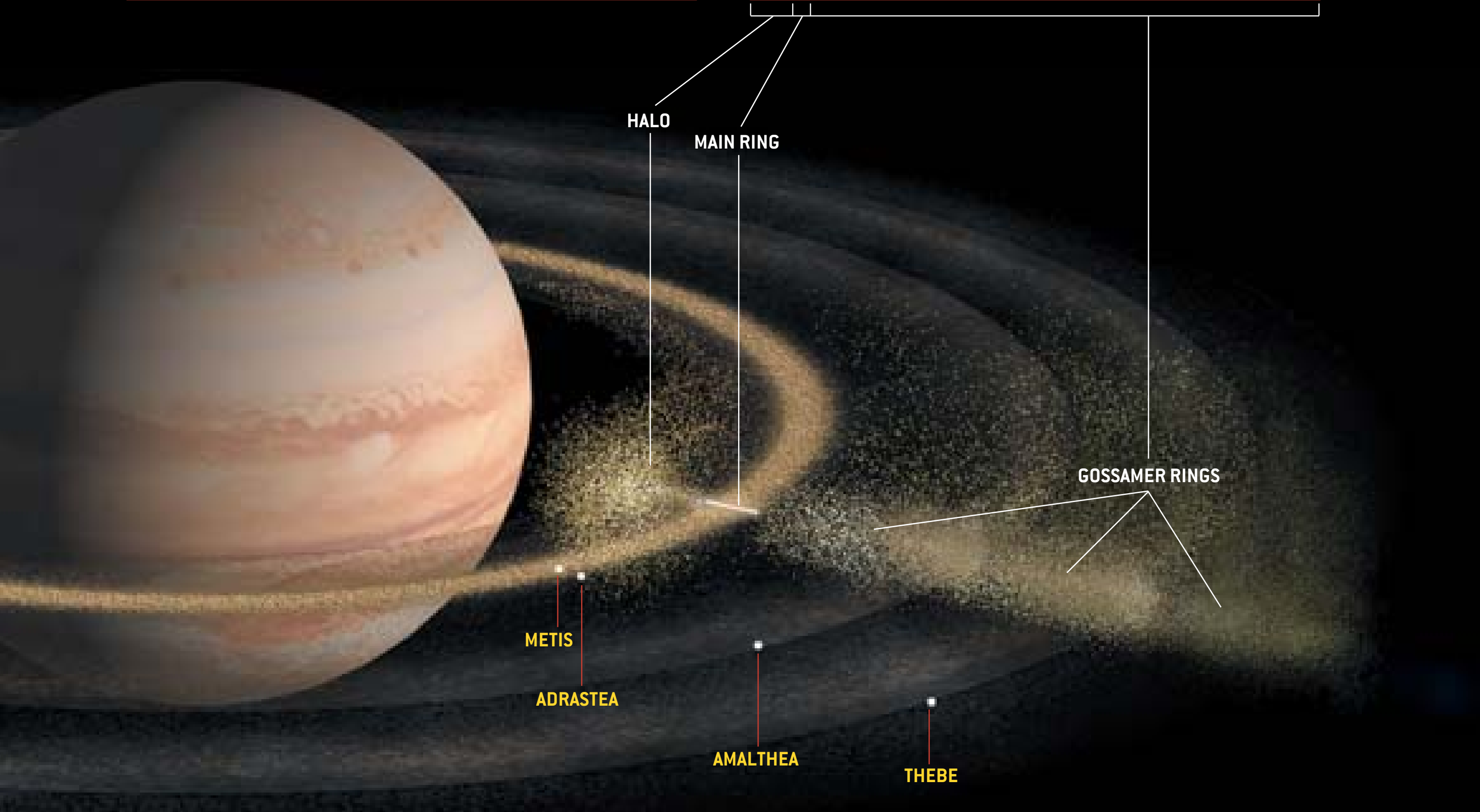
The largest planet in the solar system has rings of puzzling subtlety. They are composed of finer particles and are less flattened than the rings around other planets.

This mosaic by the Galileo spacecraft shows Jupiter in eclipse, highlighting its upper atmosphere and rings.

A tenuous, puffy halo rises up from the main ring's inner edge.

Faint gossamer rings (*yellow, red and blue bands*) extend beyond the main ring and halo (*black-and-white blob at left*).





Saturn

Saturn's rings, the most baroque, seem to get more complicated the closer scientists look. The famous Voyager images may pale in comparison to what the Cassini spacecraft finds in 2004.

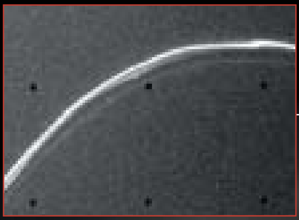
A RING (*artist's conception*)

In at least one place, the meter-size snowballs are cleared away by satellites.

ENCKE GAP

Pried open by the tiny satellite Pan.

F RING

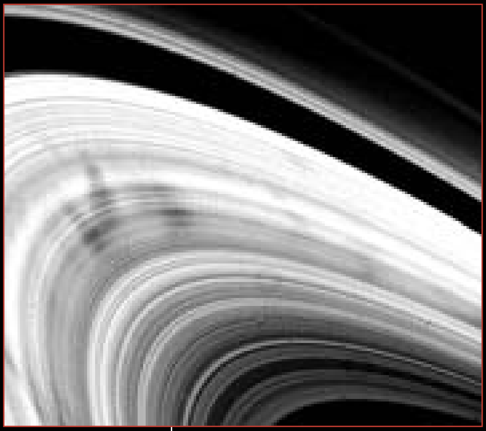


The multiple strands are knotted by the tugs of two nearby satellites.

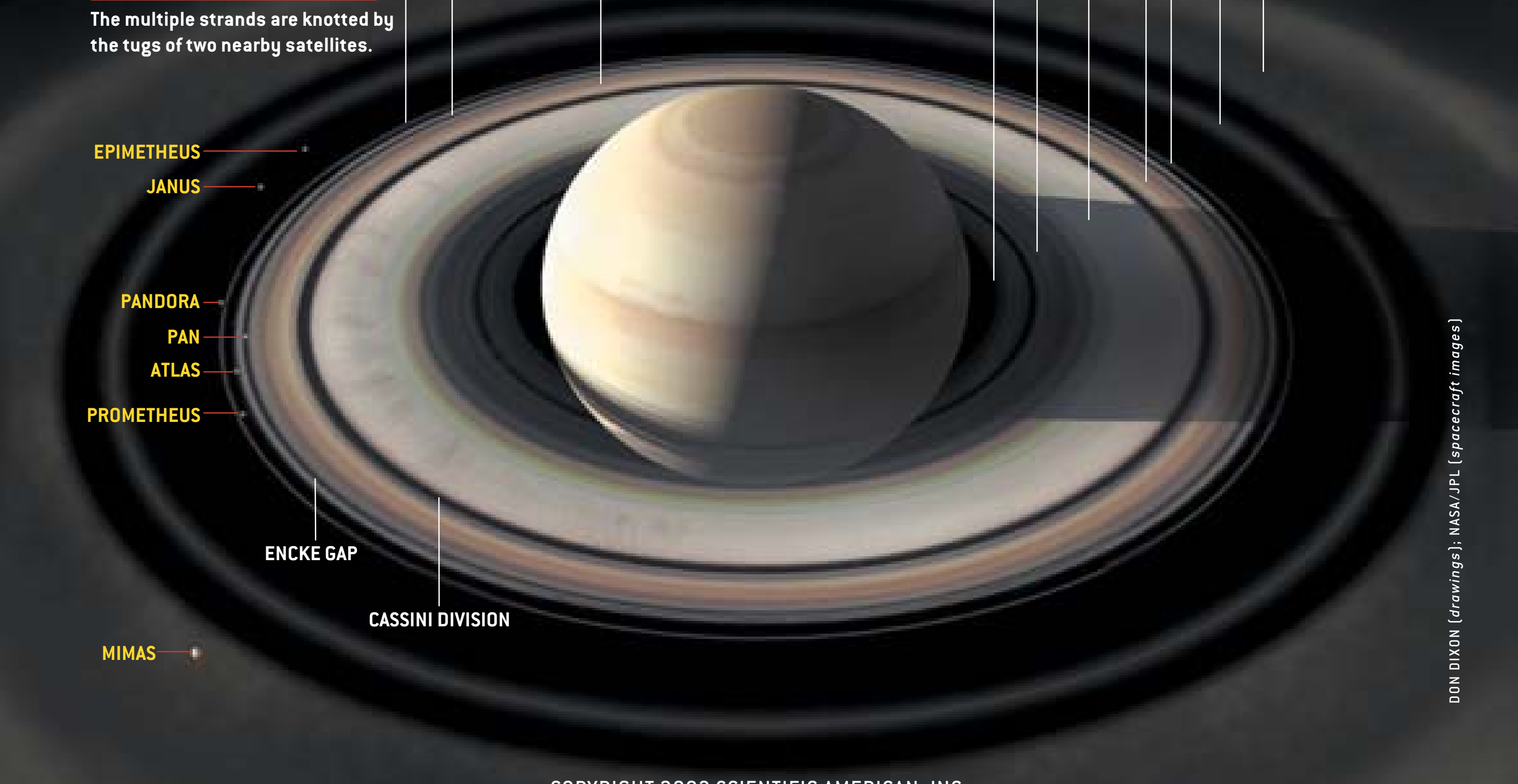
C RING

This image exaggerates the slight color differences between the C ring (*blue*) and B ring (*gold*).

B RING



The "spokes" are fleeting smudges made of levitating dust grains. The innumerable ringlets remain unexplained.



DON DIXON (*drawings*); NASA/JPL (*spacecraft images*) particles—chunks of rock and ice—that independently circle the central planet while gently jostling one another. Rings fall into two general categories based on how densely packed the particles are, as described by the optical depth, a measure of the exponential decay of light as it penetrates perpendicularly through the ring. For the densest rings, such as Saturn's main rings (designated A and B) and the Uranian rings (designated by numbers and Greek letters), the optical depth can be as high as 4, which means that a mere 2 percent of the light leaks through. The most tightly packed of these rings contain particles that range from a few centimeters to several meters in diameter.

Particles in a dense ring system collide frequently, often several times during each orbit around the planet. In the process, energy is lost and angular momentum is redistributed. Because particles nearer to the planet move at a higher speed than do particles farther out, collisions hold back the inner particles (which then fall toward the planet) and push forward the outer ones (which then move away from the planet). Thus, a ring tends to spread radially. But the spreading takes time, and in this regard, a ring may be thought of as a viscous fluid that slowly diffuses inward and outward. Saturn's rings have an effective kinematic viscosity like that of air.

The energy loss, combined with angular-momentum redistribution, causes a dense ring system to flatten. Whatever its initial shape, the system quickly becomes a thin, near-equatorial disk. Saturn's rings are only tens of meters from top to bottom even though they stretch across several hundred thousand kilometers; they are proportionally as thick as a sheet of tissue paper spread over a football field. A similar effect flattens the debris disks around stars and the gaseous disks of spiral galaxies.

RESONANCE between a satellite and a ring particle means that their two orbits are choreographed: in this case, the particle goes around exactly twice in the time it takes the satellite to trundle around once. Because the bodies always encounter each other at the same position, gravitational tugs can add up.

DON DIXON

Another consequence of dense packing is to strengthen the particles' own mutual gravitational attraction. This may be why Uranus's rings are slightly out of round: their self-gravity resists the tendency to smear into a circular band.

At the other extreme, the faintest known rings, such as Jupiter's rings and Saturn's outermost rings, have optical depths between 10^{-8} and 10^{-6}. Particles are as spread out as baseball outfielders. Because they collide infrequently, they tend not to settle into a flat disk. As we know from how these rings scatter light, the particles are fine dust, typically microns in size, comparable to the size of smoke particles. So these structures are literally smoke rings. The particles display unusual dynamics because, being so small, they are significantly affected by electromagnetic and radiation forces in addition to gravity.

Neptune's rings do not fall into this neat dichotomy; their optical depth lies between the two extremes. The Neptunian system is anomalous in other respects as well. Its densest ring is not a smooth band; it contains discontinuous arcs that together encompass less than a tenth of the circumference. Without some confinement mechanism at work, these structures should spread fully around the planet in about a year. Yet recent Hubble images and ground-based observations find that the positions of the arcs have shifted little in 15 years.

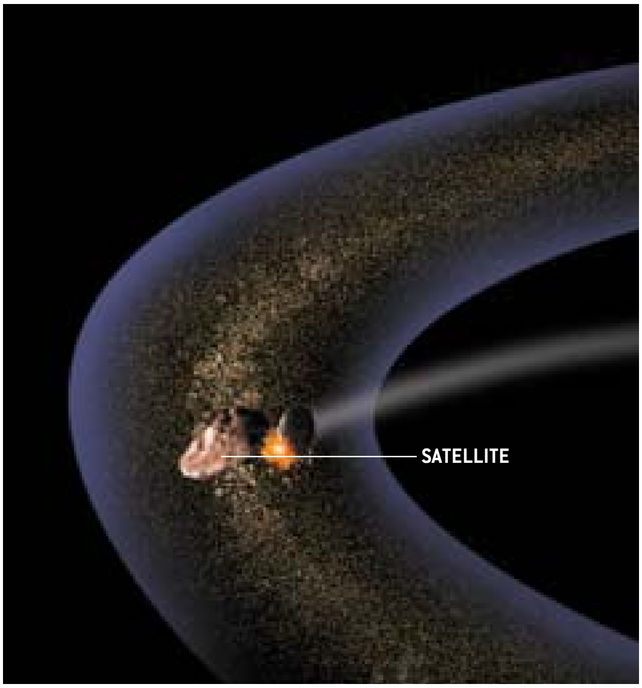
Lords of the Rings

ALL DENSE RING SYSTEMS nestle close to their planets, extending no farther than the so-called Roche limit, the radius within which the planet's tidal forces overwhelm the tendency of ring particles to agglomerate into larger bodies. Just outside the Roche limit is a zone where small, irregularly shaped moons can coexist with the rings. The interactions between rings and ring moons are implicated in many of the strangest aspects of rings.

For example, Saturn's E ring reaches across a broad region that encompasses the satellites Mimas, Tethys, Dione and Rhea, peaking in brightness at the orbit of the smooth, icy moon Enceladus. The narrow F ring, a tangle of several lumpy strands, sits isolated just beyond Saturn's A ring and is also straddled by two moons, Pandora and Prometheus. Correlations of satellite positions and ring features occur in the Jovian, Uranian and Neptunian systems as well.

Explaining how satellites wield such power has been the major advance in ring science over the past two decades. Three basic processes appear to be at work. The first is the orbital resonance, a tendency of gravitational forces to be magnified at positions where a particle's orbital period matches an integer ratio (say, $m:n$) of a satellite's orbital period. For instance, a particle at the outer edge of Saturn's B ring is in a 2:1 resonance with Mimas, meaning that it goes around the planet precisely twice for each lap the satellite completes. In another example, the exterior boundary of Saturn's A ring is in a 7:6 resonance with the satellites Janus and Epimetheus.

Orbits that lie near resonant locations suffer unusually large distortions because the gentle tugs of moons are repeated systematically and therefore build up over time. Resonances are stronger for particles in orbits near a moon, but when the orbits are too close, different resonances vie for control, and motions become chaotic. Resonances are strongest when $m = n + 1$ (for example, 2:1 or 43:42) and weaken rapidly as m and n differ more and more. Throughout Saturn's enormous rings, only a few dozen ring locations may respond to strong satellite resonances.



IF SOMETHING SLAMS into a satellite, material flies off and becomes part of a ring. Conversely, the satellite steadily sweeps up material. The balance of these competing effects determines the density of faint rings.

DON DIXON

The outcome of these resonant perturbations varies. Strong ones clear material, accounting for the outer edges of Saturn's A and B rings. In some places, gaps are opened. Such a resonance may account for Neptune's discontinuous ring. Analogous resonances explain the distribution of material in the asteroid belt, for which the sun plays the role of the planet and Jupiter plays the role of the satellite.

Elsewhere in the A ring, resonances generate waves. If the satellite has an elliptical orbit, the result is a spiral wave, a miniature version of the pinwheel pattern of our galaxy. If the satellite has a tilted orbit, the result is a series of vertical bending waves, an out-of-plane corrugation—small ripples in a cosmic carpet.

Although resonances typically involve satellites, any force that repeats periodically at an integer ratio of the orbital period—such as lumpy planetary gravitational fields or variable electromagnetic forces—will be similarly effective. The Jovian system has become infamous for such resonances. Inward of a radius of 120,000 kilometers, the ring abruptly puffs up from a flat disk to a thick torus. A ring particle at that radius orbits three times for every two planetary spins; thus, the planet's tilted magnetic field pushes it ever upward. Still closer to the planet, at a radius of 100,000 kilometers, the brightness of the Jovian ring drops sharply. That happens to be the location of the 2:1 electromagnetic resonance. Particles that drift to this position are spread so thinly that they vanish against the giant planet's glare.

Ringmaster

THE SECOND BASIC WAY that satellites govern ring structures is by influencing the paths of ring particles. The gravitational interaction of a satellite and a nearby particle is somewhat counterintuitive. If these two bodies were isolated in deep space, their close encounters would be symmetrical in space and time. The particle would approach the satellite, accelerate, zip around, emerge on the other side and decelerate (assuming that it did not collide). The departure leg would be the mirror image of the inbound path (a hyperbola or parabola). Although the particle would have changed direction, it would eventually return to its original speed.

In a ring system, however, a satellite and particle are not isolated—they are in orbit around a third object, the planet. Whichever body is nearer to the planet orbits faster. Suppose it is the particle. During the close encounter, the gravity of the satellite nudges the particle into a new orbit. The event is asymmetrical: the particle moves closer to the satellite, and the gravitational interaction of the two bodies strengthens. So the particle is unable to regain the velocity it once had; its orbital energy and angular momentum have decreased. Technically, that means its orbit is distorted from a circle to an ellipse of slightly smaller size; later, collisions within the ring will restore the orbit to a circle, albeit a shrunken one.

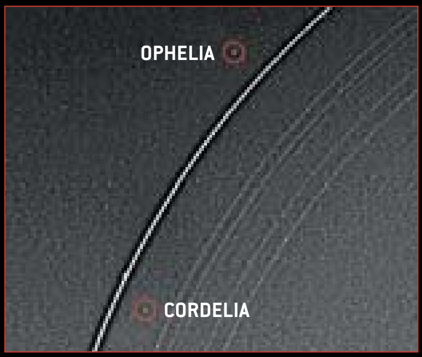
The net effect is that the particle is pushed inward. Its loss is the satellite's gain, although because the satellite is more massive, it moves proportionately less. If the positions are reversed, so are the roles: with the satellite on the inside, the particle will be pushed outward and the satellite inward. In both cases, the attractive gravity of a satellite appears to *repulse* ring material. None of Newton's laws have been broken; this bizarre outcome occurs when two bodies in orbit around a third interact and lose energy. (It is completely different from the "repulsive" gravity that occurs in theories of the expanding universe.)

Like resonances, this mechanism can pry open gaps in rings. The gaps will grow until the satellite's repulsive forces are counterbalanced by the tendency of rings to spread viscously through collisions. Such gaps are present within Saturn's A, C and D rings, as well as throughout the Cassini division, a zone that separates the A and B rings.

Conversely, the process can squeeze a narrow ring. Satellites on either side of a strand of material can shepherd that material, pushing back any particles that try to escape. In 1978 Peter Goldreich and Scott D. Tremaine, then both at the California Institute of Technology, hypothesized the shepherding process to explain the otherwise puzzling stability of the threadlike rings of Uranus [see "Rings in the Solar System," by James B. Pollack and Jeffrey N. Cuzzi; Scientific American, November 1981]. The satellites Cordelia and Ophelia keep Uranus's ε ring corralled. Saturn's F ring appears to be herded by Prometheus and Pandora. To be sure, most of the visible gaps and narrow

Uranus

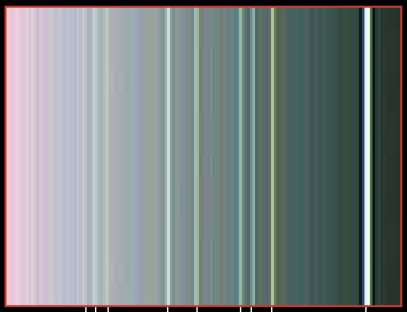
What makes the rings of Uranus so odd is that most of them are slightly elliptical and tilted. Somehow they have resisted the forces that would have circularized and flattened them.

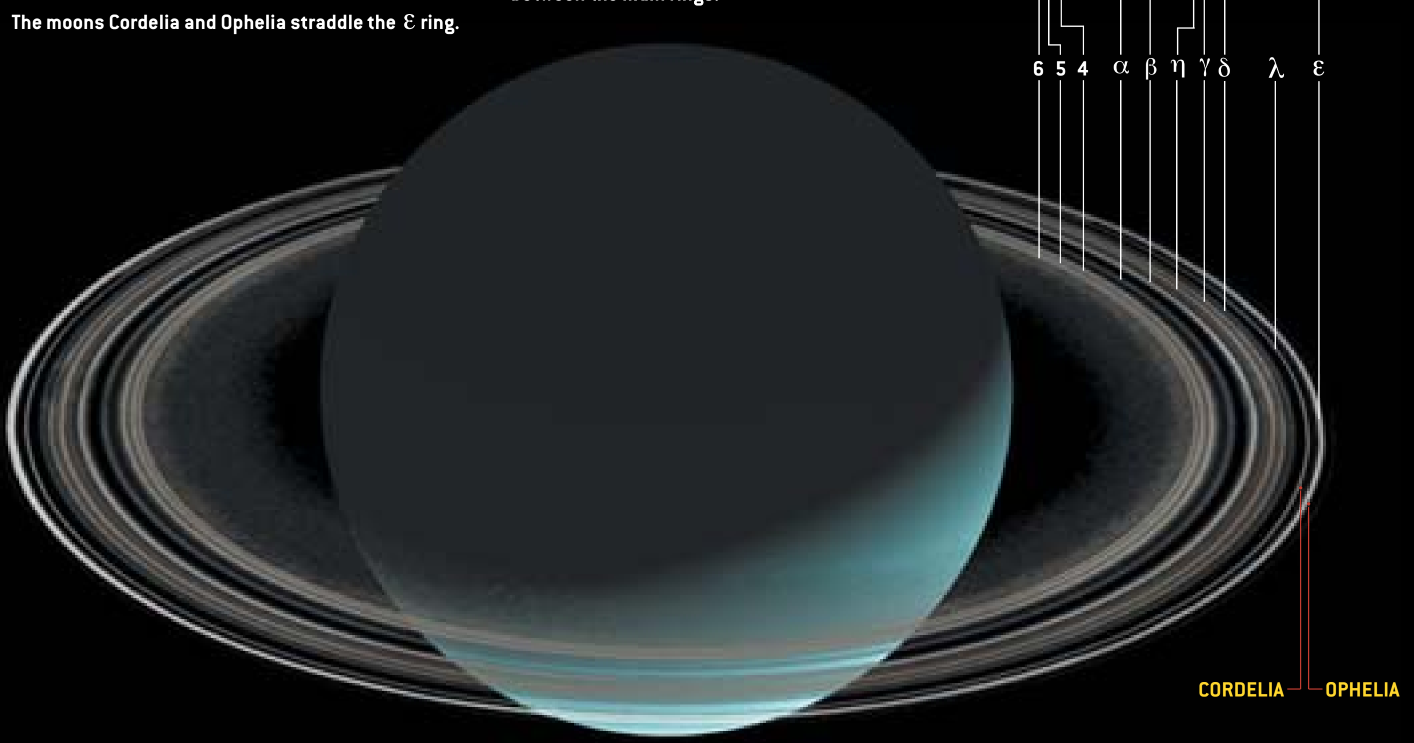


The moons Cordelia and Ophelia straddle the ε ring.

Using a different viewing angle and longer exposure, Voyager saw dust between the main rings.

This false-color image hints at differing particle properties. The dusty λ ring is too faint to see here.





ringlets remain unexplained. Perhaps they are manipulated by moons too small to see with present technology. The Cassini orbiter may be able to spy some of the hidden puppeteers.

Yet another effect of repulsive gravity is to scallop ring edges. These undulations are easiest to understand from the vantage point of the satellite. In rings, a continuous stream of particles flows past the satellite. When these particles overtake the moon, gravity modifies their circular orbits into elliptical ones of almost the same size. The particles no longer maintain a constant distance from the planet. Someone riding on the satellite would say that the particles have started to weave back and forth in concert. The apparent motion is sinusoidal with a wavelength proportional to the distance between the orbits of the satellite and the particle.

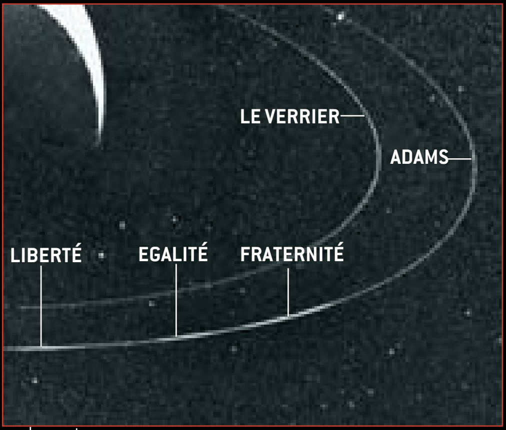
The resulting wave appears behind the satellite if the particle is on the outside and in front of the satellite if the particle is on the inside. It is akin to the wake of a boat in an unusual river where the water on one side of the boat moves faster than the boat itself. One of us (Showalter) analyzed the scalloped edges of Saturn's Encke division to pinpoint a small satellite, Pan, that had eluded observers. Another example is the F ring, whose periodic clumps seem to have been imprinted by Prometheus.

DON DIXON (*drawing*); NASA/JPL (*spacecraft images*)

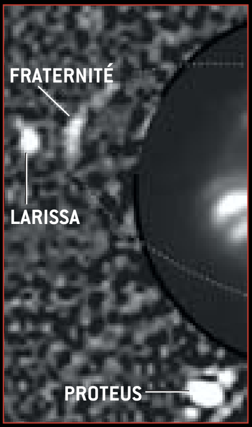
Neptune

The least understood rings are those of Neptune. The outer ring contains clumps—the so-called arcs—but little is known about the fainter inner rings.

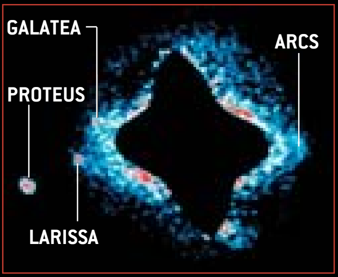
LE VERRIER AND ADAMS RINGS



Voyager images from 1989 reveal clumps in the outermost ring, perhaps the result of a complex satellite resonance.



The ring arcs also appear in this Hubble Space Telescope image from 1998. The arcs have persisted for nearly two decades, perhaps stabilized by periodic gravitational tugs from the nearby satellite Galatea.



Ground-based observations in 1998 agree with Hubble's conclusions.

Those Dirty Rings

THE THIRD AND FINAL effect of moons on rings is to spew out and soak up material. This role, especially vital for faint, dusty rings such as those around Jupiter, has come into clear view only with the Galileo mission to Jupiter. Earlier the Voyager spacecraft had discovered Jupiter's rings as well as two small moons, Adrastea and Metis, close to the main ring's outer edge. But its camera was not sharp enough to tell us what the satellites actually did. Were they shepherds that prevented the rings' outward spread? Or were they the source of ring material that, once placed into orbit, drifted inward? Neither could Voyager make sense of a faint outer extension—a gossamer ring that accompanied the main one.

Galileo's imaging system found that the gossamer ring vanished abruptly beyond the orbit of the moon Amalthea. It discovered another, fainter gossamer ring that extended as far as the moon Thebe and no farther. On the flight home from the meeting at which these images were first available, one of us (Burns) noticed the smoking gun: the vertical extent of the innermost gossamer ring was equal to the orbital tilt of Amalthea, and the thickness of the outer gossamer ring perfectly matched the inclination of Thebe. Furthermore, both gossamer rings were brightest along their top and bottom edges, indicating a pileup of material—which is exactly what one would expect if particles and satellites shared the same orbital tilt. This tight association is most naturally explained if the particles are debris ejected by meteoroid impacts onto the satellites.

Ironically, small moons should be better sources of material than big ones: though smaller targets, they have weaker gravity, which lets more debris escape. In the Jovian system the most effective supplier is calculated to be 10 or 20 kilometers across—just about the size of Adrastea and Metis, explaining why they generate more formidable rings than do Amalthea and Thebe, which are much larger.

An odd counterexample is Saturn's 500-kilometer-wide moon, Enceladus, which appears to be the source of the E ring. Powerful impacts by ring particles, as opposed to interplanetary projectiles, might explain how Enceladus manages to be so prolific. Each grain that hits Enceladus generates multiple replacement particles, so the E ring could be self-sustaining. Elsewhere such collisions usually result in a net absorption of material from the ring.

Ring Out the Old

THE EVIDENT IMPORTANCE of sources and sinks reopens the classic question of whether rings are old and permanent or young and fleeting. The former possibility implies that rings could date to the formation of the solar system. Just as the protosun was surrounded by a flattened cloud of gas and dust out of which the planets are thought to have emerged, each of the giant planets was surrounded by its own cloud, out of which satellites emerged. Close to each planet, within the Roche limit, tidal forces prevented material from agglomerating into satellites. That material became a ring instead.

Alternatively, the rings we see today may have arisen much later. A body that strayed too close to a planet may have been torn asunder, or a satellite may have been shattered by a high-speed comet. Once a satellite is blasted apart, the fragments will reagglomerate only if they lie beyond the Roche limit. Even then, they will be unconsolidated, weak rubble piles susceptible to later disruption.

Several lines of evidence now indicate that most rings are young. First, tiny grains must lead short lives. Even if they survive interplanetary micrometeoroids and fierce magnetospheric plasma, the subtle force exerted by radiation causes their orbits to spiral inward. Unless replenished, faint rings should disappear within only a few thousand years. Second, some ring moons lie very close to the rings, even though the back reaction from spiral density waves should quickly drive them off.

Third, Saturn's icy ring particles should be darkened by cometary debris, yet they are generally bright. Fourth, satellites just beyond Saturn's rings have remarkably low densities, as though they are rubble piles. Finally, some moons are embedded within rings. If rings are primordial material that failed to agglomerate, how did those moons get there? The moons make the most sense if they are merely the largest remaining pieces of a shattered progenitor.

So it seems that rings are not quite the timeless fixtures they appear to be. Luke Dones of the Southwest Research Institute in Boulder, Colo., has suggested that Saturn's elaborate adornments are the debris of a shattered moon roughly 300 to 400 kilometers across. Whether all rings have such a violent provenance, we now know that they were not simply formed and left for us to admire. They continually reinvent themselves. Joshua E. Colwell and Larry W. Esposito of the University of Colorado envision recycling of material between rings and ring moons. Satellites gradually sweep up the particles and subsequently slough them off during energetic collisions. Such an equilibrium could determine the extent of many rings. Variations in the composition, history and size of the planets and satellites would naturally account for the remarkable diversity of rings.

Indeed, the emerging synthesis explains why most of the inner planets are ringless: they lack large retinues of satellites to provide ring material. Earth's moon is too big, and any micron-size dust that does escape its surface is usually stripped away by solar gravitational and radiation forces. Mars, with its two tiny satellites, probably does have rings. But two of us (Hamilton and Showalter) were unable to find any rings or smaller satellites in Hubble observations in 2001. If a Martian ring does exist, it must be exceedingly tenuous, with an optical depth of less than 10^{-8}.

As often happens in science, the same basic principles apply to phenomena that at first seem utterly unrelated. The solar system and other planetary systems can be viewed as giant, star-encircling rings. Astronomers have seen hints of gaps and resonances in the dusty disks around other stars, as well as signs that source bodies orbit within. The close elliptical orbits of many large extrasolar planets are best understood as the end result of angular momentum transfer between these bodies and massive disks [see "Migrating Planets," by Renu Malhotra; SCIENTIFIC AMERICAN, September 1999]. Planetary rings are not only striking, exquisite structures; they may be the Rosetta stones to deciphering how planets are born.

MORE TO EXPLORE

Observations of Saturn's Ring-Plane Crossings in August and November 1995. Philip D. Nicholson et al. in *Science,* Vol. 272, pages 509–515; April 26, 1996.

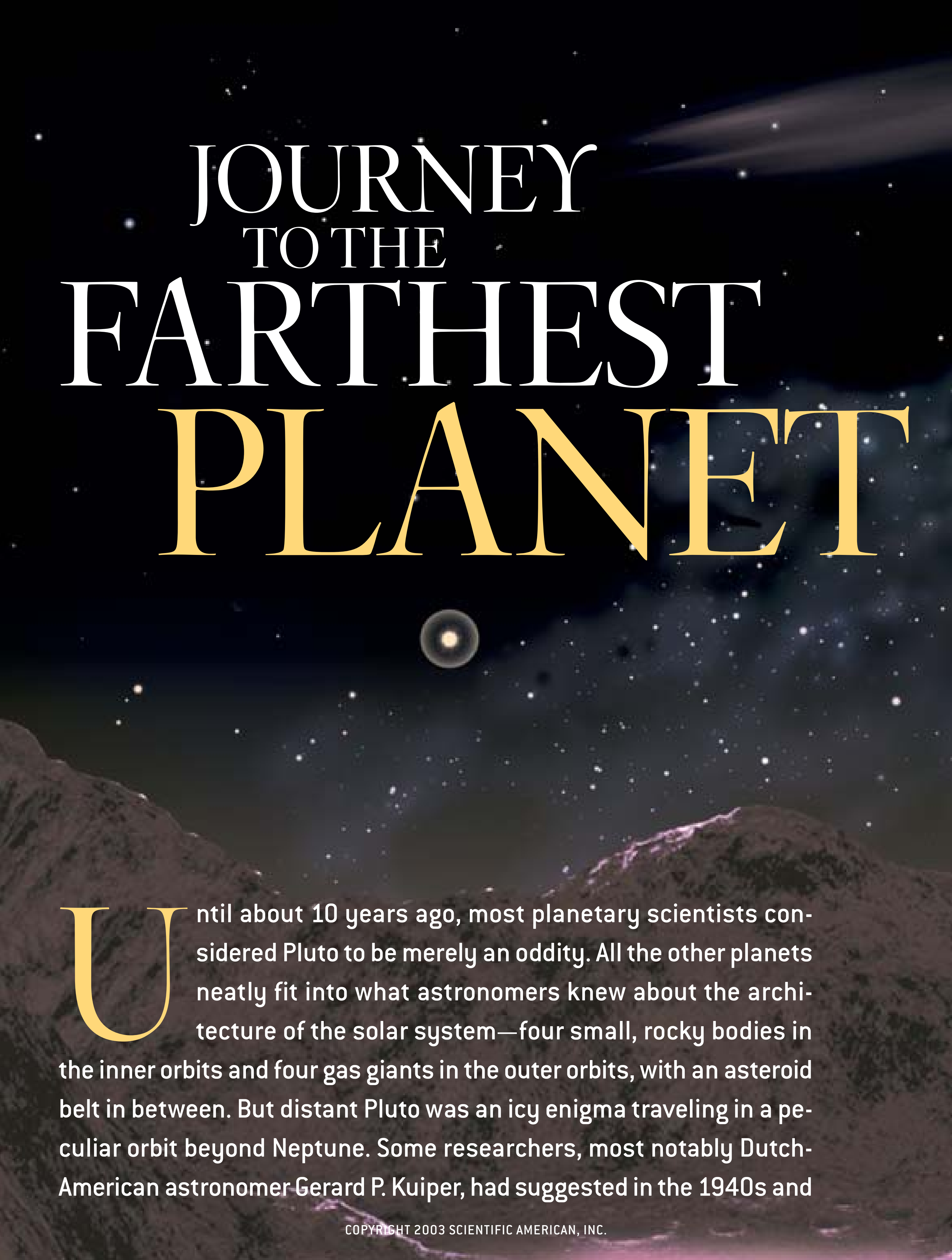
The Formation of Jupiter's Faint Rings. Joseph A. Burns et al. in *Science,* Vol. 284, pages 1146–1150; May 14, 1999.

Stability of Neptune's Ring Arcs in Question. Christophe Dumas et al. in *Nature,* Vol. 400, pages 733–735; August 19, 1999.

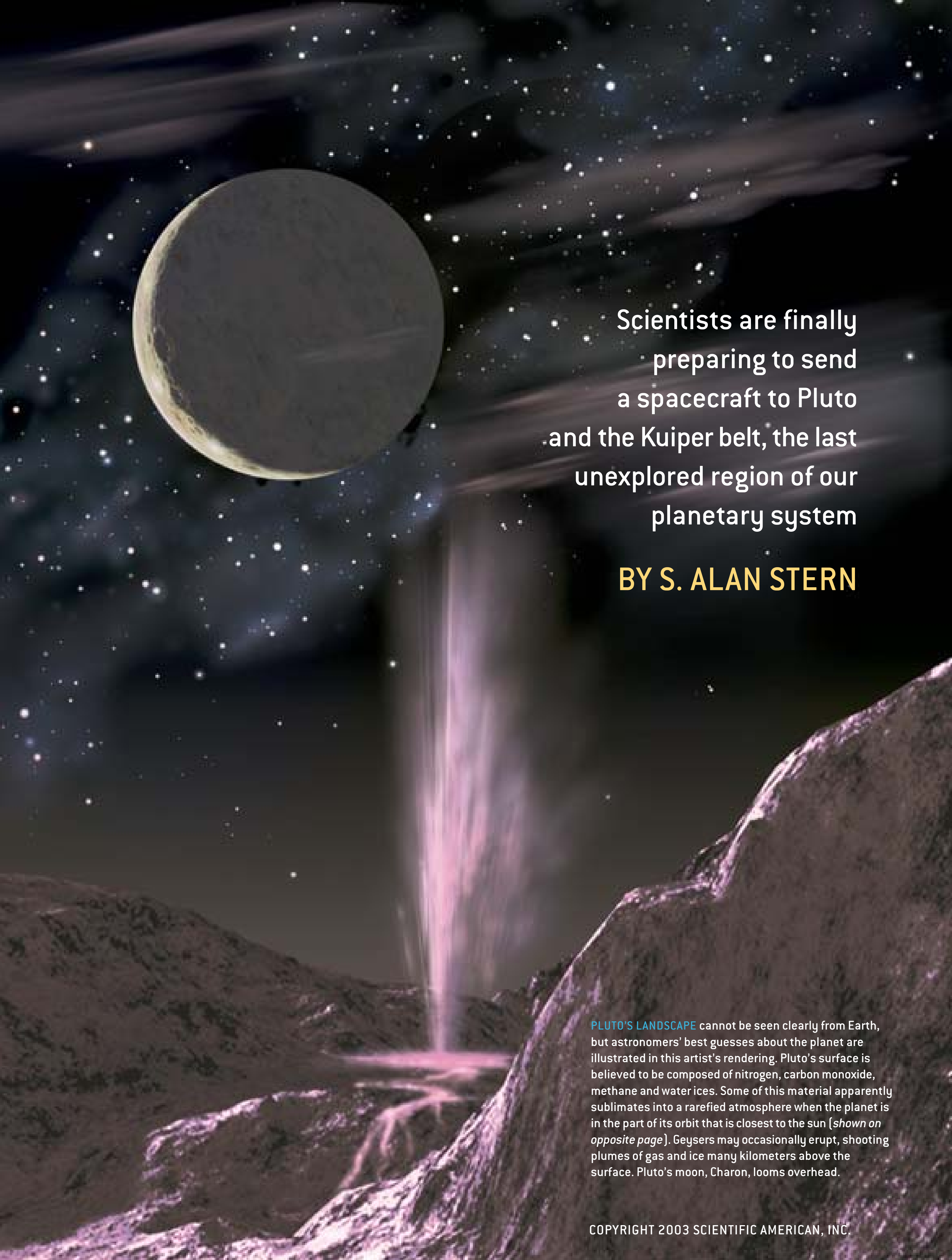
Planetary Rings. Joseph A. Burns in *The New Solar System.* Fourth edition. Edited by J. Kelly Beatty, Carolyn Collins Petersen and Andrew Chaikin. Cambridge University Press, 1999.

DON DIXON (*drawing*); NASA/JPL (*left*); FROM *NATURE,* VOL. 400, NO. 6746, 1999: B. SICARDY ET AL.,
CANADA-FRANCE-HAWAII TELESCOPE (*right*) AND C. DUMAS ET AL., HUBBLE SPACE TELESCOPE, JPL AND CALIFORNIA INSTITUTE OF TECHNOLOGY (*center*)

JOURNEY TO THE FARTHEST PLANET



Until about 10 years ago, most planetary scientists considered Pluto to be merely an oddity. All the other planets neatly fit into what astronomers knew about the architecture of the solar system—four small, rocky bodies in the inner orbits and four gas giants in the outer orbits, with an asteroid belt in between. But distant Pluto was an icy enigma traveling in a peculiar orbit beyond Neptune. Some researchers, most notably Dutch-American astronomer Gerard P. Kuiper, had suggested in the 1940s and



Scientists are finally preparing to send a spacecraft to Pluto and the Kuiper belt, the last unexplored region of our planetary system

BY S. ALAN STERN

PLUTO'S LANDSCAPE cannot be seen clearly from Earth, but astronomers' best guesses about the planet are illustrated in this artist's rendering. Pluto's surface is believed to be composed of nitrogen, carbon monoxide, methane and water ices. Some of this material apparently sublimates into a rarefied atmosphere when the planet is in the part of its orbit that is closest to the sun (*shown on opposite page*). Geysers may occasionally erupt, shooting plumes of gas and ice many kilometers above the surface. Pluto's moon, Charon, looms overhead.

1950s that perhaps Pluto was not a world without context but the brightest of a vast ensemble of objects orbiting in the same region. This concept, which came to be known as the Kuiper belt, rattled around in the scientific literature for decades. But repeated searches for this myriad population of frosty worlds came up empty-handed.

In the late 1980s, however, scientists determined that something like the Kuiper belt was needed to explain why many short-period comets orbit so close to the plane of the solar system. This circumstantial evidence for a distant belt of bodies orbiting in the same region as Pluto drove observers back to their telescopes in search of faint, undiscovered objects beyond Neptune. By the 1980s telescopes were being equipped with electronic light detectors that made searches far more sensitive than work done previously with photographic plates. As a result, success would come their way.

In 1992 astronomers at the Mauna Kea Observatory in Hawaii discovered the first Kuiper belt object (KBO), which was found to be about $1/10$ the size of Pluto and almost $1/10{,}000$ as bright [see "The Kuiper belt," by Jane X. Luu and David C. Jewitt; Scientific American, May 1996]. Since then, observers have found more than 700 KBOs, with diameters ranging from 50 to almost 1,200 kilometers. (Pluto's diameter is about 2,400 kilometers.)

And that's just the tip of the iceberg, so to speak. Extrapolating from the small fraction of the sky that has been surveyed so far, investigators estimate that the Kuiper belt contains approximately 100,000 objects larger than 100 kilometers across. As a result, the Kuiper belt has turned out to be the big brother to the asteroid belt, with far more mass, far more objects (especially of large sizes), and a greater supply of ancient, icy and organic material left over from the birth of the solar system.

It is thus clear that Pluto is not an anomaly. Instead it lies within a vast swarm of smaller bodies orbiting between about five billion and at least eight billion kilometers from the sun. Because this far-off region may hold important clues to the early development of the solar system, astronomers are keenly interested in learning more about Pluto, its moon, Charon, and the bodies making up the Kuiper belt. Unfortunately, the immense distance between this region of the solar system and Earth has limited the quality of observations. Even the exquisite Hubble Space Telescope, for example, shows only blurry regions of light and dark on Pluto's surface. And although the Pioneer, Voyager and Galileo spacecraft have provided scientists with marvelous close-up images of Jupiter, Saturn, Uranus and Neptune, no space probe has ever visited the Pluto-Charon system or the Kuiper belt.

Recognizing the importance of this region of the solar system, scientists urged NASA to put Pluto on its planetary exploration agenda for more than a decade. In response, the space agency studied mission concepts ranging from houseboat-size, instrument-laden spacecraft similar to the Cassini probe (now on its way to Saturn) to hamster-size craft carrying just a camera. In the late 1990s NASA settled on a midsize concept called Pluto-Kuiper Express that would be built by the Jet Propulsion Laboratory in Pasadena, Calif. But the projected cost of that mission quickly rose toward $800 million, which was considerably more than NASA wanted to invest. As a result, the agency scrapped Pluto-Kuiper Express in the fall of 2000.

But this cancellation didn't go down easily. Scientists, space exploration advocates and schoolchildren flooded NASA with requests to reconsider, and the agency did so, but with a twist. Rather than restarting the expensive Pluto-Kuiper Express, NASA launched a competition among universities, research labs and aerospace companies for proposals to explore Pluto, Charon and the Kuiper belt at lower cost. Never before had NASA allowed industry and universities to compete to lead a mission to the outer solar system. Given the novelty of such a competition, NASA made it clear that if none of the proposals could accomplish the specified scientific measurement objectives by 2020, and for less than $500 million, then the agency was under no obligation to choose *any* of them.

In November 2001, after a grueling selection process, NASA picked our team, called New Horizons, to carry out the Pluto–Kuiper belt mission. Early in 2003, Congress and the Bush administration approved funding in NASA's budget for the New Horizons Pluto mission, and the agency authorized the spacecraft's construction.

The New Horizons project is led by my institution, the Southwest Research Institute, based in San Antonio, Tex., and the Applied Physics Laboratory (APL) at Johns Hopkins University. A team of scientists from more than a dozen universities, research institutions and NASA centers is deeply involved in planning the scientific observations. The Southwest Research Institute is managing the project, is in charge of the mission team and is responsible for the development of the scientific instruments. APL is building and will operate the New Horizons spacecraft. Ball Aerospace, the

Overview/*New Horizons*

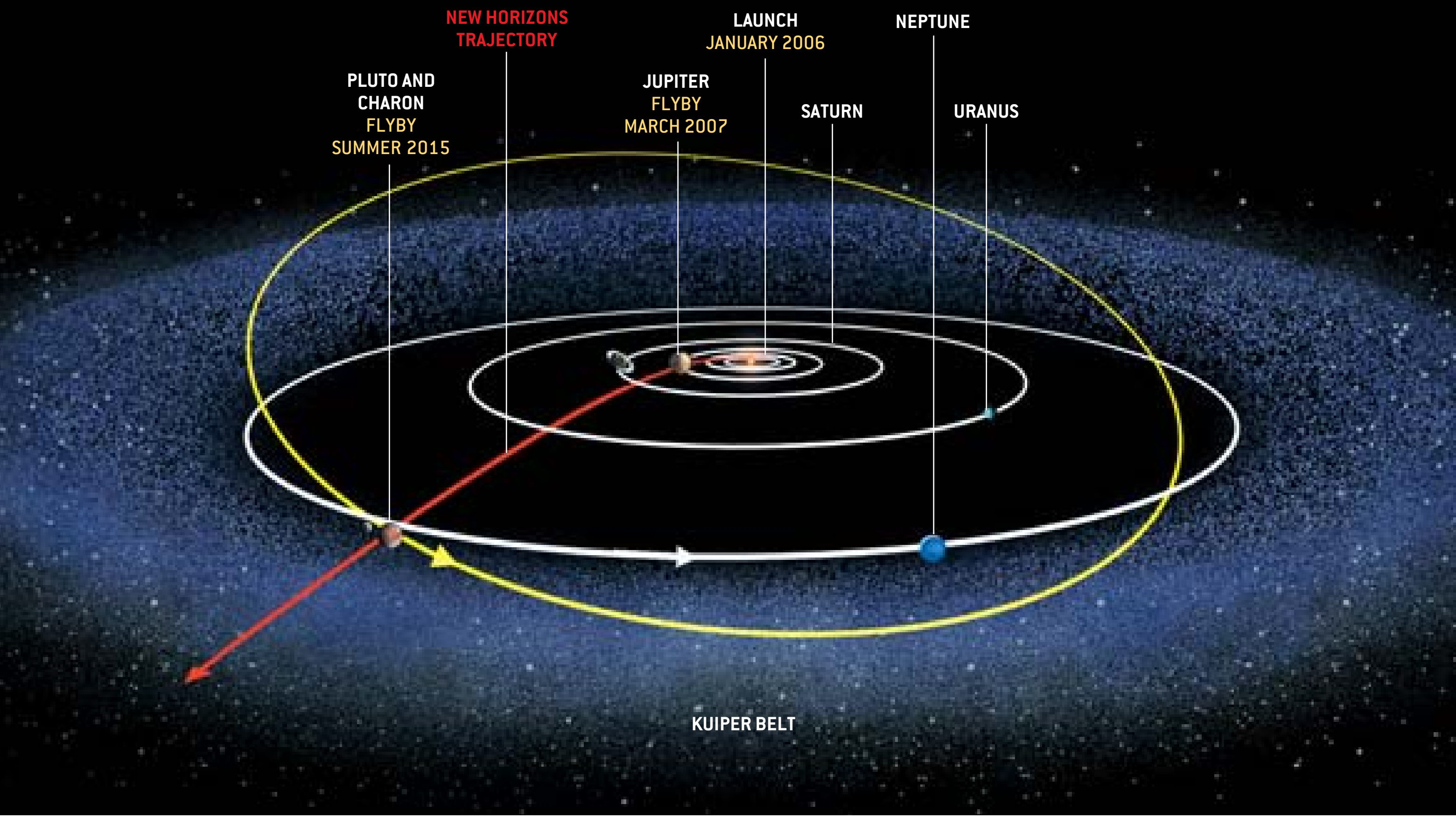
- Astronomers have recently learned that Pluto is not an anomaly, as once believed, but the brightest of a vast ensemble of objects orbiting in a distant region called the Kuiper belt. Scientists want to explore Pluto and the Kuiper belt objects because they may hold clues to the early history of the planets.
- Pluto and its moon, Charon, are also intriguing in their own right. The two bodies are so close in size that astronomers consider them a double planet. In addition, Pluto has a rapidly escaping atmosphere and complex seasonal patterns.
- NASA has chosen a team called New Horizons to build a spacecraft that will study Pluto, Charon and several Kuiper belt objects during a series of flyby encounters. The spacecraft will be launched in 2006 and will arrive at Pluto as early as 2015.

EDWIN FAUGHN (*preceding pages*)

Updated from the May 2002 issue

OUTWARD BOUND

THE JOURNEY TO PLUTO could take less than 10 years if the New Horizons spacecraft is launched in 2006. Traveling along the planned trajectory (*red line*), New Horizons would head initially for a Jupiter flyby that would use the planet's gravity to slingshot the craft toward Pluto (*yellow orbit*). After investigating Jupiter in 2007 and the Pluto-Charon system in 2015, the probe would go on to reconnoiter several of the icy bodies in the Kuiper belt.



DON DIXON

NASA Goddard Space Flight Center and Stanford University are building much of the instrument payload, and JPL will be responsible for spacecraft tracking and navigation.

By pioneering less expensive ways to build and operate a spacecraft to explore the outer solar system, New Horizons satisfied NASA's conditions: the total mission cost is less than $550 million, including more than $80 million in budgeted reserves, and the spacecraft may arrive at Pluto as early as the summer of 2015. Furthermore, New Horizons will fly more instruments and return about 10 times as much observational data as the canceled Pluto-Kuiper Express mission would have delivered and will do so for less money. During its journey, the spacecraft will also fly by and study Jupiter and its moons, and after flying by Pluto-Charon, the probe will go on to reconnoiter several KBOs at close range.

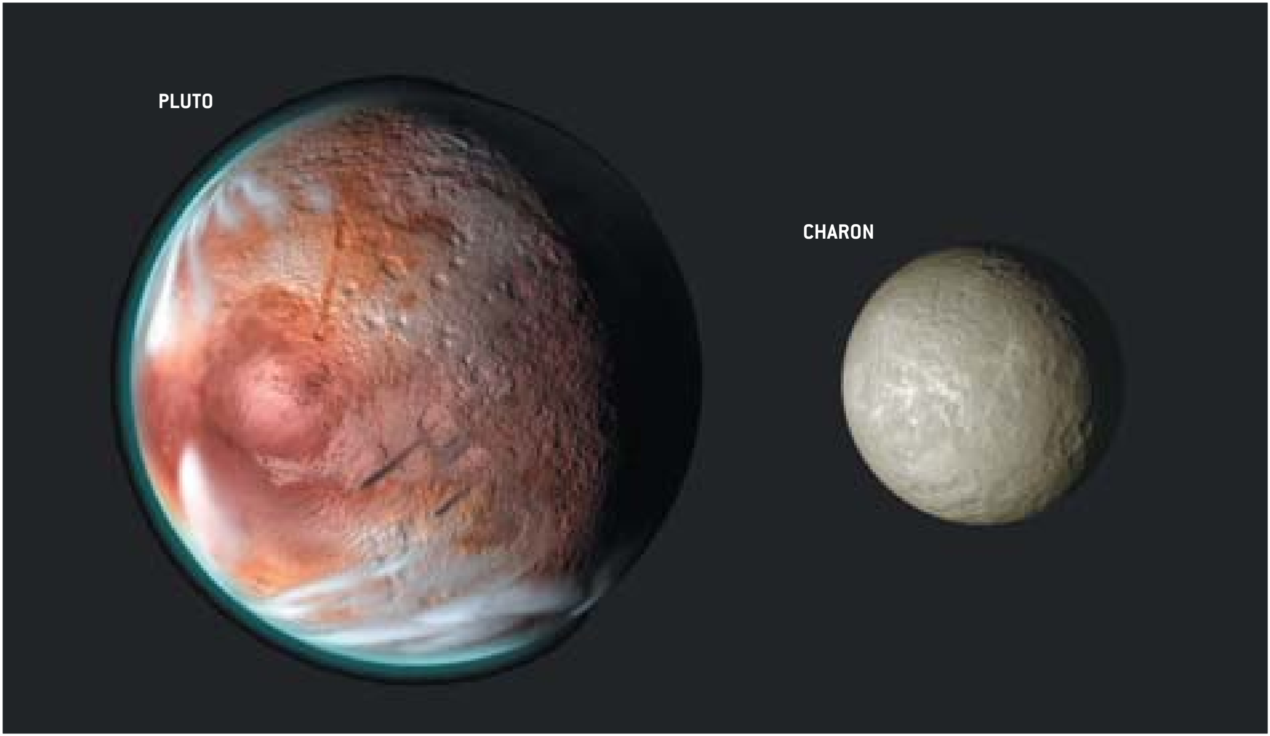
Archaeological Dig in Space

WHY ARE ASTRONOMERS so interested in studying Pluto-Charon and the Kuiper belt? I can summarize only a few of the reasons here. For one, the size, shape, mass and general nature of the Kuiper belt appear to be much like the debris belts seen around other nearby stars, such as Vega and Fomalhaut. A number of researchers, including myself, have used computer modeling techniques to simulate the formation of the KBOs more than four billion years ago as the planetary system was coalescing from a whirling disk of gas and dust. We found that the ancient Kuiper belt must have been approximately 100 times as massive as it is today to give rise to Pluto-Charon and the KBOs we see. In other words, there was once enough solid material to have formed another planet the size of Uranus or Neptune in the Kuiper belt.

The same simulations also revealed that large planets like Neptune would have naturally grown from the KBOs in a very short time had nothing disturbed the region. Clearly, something disrupted the Kuiper belt at about the time Pluto was formed, but we do not yet know the

DISTANT WORLDS

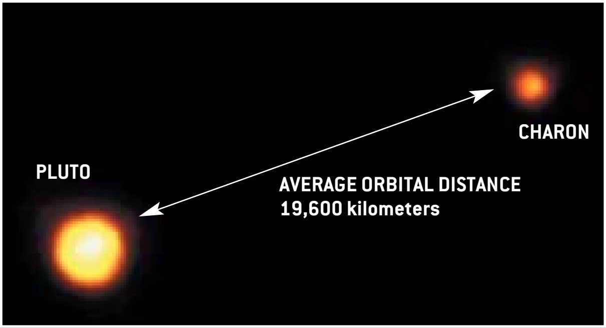
ASTRONOMERS ARE VERY INTERESTED in obtaining close-up views of Pluto and its moon, Charon, depicted here in an artist's rendering (*top*) based on the current knowledge of the two bodies. Because the Pluto-Charon system is so far from Earth, even the Hubble Space Telescope shows only blurry images of the two bodies (*bottom*).



The relative sizes of Pluto and Charon are drawn to scale, but the distance between them is not.

KEY FACTS

Diameter of Pluto: 2,400 kilometers
Diameter of Charon: 1,200 kilometers
Average distance of Pluto-Charon from sun: 5.9 billion kilometers
Orbital period around sun: 248 years
Average distance between Pluto and Charon: 19,600 kilometers
Orbital period of Charon around Pluto: 6.39 days
Rotation periods of Pluto and Charon: 6.39 days
Surface composition of Pluto: Nitrogen, carbon monoxide, methane and water ices
Surface composition of Charon: Water ice and possibly other compounds



DON DIXON (*artist's rendering*); NASA/ESA/ESO (*Hubble image*)

cause of the disturbance. Perhaps it was the formation of Neptune near the belt's inner boundary. Did the planet's gravitational influence somehow interrupt the creation of another gas giant farther out? And if so, why didn't the formation of Uranus frustrate the birth of Neptune in the same way? Perhaps instead it was the gravitational influence of a large population of planetary embryos—rocky bodies thousands of kilometers across—moving rapidly through the Kuiper belt billions of years ago after they were ejected by Uranus and Neptune from their formation zones. Or perhaps it was something else altogether. Whatever the cause, the Kuiper belt lost most of its mass, and the growth of bodies in the region was suddenly arrested.

The KBOs are remnants of that ancient planet-building process and therefore hold extremely important clues to the formation of the outer solar system. Exploring Pluto and the Kuiper belt is the equivalent of conducting an archaeological dig into the history of the outer solar system—a place where researchers can get a valuable glimpse of the long-gone era of planetary formation.

Furthermore, although our knowledge of Pluto and Charon is meager, what we do know indicates that they offer a scientific wonderland of their own.

For one, Charon is surprisingly large—its diameter is about 1,200 kilometers, or about half of Pluto's. Because the two bodies are so close in size, Pluto-Charon can be considered a double planet. No other planet in our solar system falls into this category—the diameters of most satellites are just a few percent of the diameters of their parent planets. But because astronomers have discovered many double asteroids and double KBOs in recent years, there is now little doubt that binary objects like Pluto-Charon are common in our solar system and most likely in others. Yet we have never visited a binary world.

We are eager to know how a system such as Pluto-Charon could form. The prevailing theory is that Pluto collided with another large body in the distant past and that much of the debris from this impact went into orbit around Pluto and eventually coalesced to form Charon. Because it appears that a similar collision led to the creation of Earth's moon, the study of Pluto and Charon is expected to shed some light on that subject as well.

Researchers also want to know why Pluto and Charon are so different in appearance. Observations from Earth and the Hubble Space Telescope show that Pluto has a highly reflective surface with distinct markings that indicate the presence of expansive polar caps. In contrast, Charon's surface is far less reflective, with indistinct markings. And whereas Pluto has an atmosphere, Charon apparently does not. Is the sharp dichotomy between these two neighboring worlds a result of divergent evolution, perhaps because of their different sizes and compositions, or is it a consequence of how they originally formed? We do not know.

Also intriguing is the fact that Pluto's density, size and surface composition are strikingly similar to those of Neptune's largest satellite, Triton. One of the great surprises of Voyager 2's exploration of the Neptune system was the discovery of ongoing, vigorous volcanic activity on Triton. Will Pluto also display such activity? Will the KBOs as well? Our present-day knowledge of planetary processes suggests that they should not, but Triton's activity was not expected either. Perhaps Triton is showing us that we do not yet understand the nature of small worlds. By exploring Pluto and the KBOs, we expect to gain a better comprehension of this fascinating class of bodies.

Yet another of Pluto's alluring features is its bizarre atmosphere. Although Pluto's atmosphere is about $^1/_{30,000}$ as dense as Earth's, it offers some unique insights into the workings of planetary atmospheres. Whereas Earth's atmosphere contains only one gas (water vapor) that regularly undergoes phase transitions between solid and gaseous states, Pluto's atmosphere contains three: nitrogen, carbon monoxide and methane. Furthermore, the current temperature on Pluto varies by about 50 percent across its surface—from about 40 to about 60 kelvins. Pluto reached its closest approach to the sun in 1989. As the planet moves farther away, most astronomers believe that the average surface temperature will drop and that most of the atmosphere will condense and fall as snow. Pluto may well have the most dramatic seasonal patterns of any planet in the solar system.

What is more, Pluto's atmosphere bleeds into space at a rate much like a comet's. Most of the molecules in the upper atmosphere have enough thermal energy to escape the planet's gravity; this extremely fast leakage is called hydrodynamic escape. Although this phenomenon is not seen on any other planet today, it may have been responsible for the rapid loss of hydrogen from Earth's atmosphere early in our planet's history. In this way, hydrodynamic escape may have helped make Earth suitable for life. Pluto is now the only planet in the solar system where this process can be studied.

An important connection between Pluto and the origin of life on Earth is the presence of organic compounds, such as frozen methane, on Pluto's surface and water ice in its interior. Recent observations of KBOs show that they, too, probably harbor large quantities of ice and organics. Billions of years ago such objects are believed to have routinely strayed into the inner part of the solar system and helped to seed the young Earth and Mars with the raw materials of life.

A Grand Tour Indeed

GIVEN SO MANY compelling scientific motivations, it is not hard to understand why the planetary research community wants to send a spacecraft to Pluto and the Kuiper belt. And given the romance and adventure of exploring uncharted worlds, it is not surprising that so many citizens and grade school children have also become excited about this mission to new frontiers.

NASA's request in 2001 for Pluto–Kuiper belt mission proposals specified three top priorities for scientific observations. First, the craft must map the surfaces of Pluto and Charon with an average resolution of one kilometer. (In contrast, the Hubble Space Telescope cannot do better than about 500-kilometer resolution when it views Pluto and Charon.) Second, the probe must map the surface composition across the various geologic provinces of the two bodies. Third, the craft must determine the composition and structure of Pluto's atmosphere, as well as its escape rate. NASA also outlined a list of lower priorities, including the measurement of surface temperatures and the search for additional satellites or rings around Pluto. And the agency required that the spacecraft accomplish the same objectives for at least one KBO beyond Pluto.

When NASA selected our proposal, it stated that the New Horizons mission offered "the best scientific return and the lowest risk of schedule delays and cost overruns." This was, in part, because of the robust capabilities of the spacecraft we proposed and the experience of our

THE AUTHOR

S. ALAN STERN is a planetary scientist and the principal investigator of NASA's New Horizons mission to Pluto and the Kuiper belt. He has participated in and led numerous space experiments and flies on board NASA F-18s and other high-performance aircraft to conduct high-altitude airborne astronomical research. Stern received his Ph.D. in planetary science and astrophysics from the University of Colorado in 1989. He is director of the Southwest Research Institute's department of space studies in Boulder, Colo.

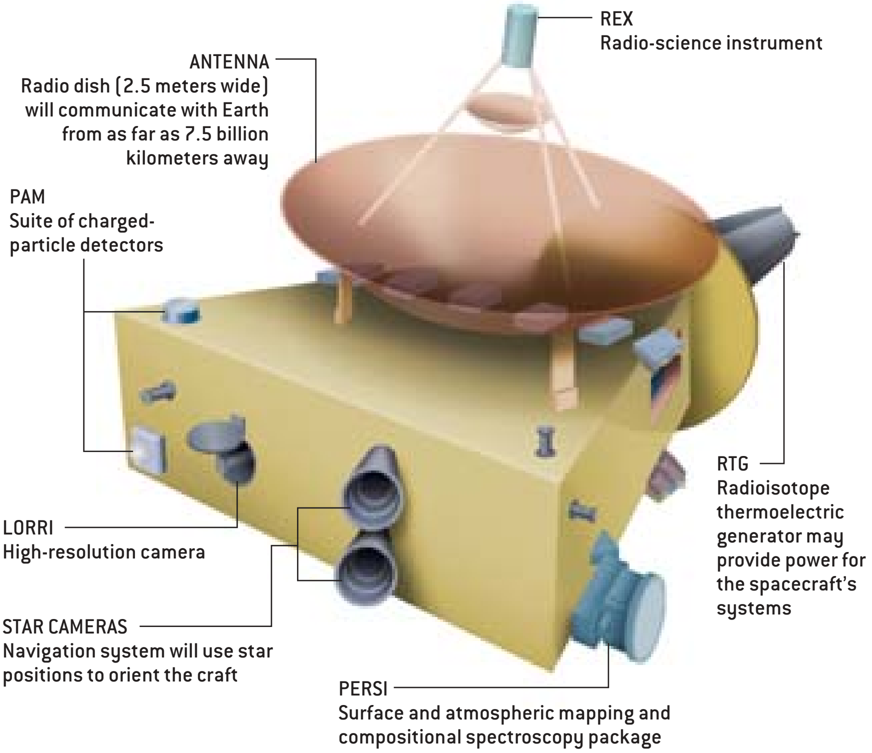
team-member institutions at delivering space missions on schedule and at or below cost.

The New Horizons spacecraft we designed is lean, with a planned mass of just 445 kilograms (981 pounds)—heavier than the early Pioneer probes but lighter than the Voyagers. This mass includes the hydrazine maneuvering propellant that will be used to adjust the craft's trajectory in flight. Most of the spacecraft's subsystems, such as its computers and its propulsion-control system, are based on designs used in other APL spacecraft. This reduces New Horizons's costs and lowers the risk of both technical and schedule problems. Almost all our spacecraft subsystems include spare equipment to increase reliability on the long flight to Pluto and the Kuiper belt.

The spacecraft will carry four instrument packages. A mapping and compositional spectroscopy package, PERSI, will make observations in the visible, ultraviolet and infrared parts of the spectrum. PERSI's infrared imaging spectrometer will be essential for mapping the composition and physical state (including temperature) of the surface ices on Pluto and Charon. A radio-science instrument dubbed REX will probe Pluto's atmospheric structure and gauge the average surface temperatures of Pluto and Charon (on both the daysides and nightsides of the bodies) by measuring the intensity of the microwave radiation striking the spacecraft's 2.5-meter-wide radio dish. A third instrument suite, PAM, consists of charged-particle detectors designed to sample material escaping from Pluto's atmosphere and to determine its escape rate. The fourth instrument is LORRI, a high-resolution imager that will supplement PERSI's already formidable mapping capabilities. At closest approach, PERSI's global maps of Pluto-Charon and the KBOs will have an average resolution of one kilometer. But LORRI, which will image selected regions, will be able to detect objects 1⁄20 that size. A small, student-built interplanetary dust collector has recently been added as well.

NEW HORIZONS SPACECRAFT

TO EXPLORE Pluto, Charon and the Kuiper belt objects, the proposed craft will carry four instrument packages, called REX, PAM, PERSI and LORRI, and an interplanetary dust collector (*not shown*).



JUAN VELASCO

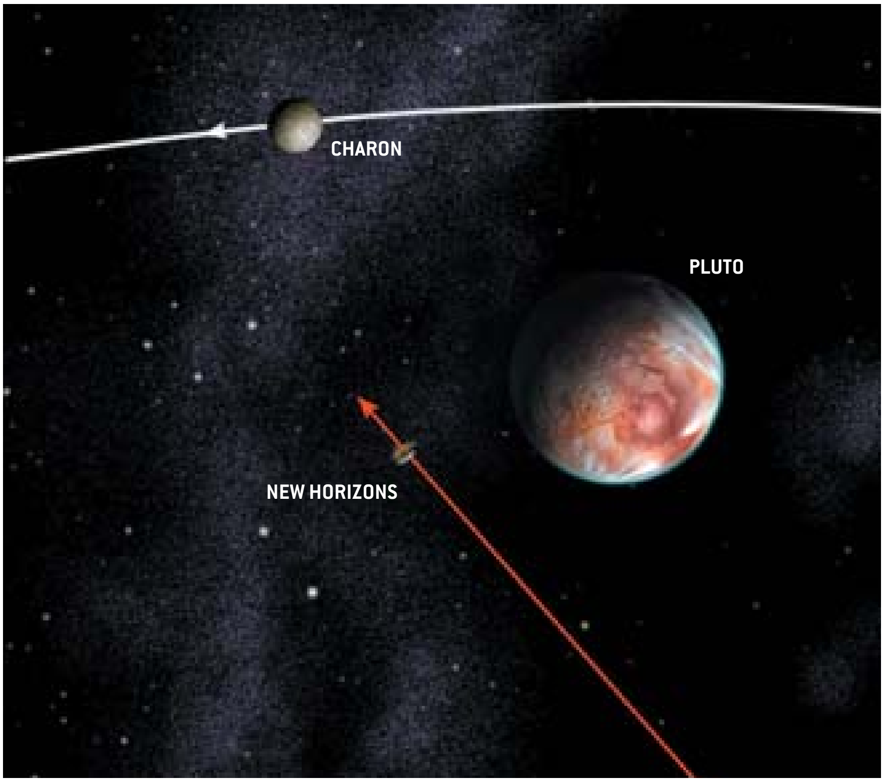
FLIGHT SPECS

- The spacecraft has a design mass of 445 kilograms (981 pounds) and is about the size of a small lifeboat.
- On the journey to Pluto, the probe will reach a top speed of about 70,000 kilometers an hour.
- The craft's computers will be able to store 48 gigabits of data and transmit the information to Earth at up to 770 bits a second from Pluto (16,000 bits a second from Jupiter).

If all goes as planned, the spacecraft will be launched in January 2006, heading initially for a flyby of Jupiter that will use the planet's gravity to slingshot the craft toward Pluto [*see illustration on page 87*]. During its Jupiter flyby, New Horizons will conduct an intensive four-month study of the planet's intriguing system of more than 20 moons, as well as its auroras, atmosphere and magnetosphere. Thanks to the gravitational assist from Jupiter, the spacecraft can reach the Pluto-Charon system as early as 2015. (The exact arrival date depends on the launch vehicle selected by NASA and the precise day we launch.)

For much of the long cruise from Jupiter to Pluto, New Horizons will slumber in electronic hibernation. Turning off unneeded systems and reducing the amount of contact with the craft lowers the chance of equipment failures and drastically decreases mission operations costs. During this hibernation the craft will continuously transmit a simple status beacon to Earth; if an unexpected problem develops, our ground-control team will respond. Once each year the craft will be awakened for about 50 days to thoroughly test the systems, make course corrections and calibrate its scientific instruments.

Unlike earlier plans for a quick flyby of Pluto-Charon, New Horizons will begin its study of Pluto-Charon six months before its closest approach to the planet. Once the craft is about 100 million kilometers from Pluto—about 75 days before closest approach—its images of the planet will be better than those of the Hubble Space Telescope, and the results will improve with each passing day. In



FLYBY OF PLUTO by the New Horizons spacecraft will reach its climax at closest approach, when the craft may come as near as a few thousand kilometers from the planet's surface. The flyby is shown from a perspective within Charon's orbit around Pluto and slightly above the orbital plane.

DON DIXON

the weeks before closest approach, our mission team will be able to map Pluto-Charon in increasing detail and observe phenomena such as Pluto's weather by comparing the images of the planet over time. And using LORRI's high-resolution imaging capabilities, we will get "zoom lens" views of Pluto and Charon that will help us decide which geologic features are worthy of special scrutiny. During the day of closest approach, when New Horizons may come as near as a few thousand kilometers from Pluto, PERSI will obtain its best maps of the entire sunlit faces of Pluto and Charon. Meanwhile LORRI will focus on producing higher-resolution maps of dozens of smaller areas on these bodies.

Once the spacecraft passes Pluto, it will turn around and map the planet's nightside, which will be softly illuminated by the reflected moonlight from Charon. And the spacecraft's antenna will receive a powerful radio beam from Earth passing through Pluto's atmosphere. By measuring the refraction of this radio beam, we will be able to plot the temperature and density profile of Pluto's atmosphere from high altitude down to the surface.

After the Pluto-Charon encounter, New Horizons will almost immediately maneuver to begin a series of what we hope will be one or more similar flybys with ancient KBOs over the next six years. The exact number of encounters will depend on how much propellant is left in the spacecraft after the Pluto flyby.

Now—or Never?

THE NEW HORIZONS mission promises to revolutionize our knowledge of both the Pluto-Charon system and the Kuiper belt. But the potential for discovery will be lost if the mission is not launched by 2007. A later departure would force the Pluto encounter to slip far into the future, beyond 2020, because of changing alignment of the planets.

By that time Pluto will be hundreds of millions of kilometers farther from the sun and significantly colder than it is today. Because of a combination of Pluto's extreme polar tilt and its motion around the sun, more than four million kilometers of terrain—much of the planet's southern hemisphere—will by then be covered in a dark polar shadow, thereby preventing it from being observed. Also, it is likely that virtually all the planet's atmosphere will have condensed by then, closing off any opportunity to study it until the 23rd century, when the atmosphere should again rise as the planet makes its next close approach to the sun.

New Horizons represents a thrilling return to first-time exploration for NASA's planetary program: for the first time since 1989, when Voyager 2 flew by Neptune, a spacecraft will train its instruments on a new world. The mission offers a scientific bonanza of proportions reminiscent of NASA's historic explorations. And by selecting New Horizons through competitive bidding, NASA has reduced costs to dimes on the dollar compared with recent missions to the outer solar system.

The scientific community and the public were successful in convincing Congress and the Bush administration to fund New Horizons. As a result, the exploration of Pluto-Charon and the Kuiper belt will commence with a series of flyby encounters beginning just over a dozen years from now, in the summer of 2015. In supporting this project, the U.S. will at last complete the basic reconnaissance of our solar system that it began in the 1960s with the historic Mariner missions to Venus and Mars.

MORE TO EXPLORE

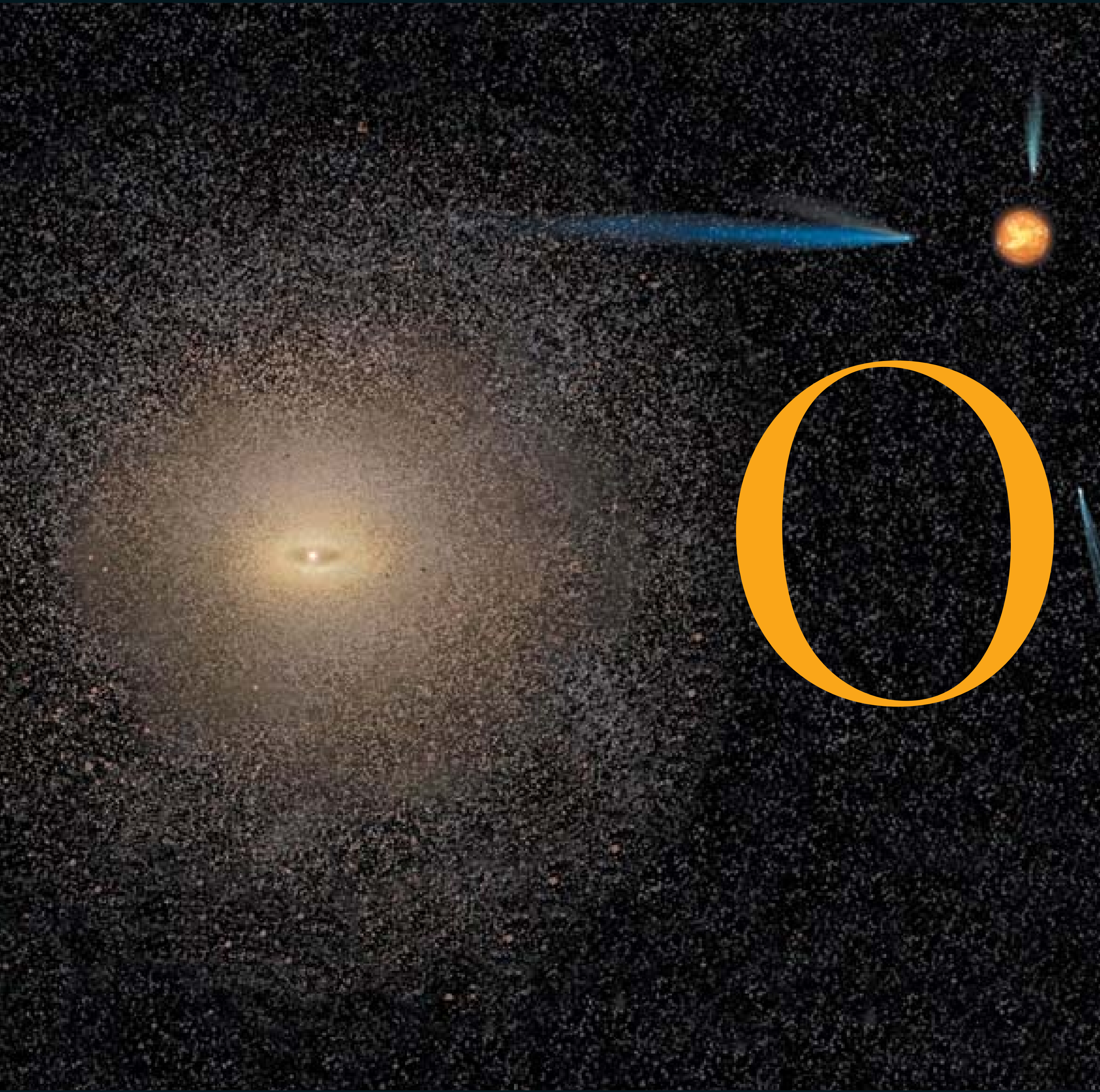
Pluto. Richard P. Binzel in *Scientific American,* Vol. 262, No. 6, pages 50–58; June 1990.

Pluto and Charon: Ice Worlds on the Ragged Edge of the Solar System. S. Alan Stern and Jacqueline Mitton. John Wiley & Sons, 1997.

Beyond Pluto: Exploring the Outer Limits of the Solar System. John Davies. Cambridge University Press, 2001.

Information about Pluto and Charon is available at **seds.lpl.arizona.edu/nineplanets/nineplanets/pluto.html**

Details of the New Horizons mission are available at **pluto.jhuapl.edu** and at **www.plutoportal.net**



O

CELESTIAL PIED PIPER, the red dwarf star Gliese 710, will crash through the Oort cloud in 1.4 million years—reanimating dormant comets, luring many out of their orbits and hurling some toward the planets. Such incursions, the result of haphazard stellar motions in our galaxy, occur every one million years on average. In this artist's conception, the distant comets are not to scale.

By Paul R. Weissman

the ORT cloud

On the outskirts of the solar system swarms a vast cloud of comets, influenced almost as much by other stars as by our sun. The dynamics of this cloud may help explain such matters as mass extinctions on Earth

It is common to think of the solar system as ending at the orbit of the most distant known planet, Pluto. But the sun's gravitational influence extends more than 3,000 times farther, halfway to the nearest stars. And that space is not empty—it is filled with a giant reservoir of comets, leftover material from the formation of the solar system. That reservoir is called the Oort cloud.

The Oort cloud is the Siberia of the solar system, a vast, cold frontier populated with exiles of the sun's inner empire and only barely under the sway of the central authority. Typical high temperatures are a frigid four degrees Celsius above absolute zero, and neighboring comets are typically tens of millions of kilometers apart. The sun, while still the brightest star in the sky, is only about as bright as Venus in the evening sky on Earth.

We have never actually "seen" the Oort cloud. But no one has ever seen an electron, either. We infer the existence and properties of the Oort cloud and the electron from the physical effects we can observe. In the case of the former, those effects are the steady trickle of long-period comets into the planetary system. The existence of the Oort cloud answers questions that people have asked since antiquity: What are comets, and where do they come from?

Aristotle speculated in the fourth century B.C. that comets were clouds of luminous gas high in Earth's atmosphere. But Roman philosopher Seneca suggested in the first century A.D. that they were heavenly bodies, traveling along their own paths through the firmament. Fifteen centuries passed before his hypothesis was confirmed by Danish astronomer Tycho Brahe, who compared observations of the comet of 1577 made from several different locations in Europe. If the comet had been close by, then from each location it would have had a slightly different position against the stars. Brahe could not detect any differences and concluded that the comet was farther away than the moon.

Just how much farther started to become clear only when astronomers began determining the comets' orbits. In 1705 English astronomer Edmond Halley compiled the first catalogue of 24 comets. The observations were fairly crude, and Halley could fit only rough parabolas to each comet's path. Nevertheless, he argued that the orbits might be very long ellipses around the sun:

> For so their Number will be determinate and, perhaps, not so very great. Besides, the Space between the Sun and the fix'd Stars is so immense that there is Room enough for a Comet to revolve, tho' the Period of its Revolution be vastly long.

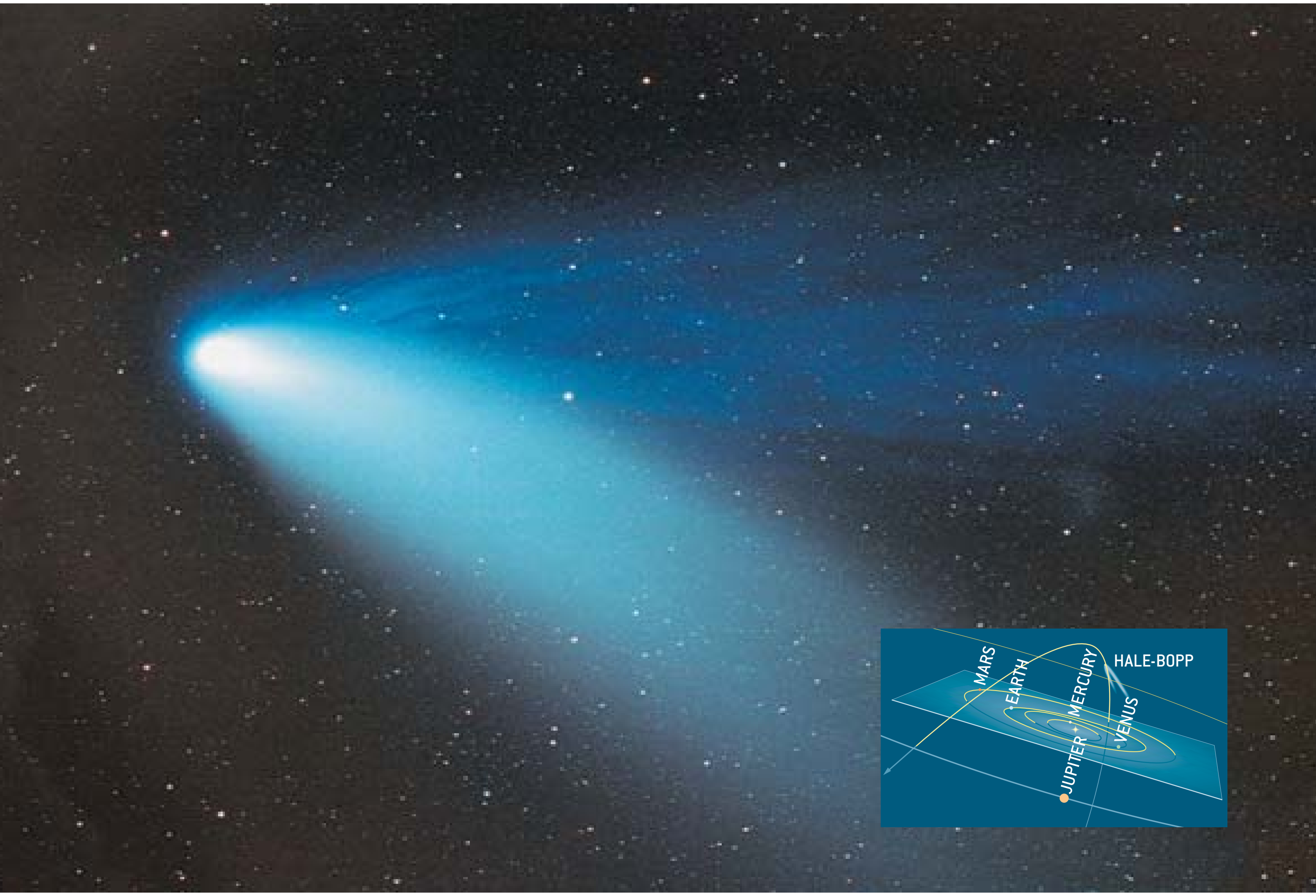
In a sense, Halley's description of comets circulating in orbits stretching between the stars anticipated the discovery of the Oort cloud two and a half centuries later. Halley also noticed that the comets of 1531, 1607 and 1682 had very similar orbits and were spaced at roughly 76-year intervals. These seemingly different comets, he suggested, were actually the same comet returning at regular intervals. That body, now known as Halley's Comet, last visited the region of the inner planets in 1986.

Since Halley's time, astronomers have divided comets into two groups according to the time it takes them to orbit the sun (which is directly related to the comets' average distance from the sun). Long-period comets, such as the recent bright comets Hyakutake and Hale-Bopp, have orbital periods greater than 200 years; short-period comets, less than 200 years. In the past decade astronomers have further divided the short-period comets into two groups: Jupiter-family comets, such as Encke and Tempel 2, which have periods less than 20 years; and intermediate-period, or Halley-type, comets, with periods between 20 and 200 years.

These definitions are somewhat arbitrary but reflect real differences. The intermediate- and long-period comets enter the planetary region randomly from all directions, whereas the Jupiter-family comets have orbits whose planes are typically inclined no more than 40 degrees from the ecliptic plane, the plane of Earth's orbit. (The orbits of the other planets are also very close to the ecliptic plane.) The intermediate- and long-period comets appear to come from the Oort cloud, whereas the Jupiter-family comets are thought to originate in the Kuiper belt, a region in the ecliptic beyond the

DON DIXON (*preceding pages*)

Updated from the September 1998 issue



COMET HALE-BOPP, which made its closest approach to Earth in March 1997, is an example of a long-period comet. It last visited the inner solar system 4,200 years ago and because of the gravitational influence of Jupiter will make its next appearance in 2,600 years. In the meantime, it will travel 370 times farther from the sun than Earth is. Like most long-period comets, Hale-Bopp has a highly inclined orbit: the plane of its orbit is nearly perpendicular to the plane of Earth's orbit (*inset*).

DENNIS DICICCO *Sky & Telescope*; MICHAEL GOODMAN (*inset*)

orbit of Neptune [see "The Kuiper Belt," by Jane X. Luu and David C. Jewitt; SCIENTIFIC AMERICAN, May 1996].

Netherworld beyond Pluto

BY THE EARLY 20th century, enough long-period cometary orbits were available to study their statistical distribution [*see box on next page*]. A problem emerged. About one third of all the osculating orbits—that is, the orbits the comets were following at the point of their closest approach to the sun—were hyperbolic. Hyperbolic orbits would originate in and return to interstellar space, as opposed to elliptical orbits, which are bound by gravity to the sun. The hyperbolic orbits led some astronomers to suggest that comets were captured from interstellar space by encounters with the planets.

To examine this hypothesis, celestial-mechanics researchers extrapolated, or integrated, the orbits of the long-period comets backward in time. They found that because of distant gravitational tugs from the planets, the osculating orbits did not represent the comets' original orbits [*see illustration on page 97*]. When the effects of the planets were accounted for—by integrating far enough back in time and orienting the orbits not in relation to the sun but in relation to the center of mass of the solar system (the sum of the sun and all the planets)—almost all the orbits became elliptical. Thus, the comets were members of the solar system, rather than interstellar vagabonds.

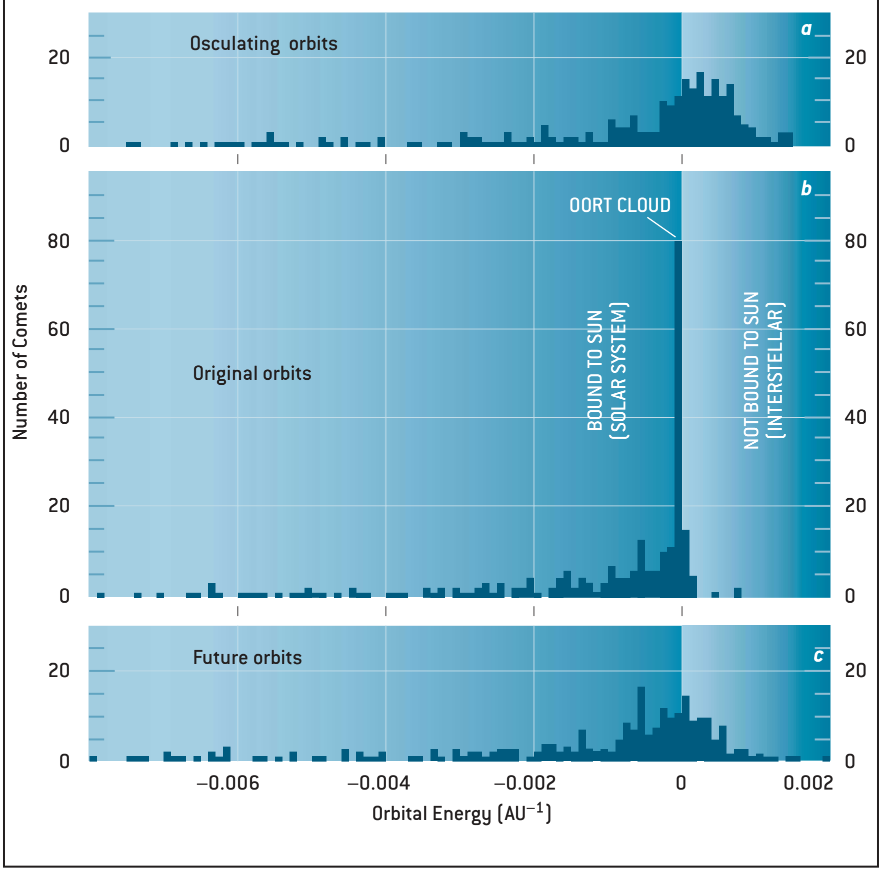
In addition, although two thirds of these orbits still appeared to be uniformly distributed, fully one third had orbital energies that fell within a narrow

THE AUTHOR

PAUL R. WEISSMAN is a senior research scientist at the Jet Propulsion Laboratory in Pasadena, Calif., where he specializes in studies of the physics and dynamics of comets. He is also an interdisciplinary scientist on the Rosetta comet rendezvous mission, which will be launched in February 2004 to Comet Churyumov-Gerasimenko. Weissman has written more than 90 refereed scientific papers and is one of three editors of the *Encyclopedia of the Solar System,* published by Academic Press in 1999.

COMET ORBITS REVEAL THE CLOUD

ORBITAL ENERGY DISTRIBUTION of known long-period comets, shown below, reveals the Oort cloud. Astronomers first calculate the osculating orbits of the comets—the orbits they would take if their motion were entirely caused by the sun's gravity. One third of these orbits have a positive energy, making them appear interstellar (*a*). But when corrected for the influence of the planets and extrapolated backward in time, the energy is slightly negative—indicating that the comets came from the edge of the solar system (*b*). A few comets still seem to be interstellar, but this is probably the result of small observational errors. As the planets continue to exert their influence, some comets will return to the Oort cloud, some will escape from the solar system and the rest will revisit the inner solar system (*c*). Technically, the orbital energy is proportional to the reciprocal of the semimajor axis, expressed in units of inverse astronomical units (AU^{-1}). (One AU is the distance from Earth to the sun.)



PAUL R. WEISSMAN

spike. That spike represented orbits that extend to very large distances—20,000 astronomical units (20,000 times the distance of Earth from the sun) or more. Such orbits have periods exceeding one million years.

Why were so many comets coming from so far away? In the late 1940s Dutch astronomer Adrianus F. van Woerkom showed that the uniform distribution could be explained by planetary perturbations, which scatter comets randomly to both larger and smaller orbits. But what about the large proportion of comets with million-year periods?

In 1950 Dutch astronomer Jan H. Oort, already famous for having determined the rotation of the Milky Way galaxy in the 1920s, became interested in the problem. He recognized that the million-year spike must represent the source of the long-period comets: a vast spherical cloud surrounding the planetary system and extending halfway to the nearest stars.

Oort showed that the comets in this cloud are so weakly bound to the sun that random passing stars can readily change their orbits. About a dozen stars pass within one parsec (206,265 astronomical units) of the sun every one million years. These close encounters are enough to stir the cometary orbits, randomizing their inclinations and sending a steady trickle of comets into the inner solar system on very long elliptical orbits [*see illustration on page 99*]. As they enter the planetary system for the first time, the comets are scattered by the planets, gaining or losing orbital energy. Some escape the solar system altogether. The remainder return and are observed again as members of the uniform distribution. Oort described the cloud as "a garden, gently raked by stellar perturbations."

A few comets still appeared to come from interstellar space. But this was probably an incorrect impression given by small errors in the determination of their orbits. Moreover, comets can shift their orbits because jets of gas and dust from their icy surfaces act like small rocket engines as the comets approach the sun. Such nongravitational forces can make the orbits appear hyperbolic when they are actually elliptical.

Shaken, Not Stirred

OORT'S ACCOMPLISHMENT in correctly interpreting the orbital distribution of the long-period comets is even more impressive when one considers that he had only 19 well-measured orbits to work with. Today astronomers have more than 15 times as many. They now know that long-period comets entering the planetary region for the first time come from an average distance of 44,000 astronomical units. Such orbits have periods of 3.3 million years.

Astronomers have also realized that stellar perturbations are not always gentle. Occasionally a star comes so close to the sun that it passes right through the Oort cloud, violently disrupting the cometary orbits along its path. Statistically a star is expected to pass within 10,000 astronomical units of the sun every 36 million years and within 3,000 astronomical units every 400 million years. Comets close to the star's path are thrown out to interstellar space, while the orbits of comets throughout the cloud undergo substantial changes.

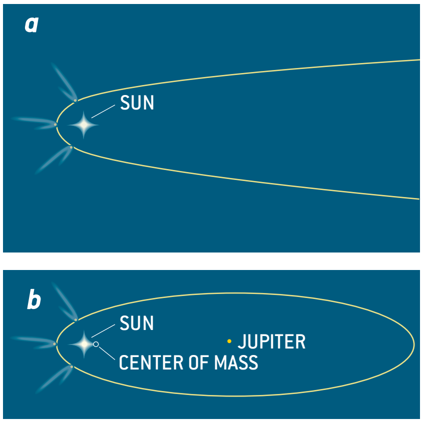
Although close stellar encounters have no direct effect on the planets—the closest expected approach of any star over the history of the solar system is 900 astronomical units from the sun—they might have devastating indirect consequences. In 1981 Jack G. Hills of Los Alamos National Laboratory suggested that a close stellar passage could send a "shower" of comets toward the planets, raising the rate of cometary impacts on the planets and possibly even causing a biological mass extinction on Earth. According to computer simulations I performed in 1985 with Piet Hut of the Institute for Advanced Study in Princeton, N.J., the frequency of comet passages during a shower could reach 300 times the normal rate. The shower would last two million to three million years.

In the late 1990s Kenneth A. Farley and his colleagues at the California Institute of Technology found evidence for just such a comet shower. Using the rare helium 3 isotope as a marker for extraterrestrial material, they plotted the accumulation of interplanetary dust particles in ocean sediments over time. The rate of dust accumulation is thought to reflect the number of comets passing through the planetary region; each comet sheds dust along its path. Farley discovered that this rate increased sharply at the end of the Eocene epoch, about 36 million years ago, and decreased slowly over two million to three million years, just as theoretical models of comet showers would predict. The late Eocene is identified with a moderate biological extinction event, and several impact craters have been dated to this time. Geologists have also found other traces of impacts in terrestrial sediments, such as iridium layers and microtektites.

Is Earth in danger of a comet shower now? Fortunately not. Joan Garcia-Sanchez of the University of Barcelona, Robert A. Preston and Dayton L. Jones of the Jet Propulsion Laboratory in Pasadena, Calif., and I have been using the positions and velocities of stars, measured by the Hipparcos satellite, to reconstruct the trajectories of stars near the solar system. We have found no evidence that a star has passed close to the sun in the past one million years. The next close passage of a star will occur in 1.4 million years, and that is a small red dwarf called Gliese 710, which will pass through the outer Oort cloud about 70,000 astronomical units from the sun. At that distance, Gliese 710 might increase the frequency of comet passages through the inner solar system by 25 percent or less—a sprinkle perhaps, but certainly no shower.

In addition to random passing stars, the Oort cloud is now known to be disturbed by two other effects. First, the cloud is sufficiently large that it feels tidal forces generated by the disk of the Milky Way and, to a lesser extent, the galactic core. These tides arise because the sun and a comet in the cloud are at slightly different distances from the midplane of the disk or from the galactic center and thus feel a slightly different gravitational tug [*see illustration on page 100*]. The tides help to feed new long-period comets into the planetary region.

Second, giant molecular clouds in the galaxy can perturb the Oort cloud, as Ludwig Biermann of the Max Planck Institute for Physics and Astrophysics in Munich suggested in 1978. These massive clouds of cold hydrogen, the birthplaces of stars and planetary systems, are 100,000 to one million times as massive as the sun. When the solar system comes close to one, the gravitational perturbations rip comets from their orbits and fling them into interstellar space. These encounters, though violent, are infrequent—only once every 300 million to 500 million years. In 1985 Hut and Scott D. Tremaine, now at Princeton University, showed that over the history of the solar system, molecular clouds have had about the same cumulative effect as all passing stars.



LONG-PERIOD COMET is so weakly bound to the sun that the planets have a decisive influence on it. Astronomers can usually see the comet only while it swings by the sun. When they apply Kepler's laws of celestial motion to plot its course—its osculating, or apparent, orbit—the comet often seems to be on a hyperbolic trajectory, implying that it came from interstellar space and will return there (*a*). A more sophisticated calculation, which accounts for the planets (especially the most massive planet, Jupiter), finds that the orbit is actually elliptical (*b*). The orbit changes shape on each pass through the inner solar system.

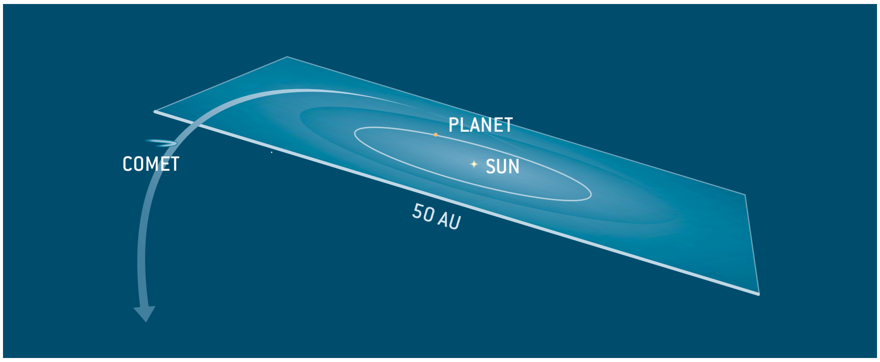
MICHAEL GOODMAN

Inner Core

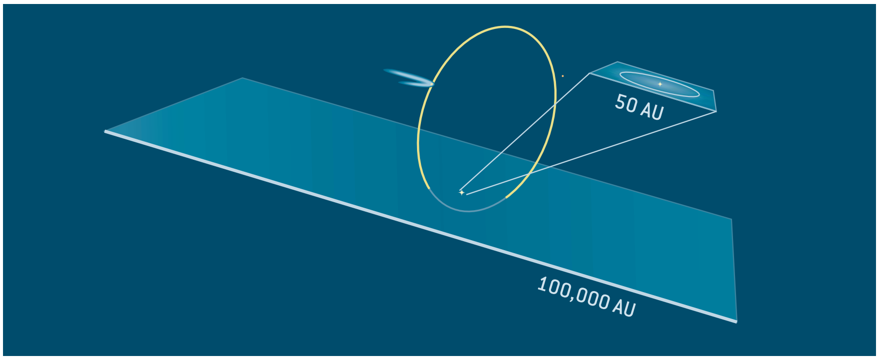
CURRENTLY THREE main questions concern Oort cloud researchers. First, what is the cloud's structure? In 1987 Tremaine, Martin J. Duncan, now at Queen's University in Ontario, and Thomas R. Quinn, now at the University of Washington, studied how stellar and galactic perturbations redistribute comets within the Oort cloud. Comets at its outer edge are rapidly lost, either to interstellar space or to the inner solar system, because of the perturbations. But deeper inside, there probably exists a dense inner core of more tightly bound comets that slowly replenishes the outer reaches.

Duncan, Quinn and Tremaine also showed that as comets fall in from the Oort cloud, their orbital inclinations tend not to change. This is a major reason why astronomers think the Kuiper belt, rather than the Oort cloud, accounts for the low-inclination, Jupiter-family comets. Still, the Oort cloud is the most likely source of the higher-inclination, intermediate-period comets, such as Halley and Swift-Tuttle. They were probably once long-period comets that the planets pulled into shorter-period orbits.

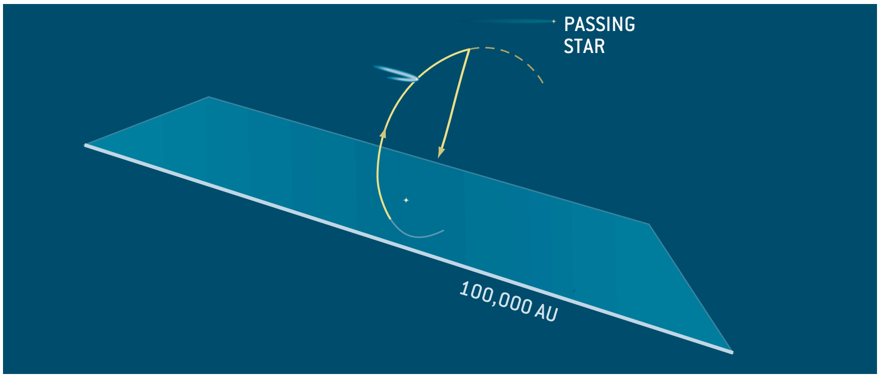
The second main question is: How many comets inhabit the Oort cloud? The number depends on how fast comets leak from the cloud into interplanetary space. To account for the observed



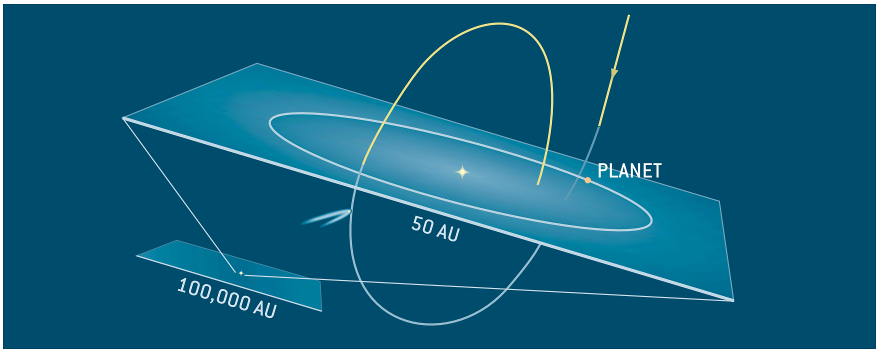
HISTORY OF A LONG-PERIOD COMET begins when it forms in an orbit among the giant planets and is catapulted by them into a wide (larger) orbit.



There the comet is susceptible to the gravitational forces of random passing stars and giant molecular clouds, as well as the tidal forces of the galactic disk and core. These forces randomly tilt the orbital plane of the comet and gradually pull it farther out.



Beyond a distance of about 20,000 astronomical units (20,000 times the Earth-sun distance), the various outside influences are capable of throwing the comet back toward the planets.



Once the comet reenters the inner solar system, the planets may pull it to a new orbit, so that it reappears on a regular basis.

MICHAEL GOODMAN

number of long-period comets, astronomers estimate the cloud has two trillion to five trillion comets, making Oort cloud comets the most abundant substantial bodies in the solar system. Only one fifth to one half of them are in the outer, dynamically active cloud first described by Oort; the rest are in the unseen central core. If the best estimate for the average mass of a comet—about 40 billion metric tons—is applied, the total mass of comets in the Oort cloud at present is about 13 to 34 times that of Earth.

Finally, from where did the Oort cloud comets originally come? They could not have formed at their current position, because material at those distances is too sparse to coalesce. Nor could they have originated in interstellar space; capture of comets by the sun is very inefficient. The only place left is the planetary system. Oort speculated that the comets were created in the asteroid belt and ejected by the giant planets during the formation of the solar system. But comets are icy bodies, essentially big, dirty snowballs, and the asteroid belt was too warm for ices to condense.

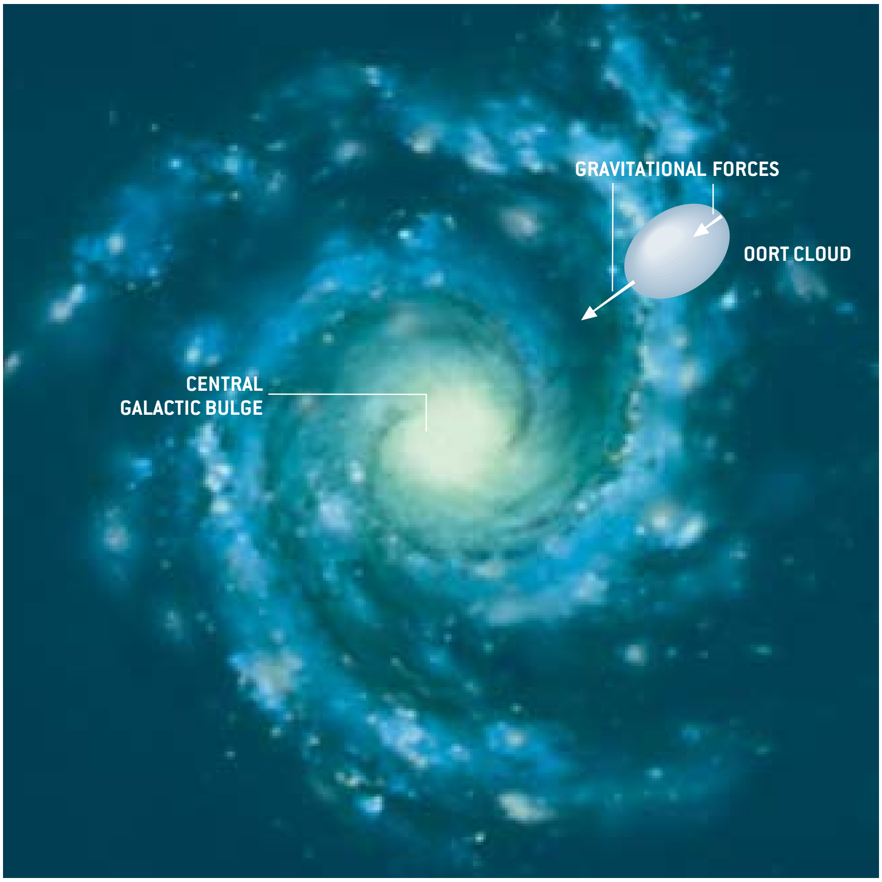
A year after Oort's 1950 paper, astronomer Gerard P. Kuiper of the University of Chicago proposed that comets coalesced farther from the sun, among the giant planets. (The Kuiper belt is named for him because he suggested that some comets also formed beyond the farthest planetary orbits and have remained there over the history of the solar system.) Comets probably originated throughout the giant planets' region, but researchers used to argue that those near Jupiter and Saturn, the two most massive planets, would have been ejected to interstellar space rather than to the Oort cloud. Uranus and Neptune, with their lower masses, could not easily throw so many comets onto escape trajectories. More recent dynamical studies have cast some doubt on this scenario. Jupiter and particularly Saturn do place a significant fraction of their comets into the Oort cloud, as well as many Uranus-Neptune comets that wander into Saturn-crossing orbits. So the populating of the Oort cloud is really a complex game of billiards among the giant planets, with all of them contributing to the process.

Therefore, the Oort cloud comets may have come from a wide range of solar distances and hence a wide range of formation temperatures. This fact may help explain some of the compositional diversity observed in comets. Indeed, recent work I have done with Harold F. Levison of the Southwest Research Institute in Boulder, Colo., has shown that the cloud may even contain asteroids from the inner planets' region. These objects, made of rock rather than ice, may constitute 2 to 3 percent of the total Oort cloud population.

The key to these ideas is the presence of the giant planets, which hurl the comets outward and modify their orbits if they ever reenter the planetary region. If other stars have giant planets, as observations over the past few years suggest, they may have Oort clouds, too. If each star has its own cloud, then as stars pass by the sun, their Oort clouds will pass through our cloud. Even so, collisions between comets will be rare because the typical space between comets is an astronomical unit or more.

The Oort clouds around each star may slowly be leaking comets into interstellar space. These interstellar comets should be easily recognizable if they were to pass close to the sun, because they would approach the solar system at much higher velocities than the comets from our own Oort cloud. To date, no such interstellar comets have ever been detected. This fact is not surprising: because the solar system is a very small target in the vastness of interstellar space, there is at best a 50–50 chance that astronomers should have seen one interstellar comet by now.

The Oort cloud continues to fascinate astronomers. Through the good fortunes of celestial mechanics, nature has preserved a sample of material from the formation of the solar system in this distant reservoir. By studying it and the cosmo-chemical record frozen in its icy members, researchers are learning valuable clues about the origin of the solar system.



TIDAL FORCES arise because gravity becomes weaker with distance. Therefore, the central bulge of our galaxy—a concentration of stars at the hub of the spiral pattern—pulls more on the near side of the Oort cloud (*not to scale*) than on the far side. The galactic disk exerts an even stronger force but in a perpendicular direction. The galactic tides are analogous to lunar tides, which arise because the side of Earth closest to the moon feels a stronger gravitational pull than the antipode does.

MICHAEL GOODMAN

Several space missions are now in flight or being readied to unlock those secrets. The Stardust spacecraft was launched in 1999 and will fly through the coma of Comet Wild 2 in January 2004. As it does, it will collect samples of cometary dust and return them to Earth in 2006 for laboratory studies. The Deep Impact mission will be launched in December 2004 and will send a 370-kilogram impactor crashing into the nucleus of Comet Tempel 1 in July 2005; the impactor will investigate the interior structure of the nucleus. The Rosetta mission will be launched in February 2004 and will fly a convoluted path through the solar system before finally going into orbit around the nucleus of Comet Churyumov-Gerasimenko in 2014. And NASA has just asked space scientists to propose plans for a comet-sample-return mission to be launched around 2009, which would bring back samples from a comet nucleus sometime during the next decade. The new millennium is turning into a wonderful time for studying comets. SA

MORE TO EXPLORE

Comets: A Chronological History of Observation, Science, Myth and Folklore. Donald K. Yeomans. John Wiley & Sons, 1991.

Comets in the Post-Halley Era. Edited by Ray L. Newburn, Jr., Marcia Neugebauer and Jurgen Rahe. Kluwer Academic Publishers, 1991.

Dynamics of Comets: Recent Developments and New Challenges. Julio A. Fernandez in *Asteroids, Comets, Meteors 1993: Proceedings of the 160th Symposium of the International Astronomical Union.* Edited by A. Milani, M. Di Martino and A. Cellino. Kluwer Academic Publishers, 1994.

Completing the Inventory of the Solar System. Edited by Terrence W. Rettig and Joseph M. Hahn. Astronomical Society of the Pacific Conference Series, Vol. 107; 1996.